# Statistik

## Corona - Notversion

M. Oettinger

06.05.2020

## Beispiel: Quartilsabstand

Für die Altersverteilung im Kurs

| Alter $x_i$ | Häufigkeit $h_i$ | rel. Häufigkeit $f_i$ | kumuliert $F_i$ |
|---|---|---|---|
| 19 | 6 | 0.250 | 0.250 |
| 20 | 5 | 0.208 | 0.458 |
| 21 | 4 | 0.167 | 0.625 |
| 22 | 3 | 0.125 | 0.750 |
| 23 | 2 | 0.083 | 0.833 |
| 24 | 1 | 0.042 | 0.875 |
| 29 | 1 | 0.042 | 0.917 |
| 33 | 1 | 0.042 | 0.959 |
| 48 | 1 | 0.042 | 1 |

kann der Quartilsabstand (die Differenz zw. unterem Quartil $\bar{x}_{0,25}$ und oberem Quartil $\bar{x}_{0,75}$ bestimmt werden.

## Beispiel: Quartilsabstand

Die Stichprobe besteht aus $n = 24$ Merkmalswerten, also ist (in a)

- für das untere Quartil $n \cdot p = 24 \cdot 0,25 = 6$ ganzzahlig. Das untere Quartil ist $\frac{x_i + x_{i+1}}{2}$ mit $i = n \cdot p = 6$:

$$\bar{x}_{0,25} = \frac{x_6 + x_7}{2} = \frac{19 + 20}{2} = 19,5$$

- für das obere Quartil $n \cdot p = 24 \cdot 0,75 = 8$ ganzzahlig. Das untere Quartil ist $\frac{x_i + x_{i+1}}{2}$ mit $i = n \cdot p = 18$:

$$\bar{x}_{0,75} = \frac{x_{18}^{+} x_{19}}{2} = \frac{22 + 23}{2} = 22,5$$

# Streuung

## Beispiel: Quartilsabstand

Der Quartilsabstand ist $22,5\text{a} - 19,5\text{ a} = 3\text{ a}$, er steigt mit der Breite der Verteilung an, ist aber als Streumaß speziell bei asymmetrischen Verteilungen nur bedingt geeignet.

Ein Box-Plot (auch Box-Whisker-Plot) ist eine grafische Darstellung der Verteilung (meist) kardinaler Merkmale. Dabei fasst man robuste Lage- und Streumaße in einer einfachen Darstellung zusammen, die einen schnellen Überblick über Lage und Verteilung von Merkmalswerten erlaubt.

# Box-Whisker-Plot

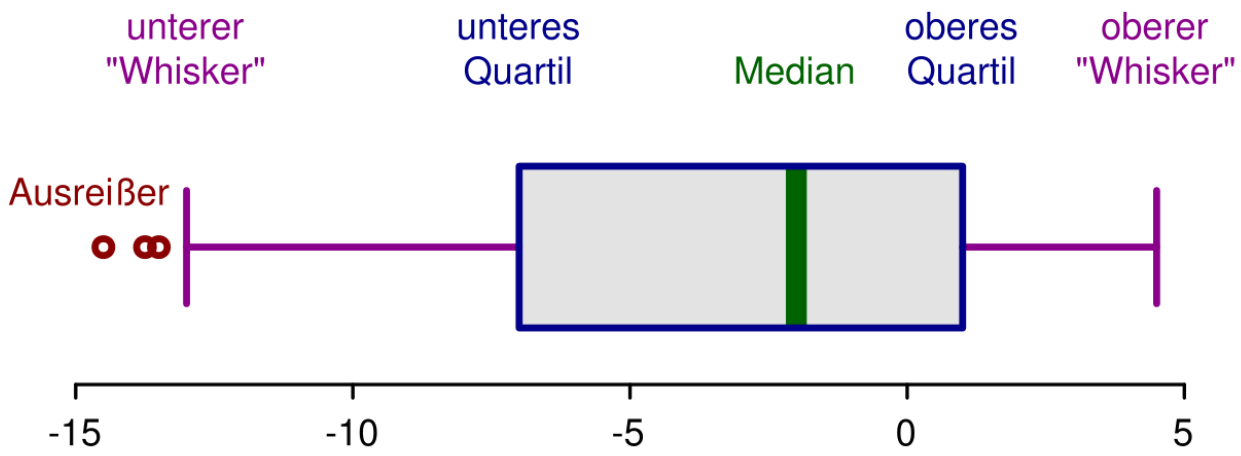



**Abbildung:** Box-Plot (wikimedia: RobSeb / CC-BY-SA-3.0)

Im Box-Plot werden die Extremwerte $x_{max}/x_{min}$, zwei Quantile (meist Quartile $\bar{x}_{0,25}$ und $\bar{x}_{0,75}$) und der Median $\bar{x}_Z$ vereinfacht dargestellt. Der Plot besteht aus einem Rechteck vom unteren bis zum oberen Quantil und zwei Linien (Whisker/Antennen), die das Rechteck verlängern. Der Strich in der Box repräsentiert den Median der Verteilung.

# Box-Whisker-Plot

## Beispiel: Altersverteilung im Kurs

Für die Altersverteilung sind die Kennzahlen bekannt (in a): $\bar{x}_Z = 21$, $\bar{x}_{0,25} = 19,5$, $\bar{x}_{0,75} = 22,5$, $x_{min} = 19$ und $x_{max} = 48$.

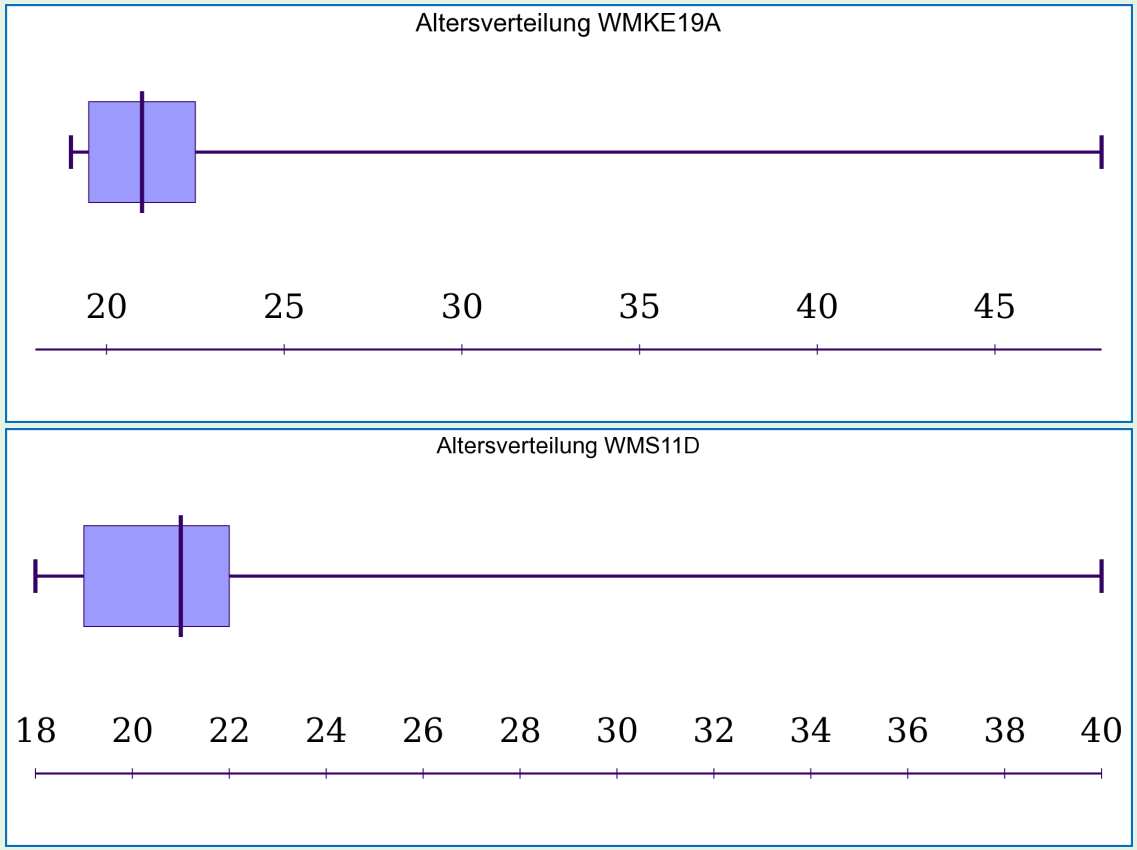



**Abbildung:** Altersverteilung zweier Statistik-Kurse

# Variationskoeffizient

die bisherigen Streumaße (Spannweite, mittlere absolute Abweichung vom Mittelwert, Standardabweichung) geben die absolute Breite einer Verteilung an. Relative Streumaße sind dagegen dimensionslos, sie geben die Abweichung bezogen auf den Mittelwert an. Ein Beispiel dafür ist der *Variationskoeffizient*.

**Definition: Definition des Variationskoeffizienten**

Für ein kardinales Merkmal $X$ mit arithmetischem Mittel $\bar{x}$ und empirischer Standardabweichung $s_X$ ist der Variationskoeffizient $v_X$ definiert durch

$$v_X := \frac{s_X}{\bar{x}}, \tag{1}$$

das absolute Streumaß $s_X$ wird ins Verhältnis zum mittleren Niveau des Merkmals $X$ gesetzt. Der Variationskoeffizient $v_X$ ist als Quotient zweier Größen gleicher Dimension und Einheiten dimensions - und einheitenlos.

# Variationskoeffizient

## Beispiel: Variationskoeffizient

Felgenrößen werden meist in Zoll angegeben (Durchmesser) und können einfach in cm umgerechnet werden (1 Zoll entspricht 2,54 cm). Haben einige Felgen die Durchmesser

| $Y$ [in cm] | 35 | 37,5 | 40 | 37,5 | 42,5 |
|---|---|---|---|---|---|
| $X$ [in Zoll] | 14 | 15 | 16 | 15 | 17 |

**Tabelle:** Felgengrößen.

ist das arithmetische Mittel der Felgengrößen nach Definition

$$\bar{x} = \frac{1}{5}(14 + 2 \cdot 15 + 16 + 17) \text{ Zoll} = 15,4 \text{Zoll}$$

$$\bar{y} = \frac{1}{5}(35 + 2 \cdot 37,5 + 40 + 42,5) \text{ cm} = 38,5 \text{ cm}$$

# Variationskoeffizient

## Beispiel: Variationskoeffizient

die Standardabweichung ist

$$s_y = \sqrt{\frac{1}{5}\left(35^2 + 2 \cdot 37,5^2 + 40^2 + 42,5^2\right) - 38,5^2}\ \text{cm} = 2,55\ \text{cm}$$

bzw. in Zoll

$$s_x = \sqrt{\frac{1}{5}\left(14^2 + 2 \cdot 15^2 + 16^2 + 17^2\right) - 15,4^2}\ \text{Zoll} = 1,0198\ \text{Zoll}$$

Die Streuung von Daten in einer Stichprobe ist natürlich unabhängig von der Wahl der Skala. Unterschiedliche Werte der Standardabweichung scheinen aber etwas anderes auszusagen.

# Variationskoeffizient

## Beispiel: Variationskoeffizient

Der Grund dafür ist, dass die Standardabweichung ein Maß für die *absolute* Streuung ist, dessen Wert von einer Einheit abhängt, in der das untersuchte Merkmal gemessen wird. Wird in einem anderen Maßstab gemessen, ändern sich die absoluten Merkmalswerte und damit auch die Standardabweichung.

Der Variationskoeefizient ist aber dimensionslos, die typische Abweichung relativ zum Mittelwert ist gleich:

$$v_x = \frac{s_x}{\bar{x}} = \frac{1,0198 \text{ Zoll}}{15,4 \text{ Zoll}} = v_y = \frac{s_y}{\bar{y}} = \frac{2,55 \text{ cm}}{38,5 \text{ cm}} = 6,6\%$$

Die empirische Standardabweichung der Felgengrößen beträgt 6,6% des Mittelwertes. Es spielt keine Rolle, in welchen Einheiten die Merkmale gemessen werden.

# Schiefe

Ob eine Verteilung symmetrisch oder unsymmetrisch (schief) ist, ist in der grafischen Darstellung leicht zu sehen. Die in Abb. 3 dargestellte Altersverteilung wird als schief - insbesondere als linkssteil oder synonym als rechtsschief - bezeichnet.

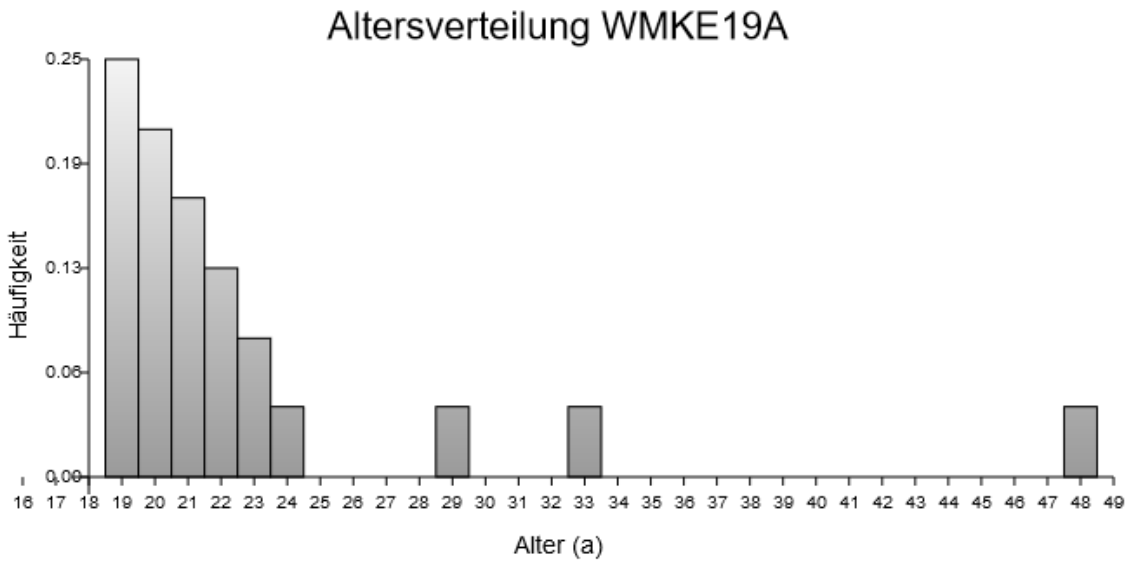



**Abbildung:** die Altersverteilung im Kurs ist linkssteil bzw. rechtsschief.

Eine einfache Möglichkeit zur Einschätzung der Schiefe eingipfliger Verteilungen mit eindeutigem Modus bietet die Lageregel.

# Lageregel

**Definition: Fechnersche Lageregel**

ist eine eingipflige Verteilung

| | | |
|---|---|---|
| linkssteil, | so gilt in der Regel | $\bar{x} \geq \bar{x}_Z \geq \bar{x}_M$ |
| rechtssteil, | so gilt in der Regel | $\bar{x} \leq \bar{x}_Z \leq \bar{x}_M$ |
| symmetrisch, | so gilt immer | $\bar{x} = \bar{x}_Z = \bar{x}_M$ |

Weil allgemein gültig $\bar{x} = \bar{x}_Z = \bar{x}_M$ bei symmetrischen, eingipfligen Verteilungen, sind mit dem arithmetischen Mittel einer symmetrischen Verteilung auch Modus und Median bekannt.

Linkssteile bzw. rechtsschiefe Verteilungen sind von großer empirischer Bedeutung. Beispielweise sind Verteilungen von Einkommens- bzw. Vermögensverhältnissen typischerweise linkssteil.

# Konzentration und Disparität

Die Konzentration misst die Ungleichheit einer Verteilung von Merkmalsausprägungen auf eine Menge von Merkmalsträgern. Sie ist bei der Verteilung von Marktanteilen von besonderem Interesse.

Im ökonomischen Sinne kann Konzentration zweierlei bedeuten:

- Die Konzentration von beispielsweise Marktanteilen, also von ökonomischer Macht, auf genau eine Wirtschaftseinheit (Monopol) oder auf lediglich einige wenige Wirtschaftseinheiten (Oligopol).
- Die Existenz erheblicher Unterschiede zwischen den Anteilen von Wirtschaftseinheiten am Gesamtbetrag eines Merkmals wie beispielsweise dem Umsatz.

# Konzentration und Disparität

## Beispiel: zur absoluten und relativen Konzentration.

Eine Aussage im Sinne einer relativen Konzentration ist beispielsweise: 2% der Bevölkerung lateinamerikanischer Staaten besitzen mehr als 90% des Geldvermögens dieser Staaten. In der Aussage tauchen ausschließlich relative Werte (angegeben in Prozenten) auf: Diese relativen Werte geben den Anteil am Gesamtwert des untersuchten Merkmals (das Geldvermögen) an, den ein bestimmter Anteil von Merkmalsträgern aufweist.

Eine Aussage im Sinne der absoluten Konzentration wäre dagegen: Auf dem deutschen Energiemarkt haben nur zwei Konzerne zusammen einen Marktanteil von etwa 80%. Die Merkmalsträger sind in absoluter Anzahl angegeben, die Zahl ist zudem sehr gering.

# Disparität

Bei der Bestimmung einer relativen Konzentration oder der Ungleichheit geht es um die Frage, ob ein großer Anteil am Gesamtwert eines Merkmals wie beispielsweise dem Energieverbrauch, um den es im folgenden Beispiel geht, auf einen geringen *Anteil* aller Merkmalsträger entfällt.

**Beispiel: fiktive Werte zum Energieverbrauch auf Gliese 581c.**

Die (fiktive) Bevölkerung des Exoplaneten Gliese 581c (GB für *G*liese-*B*evölkerung) wird politisch völlig inkorrekt in die 1, die 2. und die 3.Welt aufgeteilt.

# Disparität

## Beispiel: fiktive Werte zum Energieverbrauch auf Gliese 581c.

Der Energieverbrauch (GEV) teilt sich sortiert nach Energieverbrauch auf die drei 'Welten' folgendermaßen auf:

| | Anteil der GB $f_i$ | kumuliert $F_i$ | Anteil am GEV $q_i$ | kumuliert $Q_i$ | $(F_i; Q_i)$ |
|---|---|---|---|---|---|
| 3. Welt ($i = 1$) | 60% | 60% | 10% | 10% | $(0,6; 0,1)$ |
| 2. Welt ($i = 2$) | 30% | 90% | 30% | 40% | $(0,9; 0,4)$ |
| 1. Welt ($i = 3$) | 10% | 100% | 60% | 100% | $(1; 1)$ |

Ein geringer Anteil der Bevölkerung von etwa 10% beansprucht demnach den größten Anteil, ca. 60% der Energie, während 60% der Bevölkerung mit der geringsten Menge von 10% auskommt.

# Disparität

Die Lorenzkurve $L$ wird konstruiert mit Hilfe der Eckpunkte $(F_i; Q_i)$ aus den kumulierten relativen Anteilen einer Gruppe an der Gesamtheit

$$F_i = \sum_{k=1}^{i} f_k = f_1 + f_2 + ... + f_i \tag{2}$$

und deren kumuliertem Anteil

$$Q_i = \sum_{k=1}^{i} q_k = q_1 + q_2 + ... + q_i \tag{3}$$

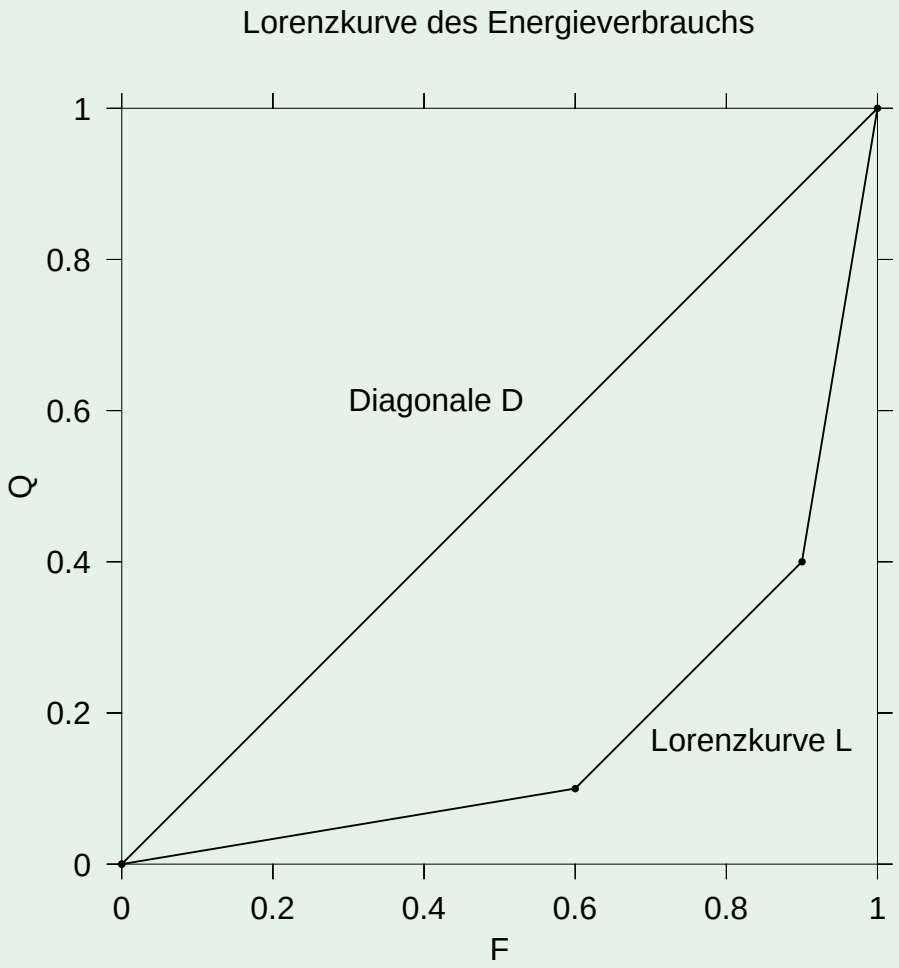
am Gesamtwert des betrachteten Merkmals. Ergänzt werden die Punkte um den Ursprung $(0; 0) = (F_0; Q_0)$ als Ausgangspunkt der Kurve. Die Lorenzkurve selbst besteht aus dem Polygonzug, der die Punkte $(F_0; Q_0), ..., (F_i; Q_i), ..., (F_n; Q_n)$ durch Geraden verbindet.

## Beispiel: Lorenzkurve des Gliese-Energieverbrauchs.

Für das obige Beispiel des Gesamtenergieverbrauchs lauten die Eckpunkte der Lorenzkurve:

$(F_0; Q_0) = (0; 0)$,
$(F_1; Q_1) = (0,6; 0,1)$,
$(F_2; Q_2) = (0,9; 0,4)$
und
$(F_3; Q_3) = (1; 1)$.

Die Lorenzkurve $L$ verbindet die Eckpunkte durch Geraden.



# Disparität

Interessant sind die beiden Grenzwerte der Kurve: bei Gleichverteilung entspricht der jeweilige Anteil des Energieverbrauchs exakt dem Bevölkerungsanteil. Die Kurve wird dann zur Diagonalen $D$ im Schaubild.

Bei maximaler Konzentration verbraucht der kleinste Teil des Bevölkerungsanteils die gesamte Energie, die Lorenzkurve besteht aus einem beinahe senkrechten Anstieg auf der rechten Seite. Je stärker eine Lorenzkurve $L$ von der Diagonalen abweicht, desto größer ist die Ungleichheit innerhalb der Verteilung der Merkmale auf einzelne Merkmalsträger. Mit anderen Worten: je größer die Abweichung der Lorenzkurve von der Diagonalen, desto stärker ist die relative Konzentration innerhalb der betrachteten Gesamtheit.

# Disparität

## Gini-Koeffizient

Ein Maß für die Abweichung der Lorenzkurve $L$ von der Diagonalen $D$ (gewissermaßen den 'Bauch' der Lorenzkurve) ist der Gini-Koeffizient. Im extremen Grenzfall, der in der Realität allerdings nicht auftreten kann, entspricht dieser Bauch gerade der gesamten Fläche unter der Diagonalen und damit der Fläche eines Dreiecks.

Der Gini-Koeffizient $G$ misst die Fläche zwischen der Diagonalen $D$ und der Lorenzkurve $L$ und setzt sie ins Verhältnis zur Fläche des Dreiecks unter der Diagonalen, die wegen der Konstruktion über kumulierte relative Werte den Betrag $1/2$ aufweist:

$$G := \frac{\text{Fläche zwischen } D \text{ und } L}{\text{Dreiecksfläche unter } D} = \frac{\text{Fläche zwischen } D \text{ und } L}{1/2} \tag{4}$$
$$= 2 \cdot \text{Fläche zwischen } D \text{ und } L$$

# Disparität

Im Falle völliger Gleichverteilung - bei der natürlich keine Konzentration vorliegt - weicht die Lorenzkurve nicht von der Diagonalen ab. Die Fläche zwischen der Diagonalen und der Lorenzkurve ist in diesem Fall Null, der Wert des GINI-Koeffizienten ist damit ebenfalls 0.

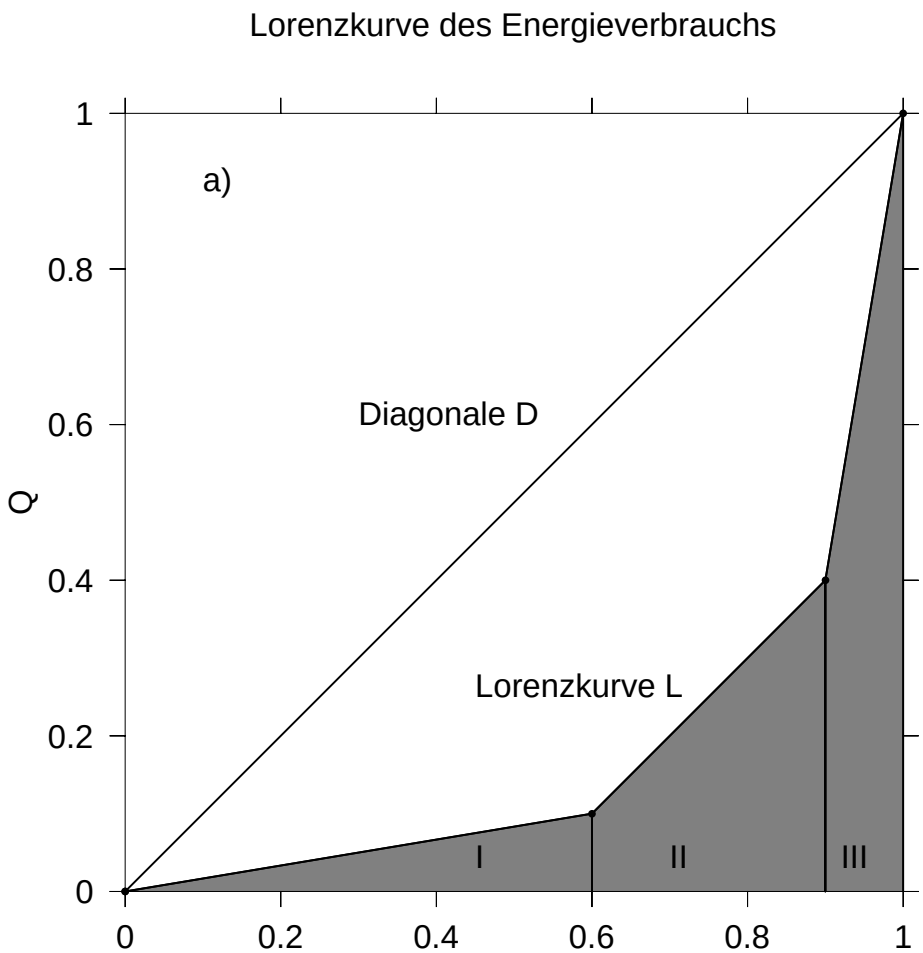
Im Falle einer extremen Ungleichverteilung kommt die Fläche zwischen der Diagonalen $D$ und der Lorenzkurve $L$ der Dreiecksfläche unter $D$ sehr nahe, allerdings ohne sie jemals zu erreichen. Durch die Division durch die Zahl 1/2 (d.h. die Multiplikation mit 2) besitzt der GINI-Koeffizient die folgende Bandbreite:

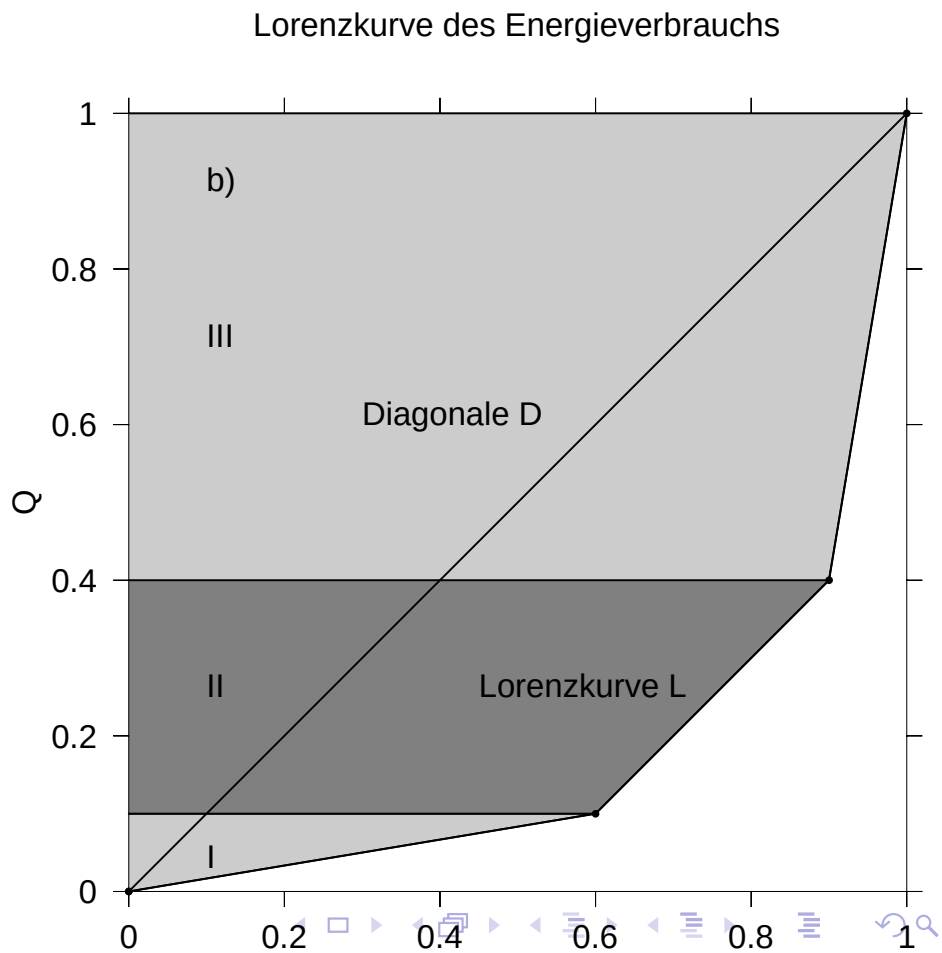
$$0 \leq G < 1.$$

Er ist eine geeignete Kennahl, die als Prozentwert der Konzentration gelesen werden kann.

# Disparität

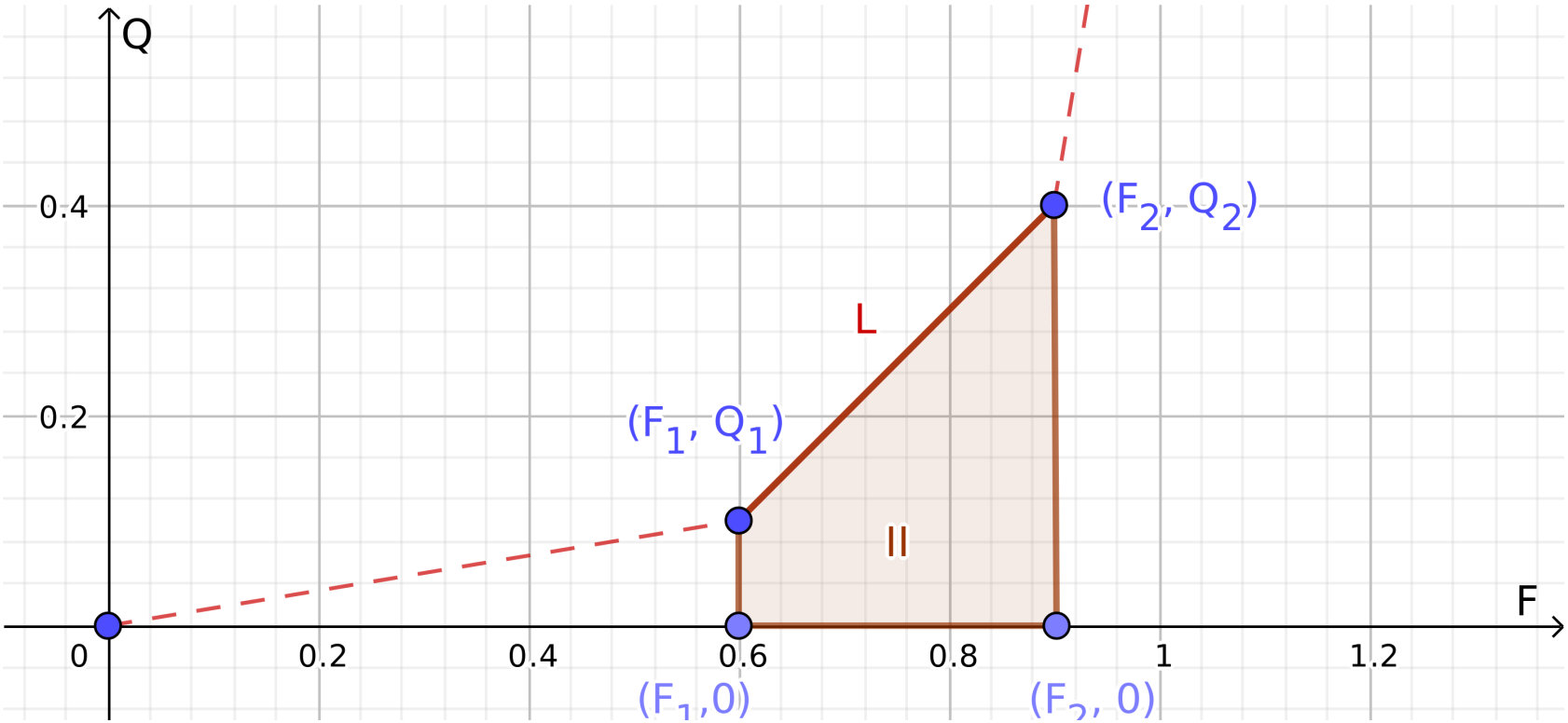
Die Fläche zwischen $D$ und $L$, und damit den GINI-Koeffizienten, gewinnt man aus der Fläche des Dreiecks unter der Diagonalen $D$ durch Subtraktion der Flächen aller Trapeze, die unterhalb der Lorenzkurve liegen.



# Disparität

Die Fläche der einzelnen Trapeze errechnet sich aus der Länge der Grundseite $F_i - F_{i-1} = f_i$, multipliziert mit der durchschnittlichen Höhe $(Q_i + Q_{i-1})/2$.



# Disparität

Die Fläche der einzelnen Trapeze errechnet sich aus der Länge der Grundseite $F_i - F_{i-1} = f_i$, multipliziert mit der durchschnittlichen Höhe $(Q_i + Q_{i-1})/2$.

$$
\begin{aligned}
G &= 2 \cdot \left( \frac{1}{2} - \text{ Summe der Flächen unter } L \text{ (Trapeze)} \right) \\
&= \frac{2}{2} - 2 \cdot \sum_{i=1}^{n} f_i \frac{Q_{i-1} + Q_i}{2} \\
&= 1 - \sum_{i=1}^{n} f_i (Q_{i-1} + Q_i) \qquad (5)
\end{aligned}
$$

Dabei gilt für die Eckpunkte der Diagonalen $D$ immer $Q_0 = 0, F_0 = 0$ und $Q_n = 1, F_n = 1$.

## Beispiel: Ungleichheit beim fiktiven Energieverbrauch auf Gliese 581c.

Der GINI-Koeffizient für unser Beispiel des fiktiven Energieverbrauchs der Bevölkerung des Exoplaneten ergibt sich nach der hergeleiteten Formel (5) zu

$$\begin{aligned} G &= 1 - \sum_{i=1}^{n} f_i(Q_{i-1} + Q_i) \\ &= 1 - (0.6 \cdot 0.1 + 0.3 \cdot (0.1 + 0.4) + 0.1 \cdot (0.4 + 1)) \\ &= 1 - 0.6 \cdot 0.1 - 0.3 \cdot (0.1 + 0.4) - 0.1 \cdot (0.4 + 1) = 0.65 \end{aligned}$$

Die relative Konzentration beträgt 65%. Auch dieser Wert ist vor allem im direkten Vergleich von Märkten sinnvoll.

# Disparität

## Beispiel: Marktmacht innerhalb einer Branche

Fünf Hersteller erzielten in einem Jahr Jahr die folgenden Umsätze:

| | $U_1$ | $U_2$ | $U_3$ | $U_4$ | $U_5$ | Summe |
|---|---|---|---|---|---|---|
| Umsatz | 600 | 1500 | 900 | 1800 | 1200 | 6000 |

Ordnet man die Unternehmen nach der Größe des Umsatzes und ermittelt die Anteile an der Gesamtheit sowie die Marktanteile (MA) in absoluter und kumulierter Form, so ergibt sich das folgende Bild:

| | Umsatz | $f_i = \frac{1}{n}$ | $F_i = i \cdot \frac{1}{n}$ | MA | kum. MA | $(F_i; Q_i)$ |
|---|---|---|---|---|---|---|
| $U_1(i=1)$ | 600 | 20% | 20% | 10% | 10% | $(0.2; 0.10)$ |
| $U_2(i=2)$ | 900 | 20% | 40% | 15% | 25% | $(0.4; 0.25)$ |
| $U_3(i=3)$ | 1200 | 20% | 60% | 20% | 45% | $(0.6; 0.45)$ |
| $U_4(i=4)$ | 1500 | 20% | 80% | 25% | 70% | $(0.8; 0.70)$ |
| $U_5(i=5)$ | 1800 | 20% | 100% | 30% | 100% | $(1.0; 1.00)$ |

## Beispiel: Marktmacht innerhalb einer Branche

Aus diesen Daten kann nun mithilfe der Formel (5) der gesuchte Wert des GINI-Koeffizienten berechnet werden. Benötigt werden lediglich die Anteile an der Gesamtheit $f_i$ (im Beispiel sind alle $f_i = 20\%$) sowie die kumulierten Marktanteile $Q_i$:

$$G = 1 - \sum_{i=1}^{n} f_i(Q_{i-1} + Q_i)$$

weil alle $f_i$ gleich

$$\begin{aligned}
&= 1 - f_i \cdot \sum_{i=1}^{n}(Q_{i-1} + Q_i) \\
&= 1 - 0.2 \cdot (0.1 + (0.1 + 0.25) + (0.25 + 0.45) \\
&+ (0.45 + 0.7) + (0.7 + 1)) \\
&= 1 - 0.2 \cdot 2 \cdot (0.1 + 0.25 + 0.45 + 0.7 + 1/2) = 0.2
\end{aligned}$$

# Disparität

## Beispiel: Marktmacht innerhalb einer Branche

Der Wert $G = 0.2$ des GINI-Koeffizienten deutet auf eine relativ geringe Konzentration des Umsatzes in der Branche hin. Dabei muss aber berücksichtigt werden, dass der maximale Wert $G_{\max}$ des GINI-Koeffizienten von der Zahl der untersuchten statistischen Einheiten $n$ abhängt, für unseren Fall von 5 Unternehmen ergibt sich ein maximaler GINI-Koeffizient von $G_{\max}(5) = 0.8$.

Diesen Wert würde man erhalten, wenn sich der gesamte Umsatz der Branche auf ein einziges Unternehmen konzentrierte.

# Disparität

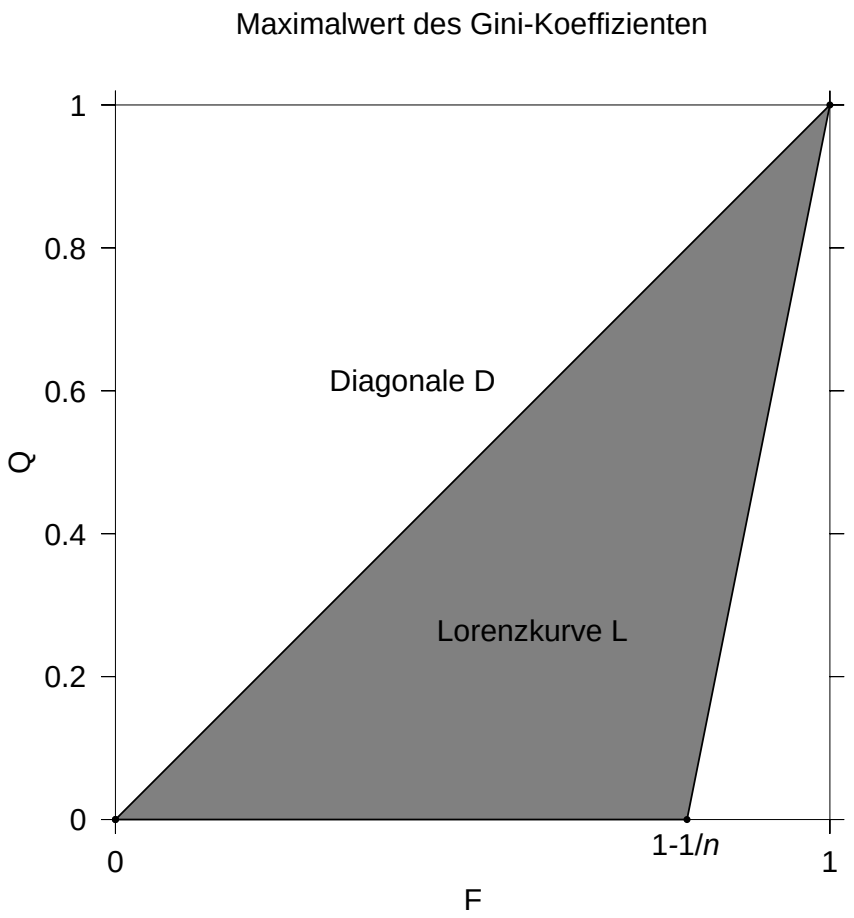



**Abbildung:** zur Bestimmung des Maximalwerts des GINI-Koeffizienten

Die Abbildung zeigt eine Lorenzkurve für den (wenig realistischen) Fall, in dem sich der Gesamtwert eines Merkmals auf eine einzelne aus $n$ untersuchten Merkmalsträgern konzentriert.

Der GINI-Koeffizient lässt sich relativ leicht berechnen, indem von der Fläche $1/2$ des Dreiecks unterhalb der Diagonalen $D$ die Fläche des Dreiecks unter der Lorenzkurve $L$ abgezogen wird

# Disparität

$$G_{\max}(n) = 2 \cdot \left( \frac{1}{2} - \frac{1}{2} \cdot \frac{1}{n} \cdot 1 \right) = 1 - \frac{1}{n} = \frac{n-1}{n}. \tag{6}$$

Damit bewegt sich der Wert des GINI-Koeffizienten $G$ im Intervall

$$0 \leq G \leq \frac{n-1}{n}.$$

Der Wert des GINI-Koeffizienten erreicht die 1 bei einer endlichen Zahl untersuchter Einheiten nie, selbst wenn sich das untersuchte Merkmal auf eine dieser Einheiten konzentriert. Im Beispiel der $n = 5$ Unternehmen kann der GINI-Koeffizient maximal $(5-1)/5 = 0.8$ erreichen, wenn sich der gesamte Umsatz auf ein einzelnes Unternehmen konzentriert - intuitiv erwartet man bei vollständiger Konzentration aber einen Wert von 1. Der berechnete Wert suggeriert eine zu geringe Konzentration.

# Disparität

Diese Überlegung legt die Bildung eines normierten GINI-Koeffizienten $G_{\text{norm}}$ nahe, der durch Division des GINI-Koeffizienten $G$ durch seinen Maximalwert $G_{\text{max}}$ gebildet wird:

**Definition: normierter GINI-Koeffizient**

$$G_{\text{norm}} := \frac{G}{G_{\text{max}}(n)} = \frac{n}{n-1} \cdot G \tag{7}$$

Damit gilt für den Wertebereich des normierten GINI-Koeffizienten

$$0 \leq G_{\text{norm}} \leq 1$$

$$\text{wobei } G_{\text{norm}} = \begin{cases} 0 & \text{bei bei gleichmäßiger Verteilung der Merkmalswerte} \\ 1 & \text{bei vollständiger Konzentration} \end{cases}$$

# Disparität

Mit Hilfe des normierten GINI-Koeffizienten lässt sich die relative Konzentration bzw. Ungleichheit zwischen Stichproben unterschiedlichen Umfangs $n$ miteinander vergleichen. Bei Stichproben mit großem Umfang ist die Normierung des GINI-Koeffizienten oft nicht nötig, denn es gilt

$$\frac{n}{n-1} \longrightarrow 1 \text{ für } n \longrightarrow \infty,$$

der normierte GINI-Koeffizient strebt für große $n$ gegen den Wert des GINI-Koeffizienten. Eine Normierung kann außerdem nur dann vorgenommen werden, wenn der Umfang der Stichprobe $n$ bekannt ist - beispielsweise auf Basis einer relativen Häufigkeitsverteilung kann zwar der GINI-Koeffizient $G$, nicht aber der normierte GINI-Koeffizient $G_{\text{max}}$ berechnet werden.

# Disparität

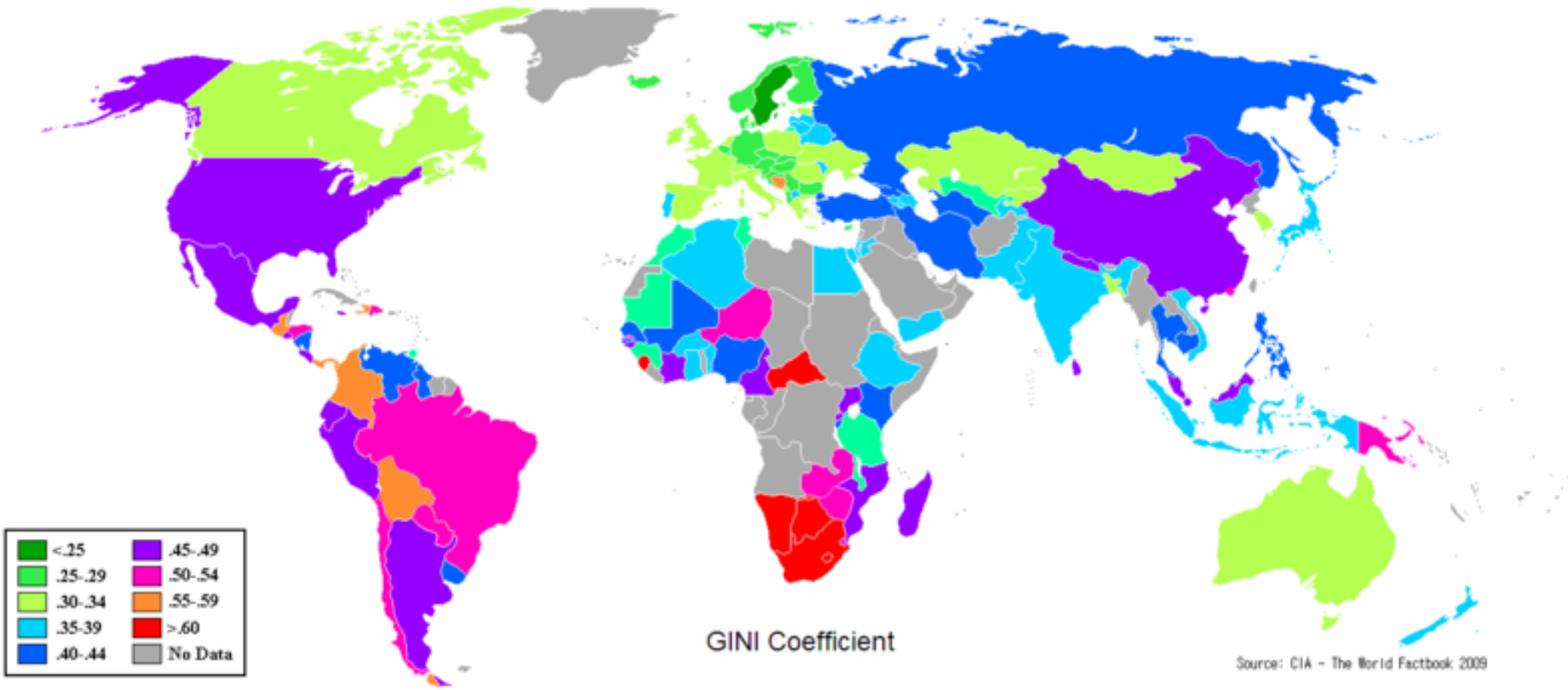



**Abbildung:** Anwendung des Gini-Koeffizienten: Disparität der Einkommensverteilung pro Familie

# Disparität

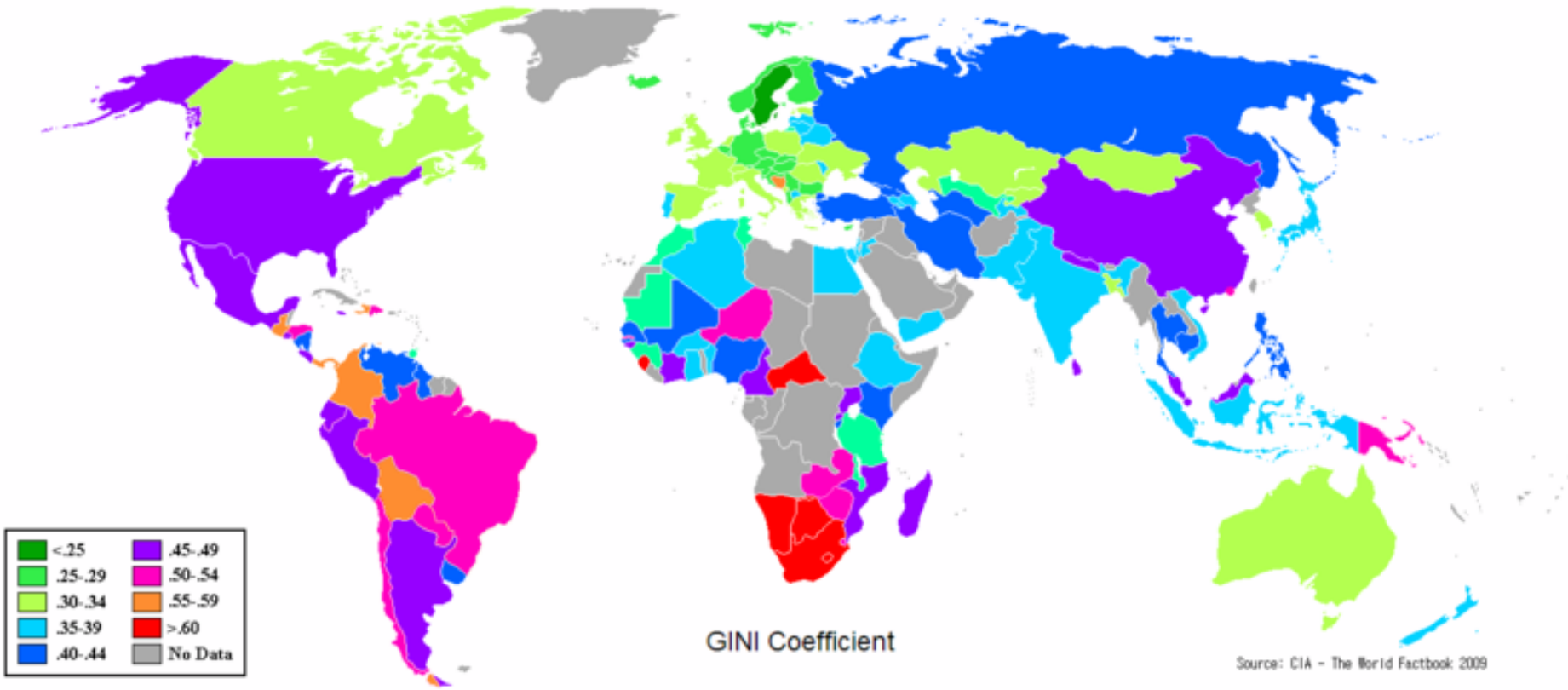



**Abbildung:** Anwendung des Gini-Koeffizienten: Disparität der Einkommensverteilung pro Familie (CIA Factbook, 2009)