6. Konzentration

6.1 Konzentrationsbegriff

Konzentration im wirtschaftlichen Sinne:

- Ballung von Verfügungsmacht bei einer oder wenigen Wirtschaftseinheiten (Anzahlaspekt)
- Existenz erheblicher Größenunterschiede (Disparität)

Beispiele:

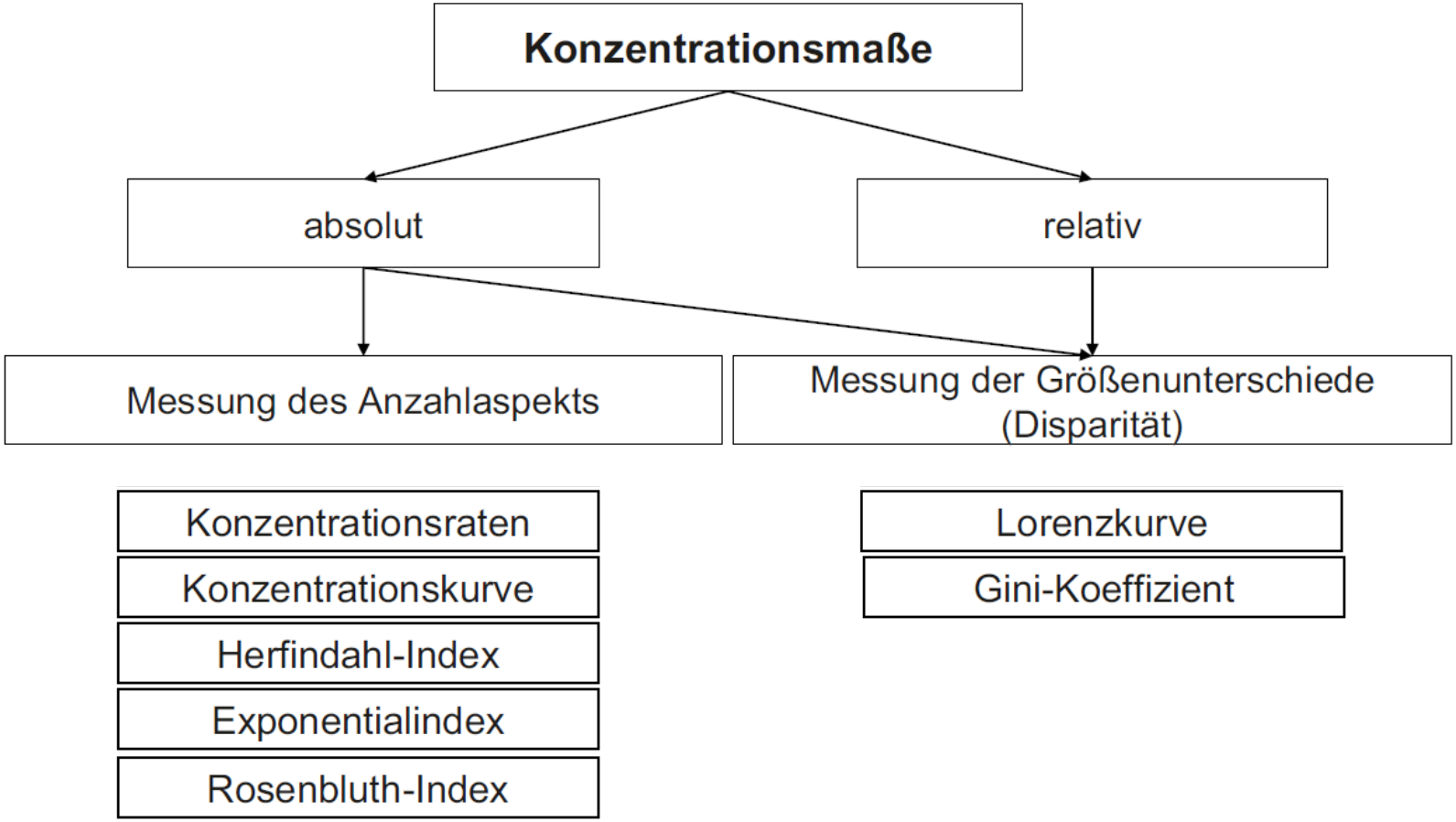
- Sind in einem Markt nur wenige Anbieter vorhanden, dann liegt eine Konzentration vor (Anzahlaspekt).
- Auch bei vielen Anbietern würde man von einer Konzentration sprechen, wenn die Umsatzanteile der Unternehmen in den Größenklassen sehr unterschiedlich sind (Disparität).

Extremsituationen:

1. **Egalitäre Verteilung („Gleichverteilung“):** Situation der minimalen Konzentration, d.h. jeder Merkmalsträger weist den gleichen Merkmalsbetrag und damit auch den gleichen Merkmalsanteil auf.

2. **Vollkommene Ungleichheit:** Situation der maximalen Konzentration, d.h. ein Merkmalsträger vereinigt die gesamte Merkmalssumme auf sich, während die übrigen statistischen Einheiten „Nullträger“ sind.

Konzentrationsmaße: Relative und absolute Konzentration



6.2 Konzentrationsraten und Konzentrationskurve

Konzentrationsraten:

- Maße der absoluten Konzentration
- Messung der Wettbewerbsintensität und Konzentrationstendenzen in der Wirtschaft

Gesetz gegen Wettbewerbsbeschränkungen (GWB), §18:

Kritische Konzentrationsraten (Vermutung einer marktbeherrschenden Stellung)

Anzahl der Unternehmen	… mit einer Konzentrationsrate von mindestens …
1	40%
3 oder weniger	50%
5 oder weniger	2/3

Bestimmung von Konzentrationsraten (I)

Konzentrationsmerkmal X (z.B. Umsatz, Absatz, Beschäftigte, Börsenwert) in Form von Einzelbeobachtungen

- Merkmalsbeträge: $x_1, x_2, \cdots, x_n$
- Geordnete Reihe (in absteigender Reihenfolge): $x^{(1)} \geq x^{(2)} \geq \cdots \geq x^{(n)}$
- Merkmalssumme:

$$S = \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x^{(i)}$$

- Merkmalsanteil des i-ten Merkmalsträgers:

$$c_i = \frac{x^{(i)}}{S}$$

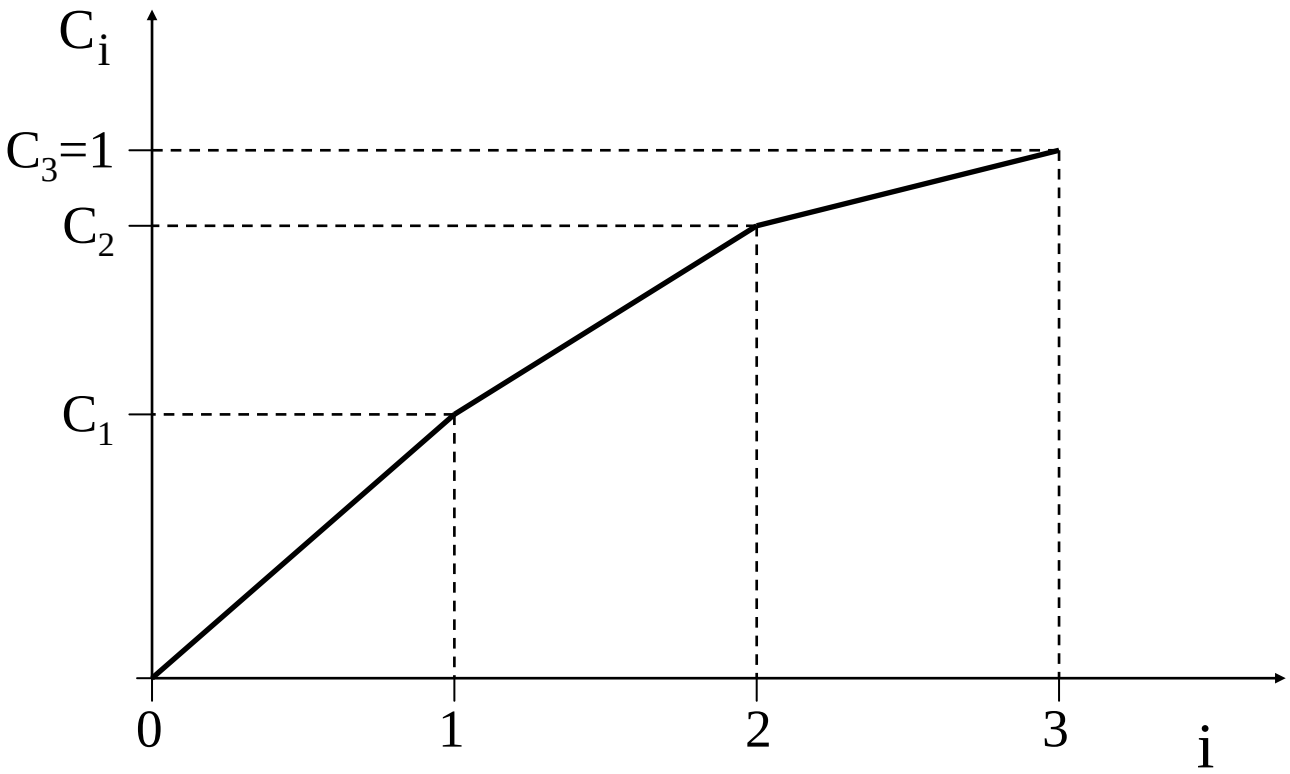
- Konzentrationsraten (kumulierte Merkmalsanteile):

$$C_i = \sum_{j=1}^{i} c_j = C_{i-1} + c_i$$

Bestimmung von Konzentrationsraten (II)

- Eigenschaft (ohne „Nullträger“): $0 < C_1 < C_2 < \cdots < C_n$
- Die **Konzentrationsrate C_i** gibt an, welchen Merkmalsanteil die i „größten“ Unternehmen auf sich vereinigen.
- **Konzentrationskurve**:

  Verbindungslinie der Punkte (0, 0), (1, C_1), (2, C_2), ... in einem kartesischen Koordinatensystem



Minimale Konzentration versus maximale Konzentration (I)

- Minimale Konzentration:
  - Alle Merkmalsträger haben den gleichen Merkmalsbetrag x_i und den gleichen Merkmalsanteil

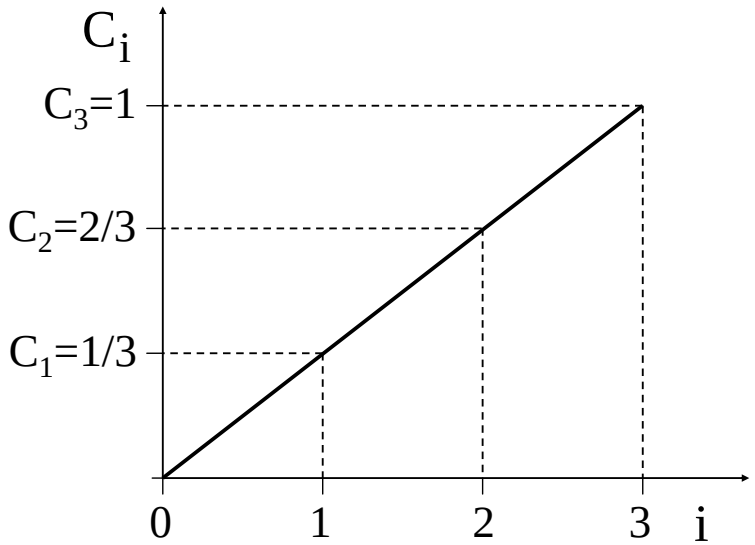
$$c_i = \frac{1}{n}, i = 1, 2, \cdots, n$$

  - Konzentrationskurve: Gerade mit Steigung $\frac{1}{n}$

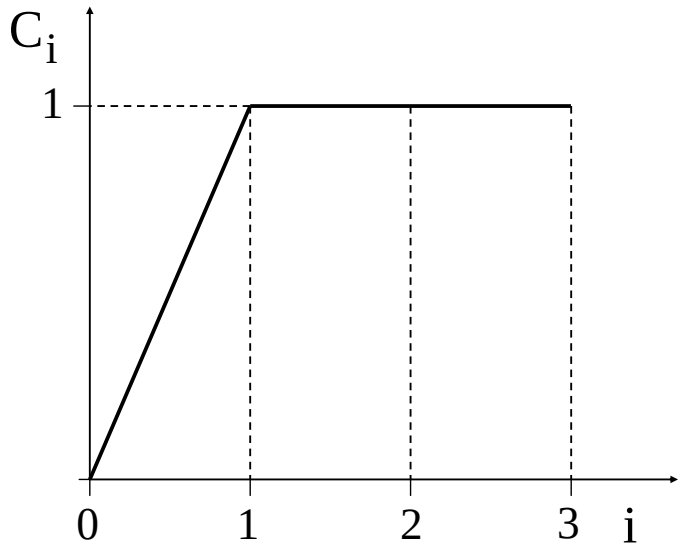
- Maximale Konzentration:
  - Ein Merkmalsträger hält den gesamten Merkmalsbetrag (d.h. 100%). Die übrigen Merkmalsträger sind „Nullträger“.
  - Merkmalsanteile: $c_1 = 1, c_2 = c_3 = \cdots = c_n = 0$

Minimale Konzentration versus maximale Konzentration (II)

Konzentrationskurven bei minimaler und maximaler Konzentration



minimale Konzentration (n = 3)



maximale Konzentration (n = 3)

- Eine empirisch ermittelte Konzentrationskurve liegt zwischen den beiden Extremzuständen. Je weiter die Kurve von der sich bei minimaler Konzentration ergebenden Geraden abweicht, desto stärker ist die Konzentration ausgeprägt.
- Darstellung der Konzentrationskurve bei einer großen Anzahl von Merkmalsträgern: Nur aus Konzentrationsraten der „größten“ Merkmalsträger oder Vervollständigung durch eine unterbrochene Verbindungslinie

Beispiel: Konzentrationen (I)

In einer Branche treten 10 Unternehmen als Anbieter auf. Die beiden größten Unternehmen erzielen einen Jahresumsatz in Höhe von 250 Mio. € und 200 Mio. €, während die Umsätze der drei mittelständischen Unternehmen 45 Mio. €, 50 Mio. € und 40 Mio. € betragen. Die kleineren Unternehmen erreichen dagegen zusammen nur einen Umsatz von 45 Mio. €.

- Merkmalssumme:

$$S = \sum_{i=1}^{10} x_i = 250 + 200 + 45 + 50 + 40 + 45 = 630[Mio.€]$$

- Geordnete Reihe:

$$x^{(1)} = 250, x^{(2)} = 200, x^{(3)} = 50, x^{(4)} = 45, x^{(5)} = 40$$

- Merkmalsanteile (= Marktanteile):

$$c_1 = \frac{x^{(1)}}{S} = \frac{250}{630} = 0{,}397; c_2 = \frac{x^{(2)}}{S} = \frac{200}{630} = 0{,}317; c_3 = \frac{x^{(3)}}{S} = \frac{50}{630} = 0{,}079;$$

$$c_4 = \frac{x^{(4)}}{S} = \frac{45}{630} = 0{,}071; c_5 = \frac{x^{(5)}}{S} = \frac{40}{630} = 0{,}063$$

Beispiel: Konzentrationen (II)

- Konzentrationsraten:

$$C_1 = c_1 = 0{,}397$$

$$C_2 = C_1 + c_2 = 0{,}397 + 0{,}317 = 0{,}714$$

$$C_3 = C_2 + c_3 = 0{,}714 + 0{,}079 = 0{,}793$$

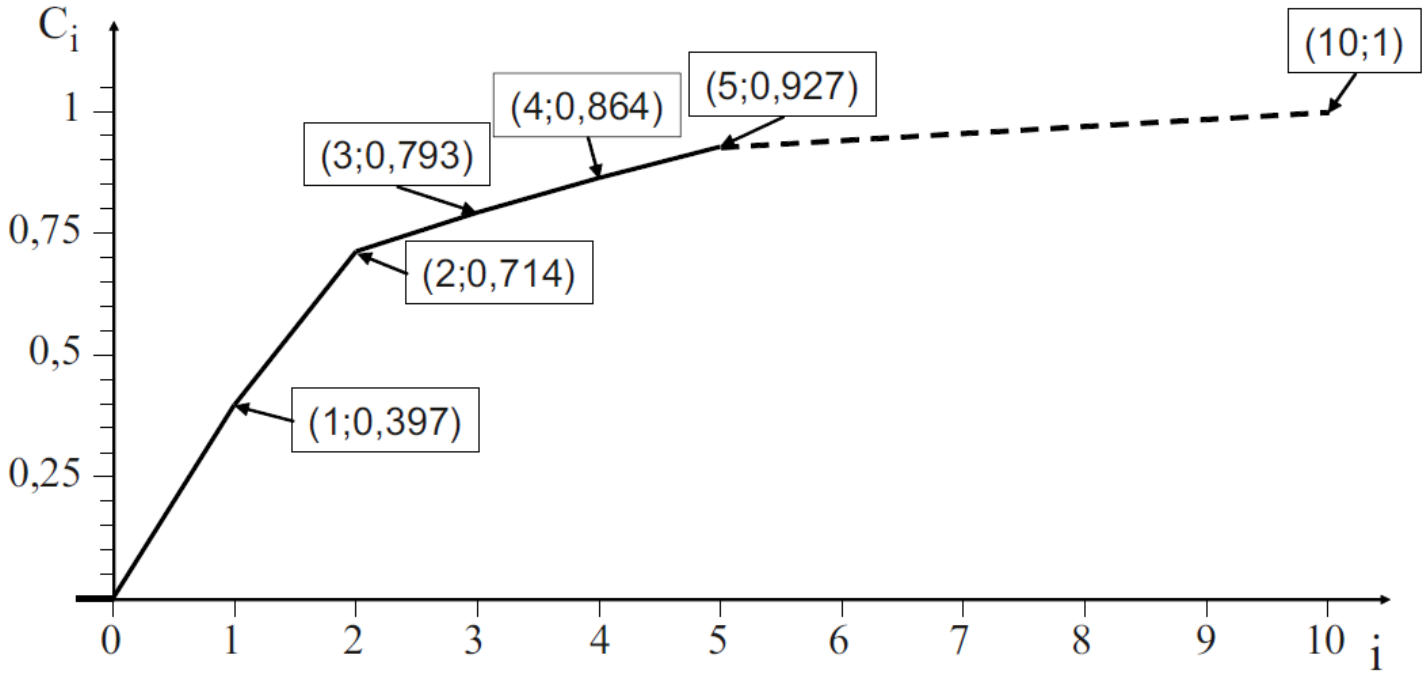
$$C_4 = C_3 + c_4 = 0{,}793 + 0{,}071 = 0{,}864$$

$$C_5 = C_4 + c_5 = 0{,}864 + 0{,}063 = 0{,}927$$

- Vermutung einer marktbeherrschenden Stellung nach §18 GWB:
  - $C_1 = 0{,}397 < 0{,}4$ allerdings
  - $C_3 = 0{,}793 > 0{,}5$ und
  - $C_5 = 0{,}927 > 2/3$

Beispiel: Konzentrationen (III)

- Konzentrationskurve (mit unterbrochener Verbindungslinie):



Beispiel: Konzentrationen (IV)

Für die kleineren Unternehmen liegen nun ebenfalls Umsatzzahlen vor: 10 Mio. €, 5 Mio. €, 20 Mio. €, 8 Mio. € und 2 Mio. €.

- Vollständige geordnete Reihe:

$$x^{(1)} = 250, x^{(2)} = 200, x^{(3)} = 50, x^{(4)} = 45, x^{(5)} = 40$$

$$x^{(6)} = 20, x^{(7)} = 10, x^{(8)} = 8, x^{(9)} = 5, x^{(10)} = 2$$

- Merkmalsanteile:

$$c_6 = \frac{x^{(6)}}{S} = \frac{20}{630} = 0{,}032; c_7 = \frac{x^{(7)}}{S} = \frac{10}{630} = 0{,}016; c_8 = \frac{x^{(8)}}{S} = \frac{8}{630} = 0{,}013;$$

$$c_9 = \frac{x^{(9)}}{S} = \frac{5}{630} = 0{,}008; c_{10} = \frac{x^{(10)}}{S} = \frac{2}{630} = 0{,}003$$

- Konzentrationsraten:

$$C_6 = C_5 + c_6 = 0{,}927 + 0{,}032 = 0{,}959$$

$$C_7 = C_6 + c_7 = 0{,}959 + 0{,}016 = 0{,}975$$

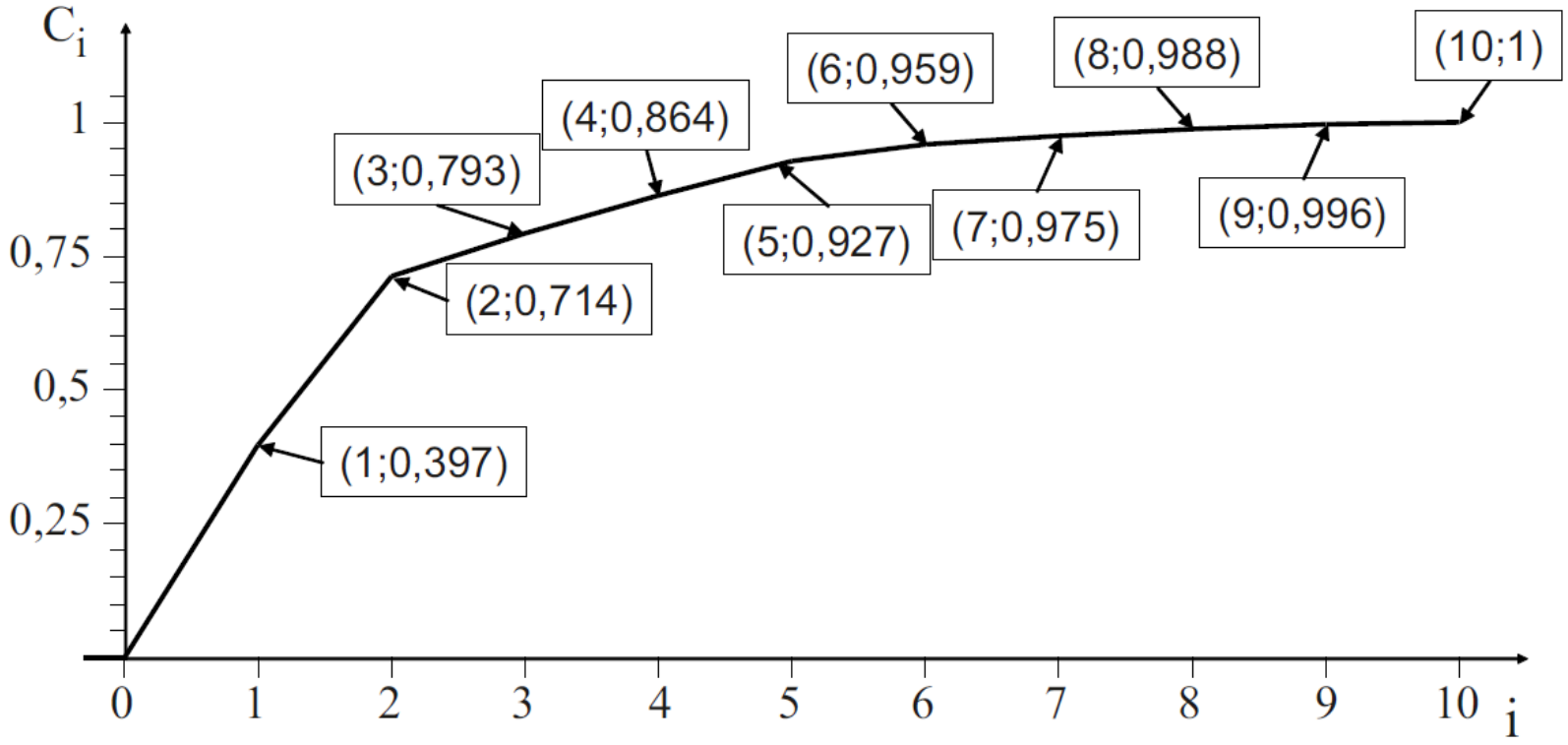
$$C_8 = C_7 + c_8 = 0{,}975 + 0{,}013 = 0{,}988$$

$$C_9 = C_8 + c_9 = 0{,}988 + 0{,}008 = 0{,}996$$

$$C_{10} = C_9 + c_{10} = 0{,}996 + 0{,}003 \approx 1$$

Beispiel: Konzentrationen (V)

- Vollständige Konzentrationskurve:



- Vergleich der Konzentrationsraten:

  Das Ausmaß der Konzentration ist entscheidend durch die großen und mittleren Unternehmen geprägt.

6.3 Herfindahl-Index

Probleme bei Konzentrationsraten:

- Sie geben nur einen statischen Einblick in die Konzentration auf einem Markt.
- Zudem sind Vergleiche von Konzentrationen schwierig (Konzentrationskurven können sich schneiden).

Herfindahl-Index:

- Globale Maßzahl der absoluten Konzentration
- Summe der quadrierten Merkmalsanteile c_i

$$K_H = \sum_{i=1}^{n} c_i^2$$

Approximativer Herfindahl-Index:

- Falls die Merkmalsbeträge nur für die n_1 wichtigsten Merkmalsträger vorliegen:

$$K_H' = \sum_{i=1}^{n_1} c_i^2$$

Wertebereich des Herfindahl-Index (Normierung)

Minimale Konzentration ($c_i = 1/n$ für alle i)**:**

$$K_H = \sum_{i=1}^{n} c_i^2 = \sum_{i=1}^{n} \left(\frac{1}{n}\right)^2 = \sum_{i=1}^{n} \left(\frac{1}{n^2}\right) = n \cdot \frac{1}{n^2} = \frac{1}{n}$$

- Untere Schranke

Maximale Konzentration ($c_1 = 1, c_2 = c_3 = \cdots = c_n = 0$)**:**

$$K_H = \sum_{i=1}^{n} c_i^2 = 1^2 + 0^2 + \cdots + 0^2 = 1$$

- Obere Schranke

Herfindahl-Index:

- Der Herfindahl-Index K_H liegt im Intervall $1/n \leq K_H \leq 1$.
- Je höher sein Wert, desto stärker ist die Konzentration ausgeprägt.

Beispiel: Herfindahl-Index

- Unter Verwendung der vorherigen fünf größten Unternehmen lässt sich der Herfindahl-Index näherungsweise bestimmen:

$$K'_H = \sum_{i=1}^{n_{1=5}} c_i^2 = 0{,}397^2 + 0{,}317^2 + 0{,}079^2 + 0{,}071^2 + 0{,}063^2$$
$$= 0{,}1576 + 0{,}1005 + 0{,}0062 + 0{,}0050 + 0{,}0040 = 0{,}273$$

- Wenn alle Merkmalsanteile bekannt sind, kann man den exakten Koeffizienten ermitteln. Sein genauer Wert unterscheidet sich von der näherungsweisen Berechnung nur auf der dritten Nachkommastelle:

$$K_H = \sum_{i=1}^{n=10} c_i^2 = 0{,}397^2 + 0{,}317^2 + 0{,}079^2 + 0{,}071^2 + 0{,}063^2$$
$$+ 0{,}032^2 + 0{,}016^2 + 0{,}013^2 + 0{,}008^2 + 0{,}003^2$$
$$= 0{,}1576 + 0{,}1005 + 0{,}0062 + 0{,}0050 + 0{,}0040$$
$$+ 0{,}0010 + 0{,}0003 + 0{,}0002 + 0{,}0001 + 0{,}0000$$
$$= 0{,}275$$

Größeneinschätzung des Herfindahl-Index

- Der Herfindahl-Index liegt meist im unteren Drittel seines Wertebereichs
- Größeneinschätzung anhand der US-Fusionsrichtlinien:
  - Niedriger Konzentrationsgrad: $K_H < 0{,}1$
  - Mittlerer Konzentrationsgrad: $0{,}1 \leq K_H \leq 0{,}18$
  - Hoher Konzentrationsgrad: $K_H > 0{,}18$

Beziehung zwischen Herfindahl-Index und Variationskoeffizient v

$$K_H = \frac{v^2 + 1}{n}$$

- Anzahlaspekt: K_H nimmt zu, wenn die Anzahl der Merkmalsträger n sinkt.
- Disparität: K_H nimmt zu, wenn die relative Streuung (v^2) des Konzentrationsmerkmals steigt.
- Als Maß der absoluten Konzentration fängt der Herfindahl-Index Aspekte der relativen Konzentration mit ein.
- Unterschiedliche Auswirkungen von Konzentrationsvorgängen auf relative und absolute Konzentration: Fusionen vergrößern zwar die absolute Konzentration, nicht notwendigerweise jedoch auch die Disparität.

Auswirkungen einer Fusion auf den Herfindahl-Index

- Vor der Fusion:
  Gesamtbeitrag der beiden Unternehmen i und h zu K_H besteht aus:

$$c_i^2 + c_h^2$$

- Nach der Fusion:
  Gesamtbeitrag der beiden Unternehmen i und h zu K_H besteht aus:

$$(c_i + c_h)^2 = c_i^2 + 2c_i c_h + c_h^2$$

- Resultat:
  K_H steigt ceteris paribus nach der Fusion der beiden Unternehmen um den Betrag:

$$2c_i c_h$$

Beispiel: Spezialmaschinen (I)

Auf einem Markt für Spezialmaschinen sind vier Unternehmen tätig, von denen jedes einen Absatz von 20 Stück erzielt. Wie verändert sich der Herfindahl-Index, wenn jeweils zwei Unternehmen fusionieren? Aus der folgenden Tabelle geht hervor, dass die Streuungen vor und nach der Fusion null sind. Durch die Fusion wird hier also nicht die Disparität berührt. Der Anstieg des Herfindahl-Indexes ist damit ausschließlich auf den Anzahlaspekt zurückzuführen.

- Arithmetisches Mittel (vor Fusion)

$$\bar{x} = \frac{1}{4} \cdot \sum_{i=1}^{4} x_i$$

$$= \frac{1}{4} \cdot (20 + 20 + 20 + 20) = 20$$

- Varianz (vor Fusion)

$$s^2 = \frac{1}{4} \cdot \sum_{i=1}^{4} (x_i - \bar{x})^2$$

$$= \frac{1}{4} \cdot \left[(20-20)^2 + (20-20)^2 + (20-20)^2 + (20-20)^2 \right]$$

$$= 0$$

- Variationskoeffizient (vor Fusion)

$$v = \frac{s}{\bar{x}} = \frac{\sqrt{0}}{20} = 0$$

- Herfindahl-Index (vor Fusion)

$$K_H = \frac{v^2 + 1}{n} = \frac{0+1}{4} = 0{,}25$$

- Arithmetisches Mittel (nach Fusion)

$$\bar{x} = \frac{1}{2} \cdot \sum_{i=1}^{2} x_i$$

$$= \frac{1}{2} \cdot (40 + 40) = 40$$

- Varianz (nach Fusion)

$$s^2 = \frac{1}{2} \cdot \sum_{i=1}^{2} (x_i - \bar{x})^2$$

$$= \frac{1}{2} \cdot \left[(40-40)^2 + (40-40)^2 \right]$$

$$= 0$$

- Variationskoeffizient (nach Fusion)

$$v = \frac{s}{\bar{x}} = \frac{\sqrt{0}}{40} = 0$$

- Herfindahl-Index (nach Fusion)

$$K_H = \frac{v^2 + 1}{n} = \frac{0+1}{2} = 0{,}5$$

Beispiel: Spezialmaschinen (II)

Vergleichen wir zwei Märkte, auf denen zwei Unternehmen konkurrieren. Die Marktanteile beim Umsatz liegen bei 0,60 und 0,40 sowie 0,70 und 0,30. Der höhere Herfindahl-Index auf dem zweiten Markt wird hier aufgrund der gleichen Anzahl von Unternehmen allein durch den Disparitätsaspekt bedingt:

- Markt 1:

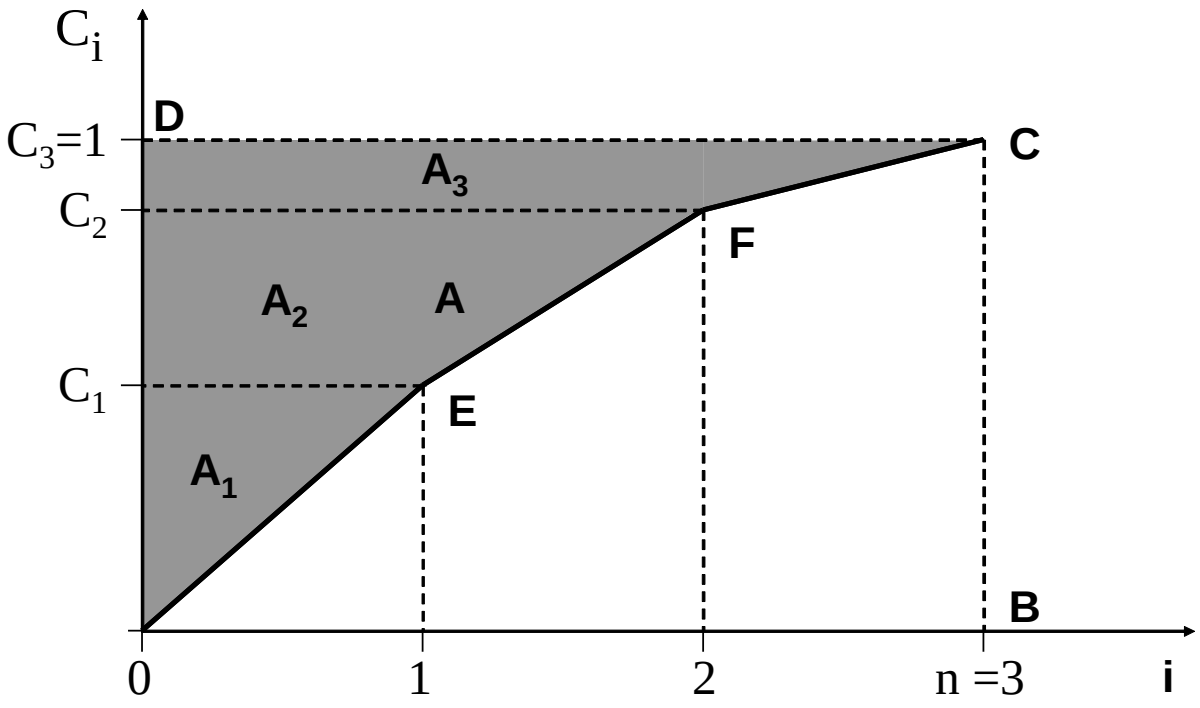
$$K_H = \sum_{i=1}^{2} c_i^2 = 0{,}6^2 + 0{,}4^2 = 0{,}52$$

- Markt 2:

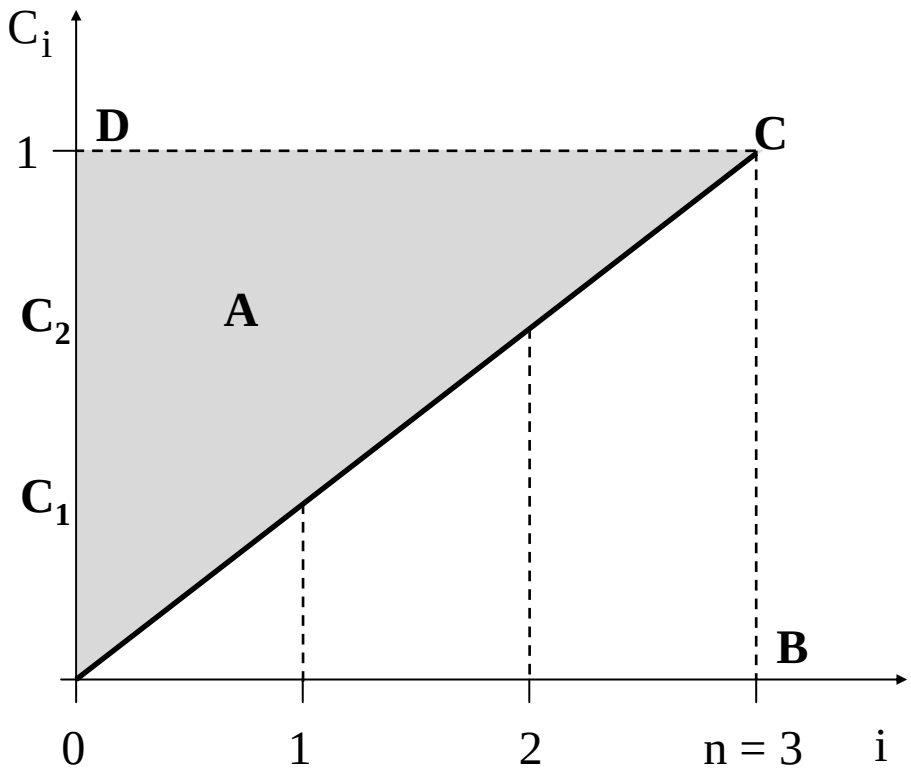
$$K_H = \sum_{i=1}^{2} c_i^2 = 0{,}7^2 + 0{,}3^2 = 0{,}58$$

6.4 Rosenbluth-Index

- Der Rosenbluth-Index K_R misst das Ausmaß der absoluten Konzentration mit Bezug auf die Konzentrationskurve.
- Der Rosenbluth-Index ist durch die Fläche A (= Dekonzentrationsfläche) oberhalb der Konzentrationskurve K innerhalb des Rechtecks 0BCD bestimmt (siehe Abbildung).
- Je größer die Fläche A ist, desto geringer ist die absolute Konzentration.

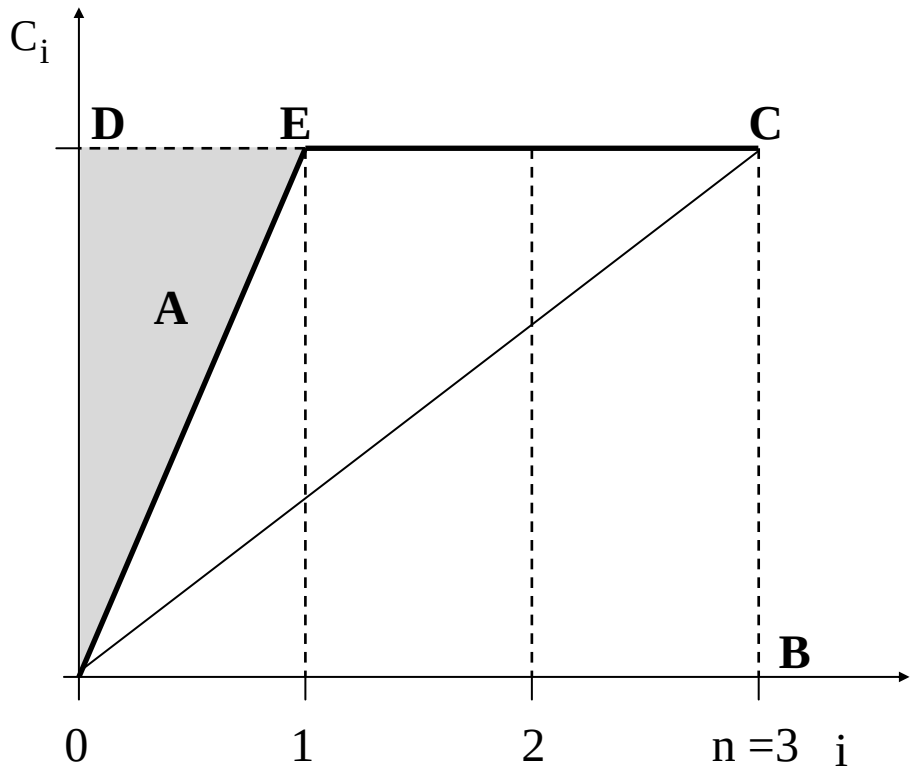


Dekonzentrationsfläche bei minimaler Konzentration



- Minimale Konzentration = maximale Dekonzentrationsfläche
- Dreiecksflächeninhalt 0CD: $A_{max} = 1/2 \cdot n = n/2$

Dekonzentrationsfläche bei maximaler Konzentration



- Maximale Konzentration = minimale Dekonzentrationsfläche
- Dreiecksflächeninhalt 0ED: $A_{min} = 1/2$

Berechnung des Rosenbluth-Indizes (I)

- Da wir nicht die Dekonzentration, sondern die Konzentration messen, ist der Rosenbluth-Index als inverse Funktion des Flächeninhalts A definiert:

$$K_R = (2 \cdot A)^{-1}$$

- Diese Definition stellt sicher, dass K_R im Intervall [0;1] liegt:
  - $K_R \rightarrow 0$: fehlende Konzentration (Dekonzentration)
  - $K_R = 1/n$: minimale Konzentration
  - $K_R = 1$: maximale Konzentration
- Flächeninhalt A (Fläche oberhalb der Konzentrationskurve K): Summe der Flächeninhalte verschiedener Trapeze (s. Abbildung auf Folie 1 des Kapitels)

$$A = \sum_{i=1}^{3} A_i$$

$$A_1 = (c_1 \cdot 1)/2 = 0{,}5 \cdot c_1$$

$$A_2 = (c_2 \cdot 1) + (c_2 \cdot 1)/2 = 1{,}5 \cdot c_2$$

$$A_3 = (c_3 \cdot 1) + (c_3 \cdot 1) + (c_3 \cdot 1)/2 = (c_3 \cdot 2) + (c_3 \cdot 1)/2 = 2{,}5 \cdot c_3$$

Berechnung des Rosenbluth-Indizes (II)

- Allgemein ist die Trapezfläche A_i gegeben durch:

$$A_i = \frac{(i-1)+i}{2} \cdot c_i = \frac{2i-1}{2} \cdot c_i$$

- Daher ergibt sich A allgemein wie folgt:

$$A = \sum_{i=1}^{n} A_i = \sum_{i=1}^{n} \frac{2i-1}{2} \cdot c_i = \frac{1}{2} \sum_{i=1}^{n} (2i-1) \cdot c_i$$
$$= \frac{1}{2} \sum_{i=1}^{n} 2 \cdot i \cdot c_i - \frac{1}{2} \sum_{i=1}^{n} c_i = \sum_{i=1}^{n} i \cdot c_i - \frac{1}{2}$$

- Einsetzen in die Formel für K_R ergibt:

$$K_R = \frac{1}{2 \cdot \sum_{i=1}^{n} i \cdot c_i - 1}$$

Beispiel: Rosenbluth-Index (I)

Der Rosenbluth-Index soll für eine Branche berechnet werden, in der 10 Unternehmen miteinander konkurrieren. Die Merkmalsanteile (= Marktanteile) c_i sind zur Darstellung der Konzentrationskurve ermittelt worden.

i	c_i	$i \cdot c_i$
1	0,397	0,397
2	0,317	0,634
3	0,079	0,237
4	0,071	0,284
5	0,063	0,315
6	0,032	0,192
7	0,016	0,112
8	0,013	0,104
9	0,008	0,072
10	0,003	0,030
Σ		2,377

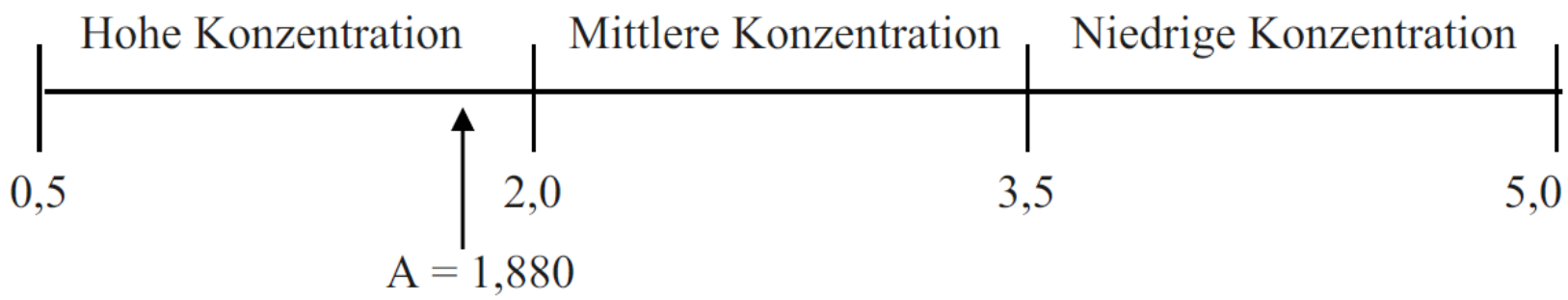
$$K_R = \frac{1}{2 \cdot \sum_{i=1}^{10}(i \cdot c_i) - 1} = \frac{1}{2 \cdot 2{,}377 - 1} = \frac{1}{3{,}754} = 0{,}266$$

Beispiel: Rosenbluth-Index (II)

Größeneinschätzung des Rosenbluth-Index (mit n = 10):

- $A_{min} = 1/2$
- $A_{max} = n/2 = 10/2 = 5$
- $A = \frac{1}{2 \cdot K_R} = \frac{1}{2 \cdot 0{,}266} = 1{,}880$

Aufteilung in drei Bereiche $\left[\left((5 - 0{,}5)\right)/3 = 1{,}5\right]$:



Der Rosenbluth-Index weist in Übereinstimmung mit dem Herfindahl-index eine hohe absolute Konzentration der Umsätze in der betrachteten Branche aus.

6.5 Lorenzkurve und Gini-Koeffizient

- Grundgesamtheit besteht aus einer Vielzahl von Einheiten: Vor allem bei Personen, Haushalten, Familien
- In diesem Fall ist ein Übergang von absoluten zu relativen Konzentrationsmaßen (Disparität) zweckmäßig

Beispiele:

- Auf eine größere Anzahl von Personen bzw. Haushalten (Gruppe, Klasse) entfällt ein hoher Anteil am Volkseinkommen oder Vermögen, während ihr dagegen Anteil an der gesamten Bevölkerung nur gering ist.
- Die relative Konzentration oder Disparität macht Aussagen über Ungleichheiten z.B. in der Form, dass 2% der Bevölkerung 80% des Produktivvermögens besitzen oder 40% der Einkommensempfänger 10% des Volkseinkommens beziehen.

Lorenzkurve (grafisches Instrument) (I)

- Wir betrachten die Konstruktion der Lorenzkurve für klassierte Daten.
- Zur Zeichnung der Lorenzkurve benötigen wir die die kumulierte relativen Klassenhäufigkeiten $H_k = H_{k-1} + h_k$ sowie die kumulierten Merkmalsanteile Q_k.
- Die **Merkmalsanteile**

$$q_k = \frac{S_k}{S} = \frac{\bar{x}_k \cdot n_k}{\sum_{k=1}^{p} \bar{x}_k \cdot n_k} = \frac{\bar{x}_k \cdot h_k}{\sum_{k=1}^{p} \bar{x}_k \cdot h_k}$$

  berechnet man als Quotient der Merkmalssumme S_k der k-ten Klasse

$$S_k = \bar{x}_k \cdot n_k$$

  sowie der Merkmalssumme S insgesamt

$$S = \sum_{k=1}^{p} S_k = \sum_{k=1}^{p} \bar{x}_k \cdot n_k$$

- Die kumulierte Merkmalsanteile:

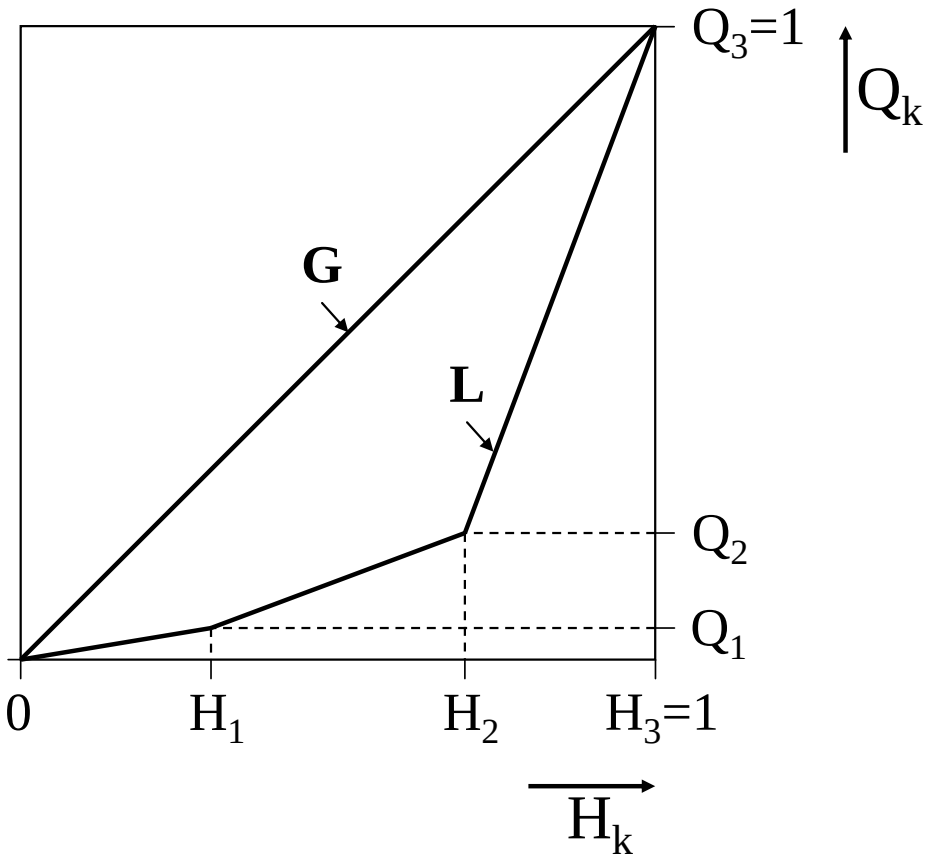
$$Q_k = Q_{k-1} + q_k$$

  geben an, welcher Anteil der Merkmalssumme des Konzentrationsmerkmals auf die ersten k Klassen entfällt.

Lorenzkurve (grafisches Instrument) (II)

- Lineare Verbindung der Punkte (0,0), (H_1,Q_1), (H_2,Q_2), ..., (1,1) in einem kartesischen Koordinatensystem
- Gleichverteilungsgerade G: 45°-Linie, die die Punkte (0,0) und (1,1) verbindet

Lorenzkurve bei drei Klassen



Eigenschaften der Lorenzkurve

- Gleichmäßige Verteilung des Konzentrationsmerkmals:
  - Merkmalsanteile q_k = relative Klassenhäufigkeiten h_k

  → Lorenzkurve L fällt mit Gleichverteilungsgeraden G zusammen
- Je „ungleicher“ die Merkmalsbeträge auf die einzelnen Klassen verteilt sind, desto größer wird die Fläche F zwischen der Lorenzkurve und der Gleichverteilungsgeraden.
- Die Lorenzkurve „hängt“ umso mehr „durch“, je größer die Abweichung von der egalitären Verteilung ist.
- Steigung der Lorenzkurve:

$$\frac{\Delta Q_k}{\Delta H_k} = \frac{Q_k - Q_{k-1}}{H_k - H_{k-1}} = \frac{q_k}{h_k} = \frac{\bar{x}_k \cdot h_k}{\sum_{k=1}^{p} \bar{x}_k \cdot h_k} \cdot \frac{1}{h_k} = \frac{\bar{x}_k \cdot h_k}{\bar{x}} \cdot \frac{1}{h_k} = \frac{\bar{x}_k}{\bar{x}}$$

- Die Steigung der Lorenzkurve wächst monoton (konvexer Verlauf):
  - $\frac{\bar{x}_k}{\bar{x}} < 1$ für $\bar{x}_k < \bar{x}$
  - $\frac{\bar{x}_k}{\bar{x}} = 1$ für $\bar{x}_k = \bar{x}$
  - $\frac{\bar{x}_k}{\bar{x}} > 1$ für $\bar{x}_k > \bar{x}$

Beispiel: Einkommensverteilung (I)

Um Aufschluss über Veränderungen in der Einkommensverteilung zu erhalten, gibt das Ministerium für Arbeit und Soziales eines Bundeslandes eine Auswertung amtlicher Daten in Auftrag. In dem Bundesland, das 12 Mio. Einwohner:innen hat, sind die Einkommen (in €) wie folgt verteilt:

Klasse	von über ... bis zu	Klassenmittelwert	Anzahl der Personen (in Mio.)
1	0 – 1000	600	2
2	1000 – 2000	1600	6
3	2000 – 3000	2500	3
4	3000 – 5000	3900	0,9
5	über 5000	10000	0,1

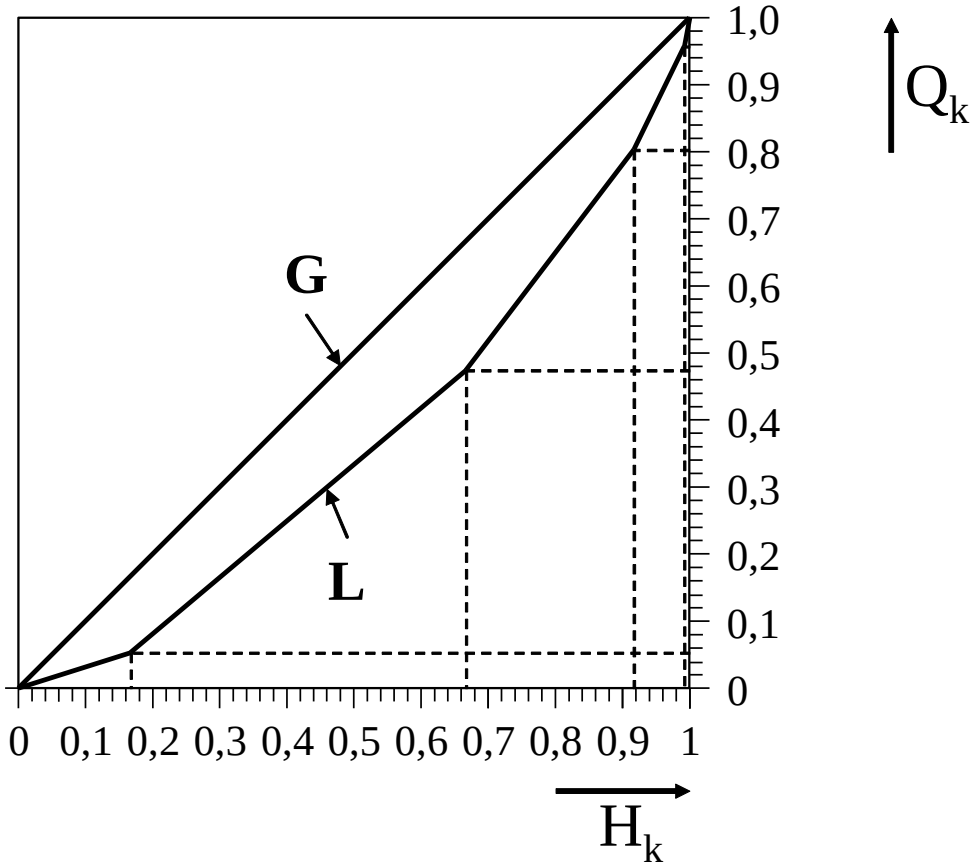
Beispiel: Einkommensverteilung (II)

Um eine Lorenzkurve der Einkommenskonzentration zu konstruieren, wird eine Arbeitstabelle erstellt:

k	h_k	H_k	$\bar{x}_k \cdot n_k$	q_k	Q_k
1	2/12 = 0,167	0,167	600 · 2 = 1200	$\frac{1200}{22810} = 0{,}053$	0,053
2	6/12 = 0,5	0,167 + 0,500 = 0,667	1600 · 6 = 9600	$\frac{9600}{22810} = 0{,}421$	0,053 + 0,421 = 0,474
3	3/12 = 0,25	0,667 + 0,250 = 0,917	2500 · 3 = 7500	$\frac{1200}{22810} = 0{,}329$	0,474 + 0,329 = 0,803
4	0,9/12 = 0,075	0,917 + 0,075 = 0,992	3900 · 0,9 = 3510	$\frac{3510}{22810} = 0{,}154$	0,803 + 0,154 = 0,957
5	0,1/12 = 0,008	0,992 + 0,008 = 1,000	10000 · 0,1 = 1000	$\frac{1000}{22810} = 0{,}044$	0,957 + 0,044 = 1,000
∑	1		22810 $\bar{x} = 1900{,}83$	1	

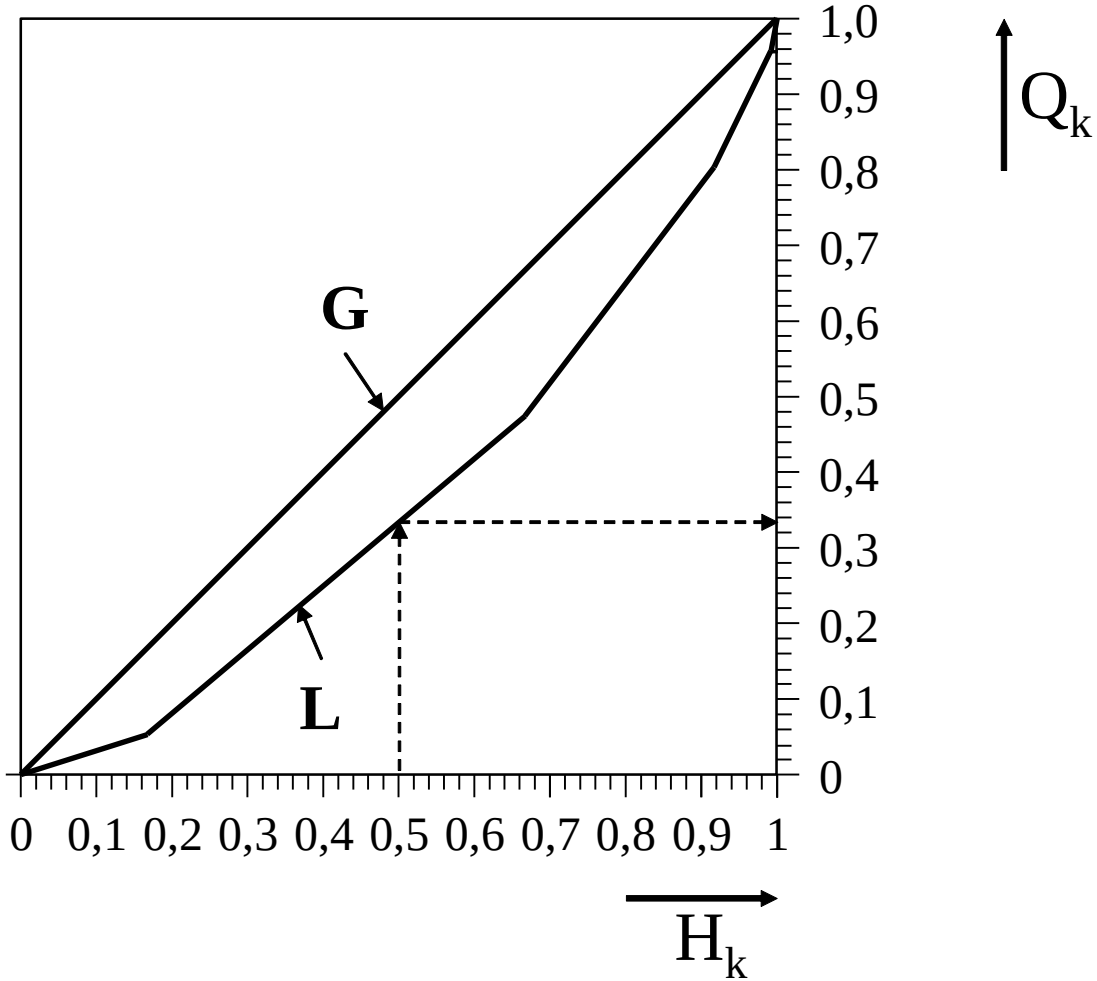
Beispiel: Einkommensverteilung (III)

- Wenn die kumulierten relativen Häufigkeiten und Merkmalsanteile in ein Koordinatensystem eingetragen werden, erhält man die Lorenzkurve
- 16,7% der Einkommensempfänger:innen (unterste Klasse) beziehen nur 5,3% des Gesamteinkommens.
- 0,8% der Einkommensempfänger:innen (oberste Klasse) beziehen 4,4% des Gesamteinkommens.



Beispiel: Einkommensverteilung (IV)

Welchen Einkommensanteil beziehen die unteren 50% der Einkommens-bezieher:innen?



Lorenzkurve bei unvollständiger Information (I)

- Klassenmittelwerte $\bar{x}_k$ unbekannt
- Berechnung der Merkmalsanteile durch Ersetzung der Klassenmittelwerte $\bar{x}_k$ durch die Klassenmitten m_k:

$$q_k = \frac{m_k \cdot n_k}{\sum_{k=1}^{p} m_k \cdot n_k} = \frac{m_k \cdot h_k}{\sum_{k=1}^{p} m_k \cdot h_k}$$

Beispiel: Umsatz

Eine Erhebung des Absatzes (in Tsd. Stück) von 10 Unternehmen einer Branche hat zu folgenden Ergebnissen geführt:

Klasse	von über ... bis zu	Anzahl der Unternehmen
1	500 – 1000	2
2	1000 – 2000	3
3	2000 – 4000	5

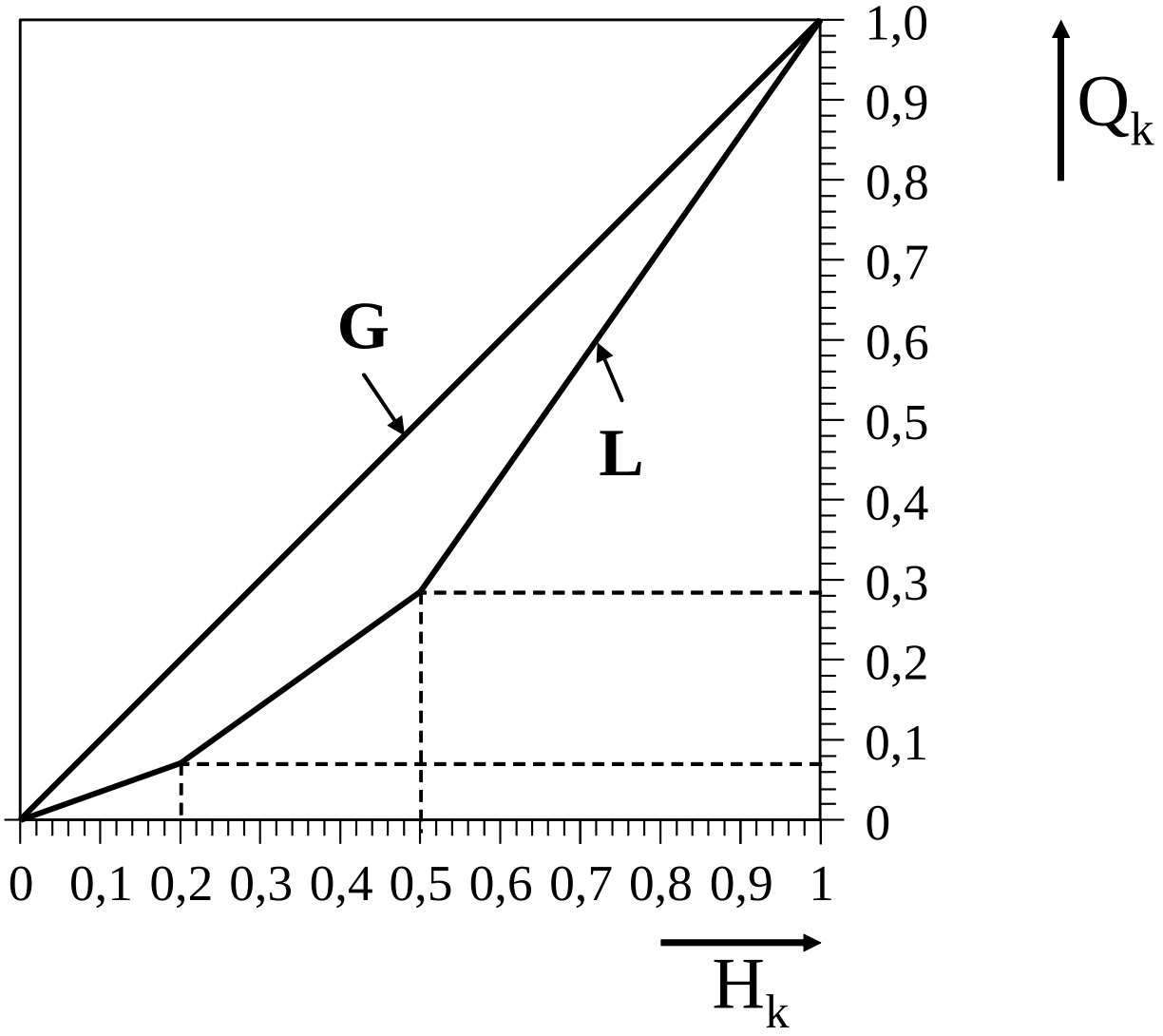
Lorenzkurve bei unvollständiger Information (II)

Berechnung der Merkmalsanteile mit Klassenmitten

k	h_k	H_k	m_k	$m_k \cdot n_k$	$q_k = \frac{m_k \cdot n_k}{\sum_{k=1}^{p} m_k \cdot n_k}$	Q_k
1	$\frac{2}{10} = 0{,}2$	0,2	$\frac{1}{2} \cdot (500 + 1000) = 750$	$750 \cdot 2 = 1500$	$\frac{1500}{21000} = 0{,}071$	0,071
2	$\frac{3}{10} = 0{,}3$	$0{,}2 + 0{,}3 = 0{,}5$	$\frac{1}{2} \cdot (1000 + 2000) = 1500$	$1500 \cdot 3 = 4500$	$\frac{4500}{21000} = 0{,}214$	$0{,}071 + 0{,}214 = 0{,}285$
3	$\frac{5}{10} = 0{,}5$	$0{,}5 + 0{,}5 = 1{,}0$	$\frac{1}{2} \cdot (2000 + 4000) = 3000$	$3000 \cdot 5 = 15000$	$\frac{15000}{21000} = 0{,}714$	$0{,}285 + 0{,}714 = 1{,}000$
∑	1			21000 $\bar{x}_m = 2100$	1	

Lorenzkurve bei unvollständiger Information (III)

Mit diesen Daten ergibt sich die folgende Lorenzkurve:



Gini-Koeffizient

- Der **Gini-Koeffizient R** gibt die Informationen aus der Lorenzkurve in einer Kennzahl wieder.
- Je höher die Werte des Gini-Koeffizienten sind (0 ≤ R < 1), desto größer ist die Konzentration.

$$R = \frac{\text{Fläche zwischen der Lorenzkurve und der Gleichverteilungsgeraden}}{\text{Fläche des Dreiecks unter der Gleichverteilungsgraden}}$$

- Formel (Herleitung siehe Lehrbuch):

$$R = 1 - \sum_{k=1}^{p} (Q_{k-1} + Q_k) \cdot h_k$$

- Interpretation:
  - $0 \leq R \leq 0{,}3$: niedrige Konzentration
  - $0{,}3 \leq R \leq 0{,}7$: mittlere Konzentration
  - $0{,}7 \leq R < 1{,}0$: starke Konzentration
  - $R = 0$: Egalitäre Gleichverteilung (Gleichverteilung)
  - $R \to 1$ bei zunehmender Konzentration

Beispiel (Fortsetzung): Einkommensverteilung

Aus dem Einkommensbeispiel sind die relativen Häufigkeiten sowie die Merkmals-anteile noch einmal wiedergegeben:

k	h_k	H_k	$\bar{x}_k \cdot n_k$	q_k	Q_k
1	0,167	0,167	1200	0,053	0,053
2	0,500	0,667	9600	0,421	0,474
3	0,250	0,917	7500	0,329	0,803
4	0,075	0,992	3510	0,154	0,957
5	0,008	1,000	1000	0,044	1,000
∑	1		22810	1	

$$R = 1 - \sum_{k=1}^{5} (Q_{k-1} + Q_k) \cdot h_k = 1 - (0{,}009 + 0{,}264 + 0{,}319 + 0{,}132 + 0{,}016)$$
$$= 1 - 0{,}740 = 0{,}260$$

Die Konzentration der Einkommen ist also relativ niedrig.

Beispiel (Fortsetzung): Umsatz

Für die Konzentration des Absatzes (in Tsd. Stück) von 10 Unternehmen einer Branche lässt sich folgender Gini-Koeffizient berechnen:

k	h_k	H_k	m_k	$m_k \cdot n_k$	q_k	Q_k
1	0,2	0,2	750	1500	0,071	0,071
2	0,3	0,5	1500	4500	0,214	0,285
3	0,5	1,0	3000	15000	0,714	1,000
∑	1			21000	1	

$$R = 1 - \sum_{k=1}^{3} (Q_{k-1} + Q_k) \cdot h_k$$
$$= 1 - [(0 + 0{,}071) \cdot 0{,}2 + (0{,}071 + 0{,}285) \cdot 0{,}3 + (0{,}285 + 1) \cdot 0{,}5]$$
$$= 1 - (0{,}014 + 0{,}107 + 0{,}643) = 1 - 0{,}764 = 0{,}236$$

R zeigt hier ebenfalls eine geringe Konzentration an.

Normierter Gini-Koeffizient (I)

- Vollkommene Ungleichheit:

$$R_{max} = 1 - \frac{1}{n}$$

- Bei kleinem Erhebungsumfang ist R_{max} damit deutlich kleiner als eins.
- Der normierte Gini-Koeffizient gewährleistet das Erreichen der maximalen Schranke von eins im Falle einer maximalen Konzentration.

$$R^* = \frac{R}{R_{max}}$$

- Bei maximaler Konzentration: $R^* = 1$

Beispiel (Fortsetzung): Einkommensverteilung

In einem Bundesland mit 12 Millionen Einwohner:innen beträgt der maximale Wert von R nahezu eins:

$$R_{max} = 1 - \frac{1}{n} = 1 - \frac{1}{12.000.000} \approx 1$$

Deshalb macht die Berechnung des normierten Gini-Koeffizienten hier keinen Sinn.

Normierter Gini-Koeffizient (II)

Beispiel (Fortsetzung): Umsatz

Bei Umsatzzahlen von zehn Unternehmen weicht der maximale Gini-Koeffizient deutlich von eins ab:

$$R_{max} = 1 - \frac{1}{n} = 1 - \frac{1}{10} = 0{,}9$$

Sinnvoll wäre deshalb die Berechnung des normierten Gini-Koeffizienten:

$$R^* = \frac{R}{R_{max}} = \frac{0{,}236}{0{,}9} = 0{,}262$$