

Konzentrationsmaße für Häufigkeitsverteilungen

Typische Fragestellungen: betreffen Analyse der Marktsituation, wer sind die Marktführen?

Konzentrationsverhältnis (concentration ratio)

Beschreibt die Marktkonzentration in einem definierten Marktsegment. Die folgende Tabelle beschreibt die Produktionsverhältnisse in ausgewählten Industriebereichen in Großbritannien in den 90er Jahren (Marktanteile der 5 größten Firmen)

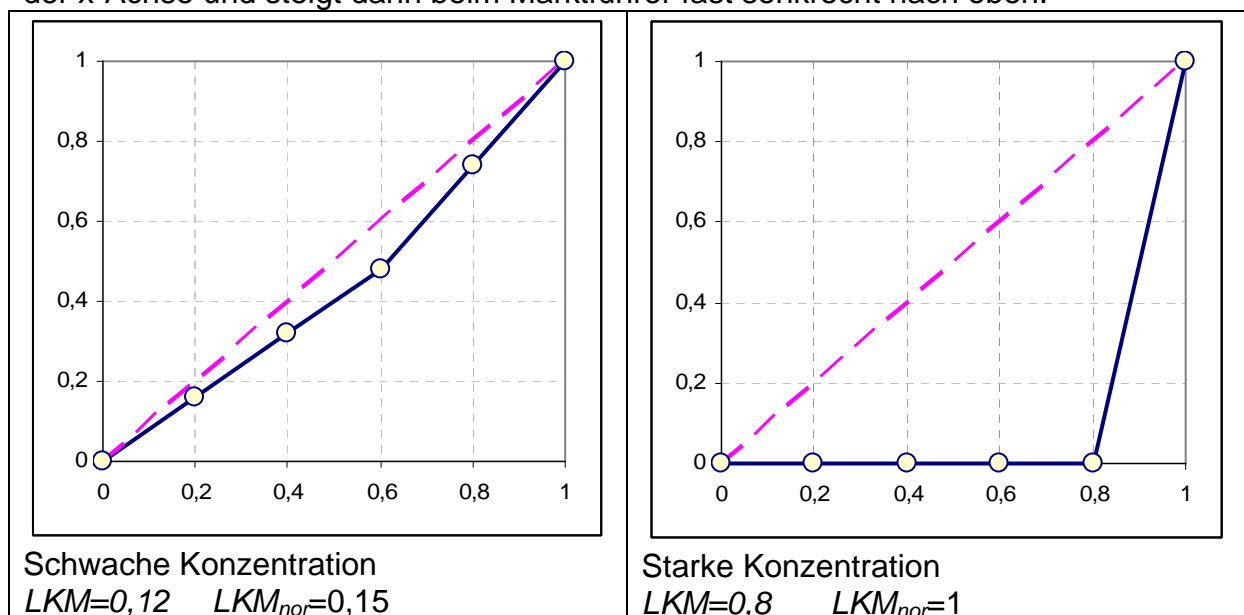
Census of Production code no.	Industry	5-firm concentration index (%)
429	Tobacco industry	99.1 %
427	Brewing and malting	45.1 %
344	Telecommunications	33.1 %
453	Clothing	16.4 %

Diese Kennzahl ist leicht zu berechnen, die Zahl 5 ist willkürlich. So gibt es Österreich nur 5 „echte“ Mobilfunkanbieter (d.h. der Index beträgt hier notwendigerweise 100%); aber mehrere hundert Autohändler.

Lorenzkurve

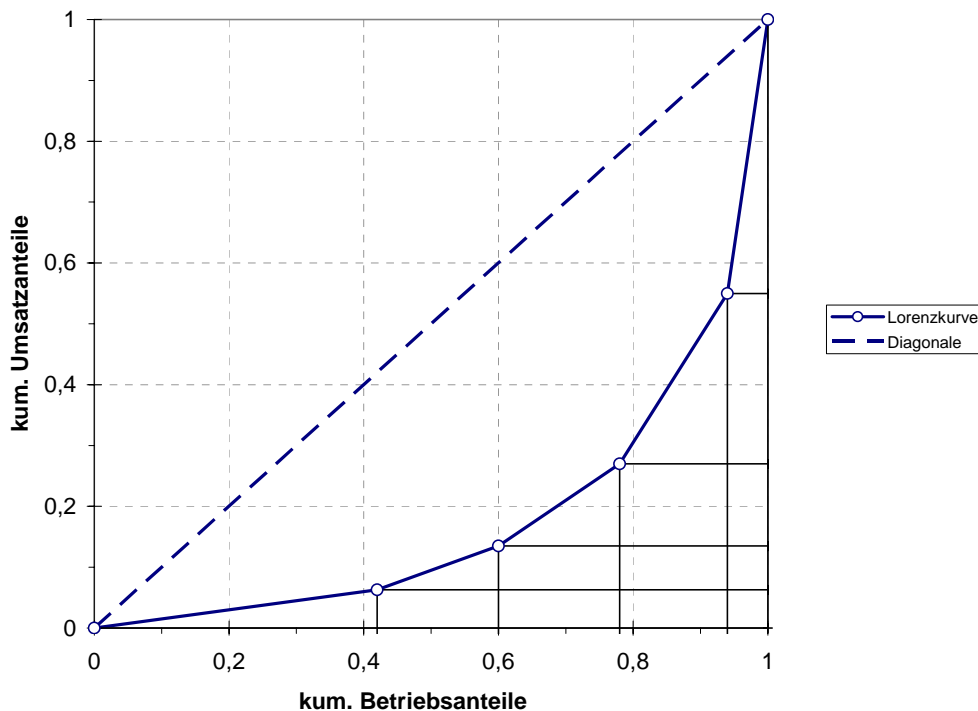
Diese Kurve hat eine größere Aussagekraft als das Konzentrationsverhältnis. Die Daten werden nach dem interessierenden Merkmal geordnet (hier nach dem Umsatz pro Gruppe), relative und kumulierte Häufigkeiten werden gebildet, und die Werte der beiden letzten Spalten werden als Kurve in einem Diagramm aufgetragen. Das Diagramm wird um die Diagonale als Referenzlinie ergänzt.

Wenn der Umsatz gleichmäßig auf alle Betriebe aufgeteilt ist (= keine Konzentration), dann fällt die Lorenzkurve mit der Diagonale zusammen. In einer monopolähnlichen Situation (=hohe Konzentration) bleibt die Lorenzkurve nahe bei der x-Achse und steigt dann beim Marktführer fast senkrecht nach oben.



durchschnittlicher Umsatz	Anzahl der Betriebe	Umsatz pro Gruppe	Rel. Häufigkeit Betriebe	Umsatzanteile	kum. Betriebsanteile	kum. Umsatzanteile
u_i	H_i	$u_i \cdot H_i$	h_i	$u_i \cdot H_i / \sum u_i \cdot H_i$	k_i	l_i
3	21	63	0,42	0,063	0,42	0,063
8	9	72	0,18	0,072	0,6	0,135
15	9	135	0,18	0,135	0,78	0,27
35	8	280	0,16	0,28	0,94	0,55
150	3	450	0,06	0,45	1	1
	n=50	1000	1	1		

Auf 42% der Betriebe entfallen also nur 6,3% des gesamten Umsatzes.



Gini-Koeffizient Lorenzsche Konzentrationsmaß LKM

Mit Hilfe der Lorenzkurve kann man eine Maßzahl für die Konzentration bilden, die bei vollständiger Konzentration den Wert nahe bei 1, bei einer Gleichverteilung der Zielgröße den Wert 0 annimmt. Das Lorenzsche Konzentrationsmaß wird definiert als der zweifache Flächeninhalt zwischen der Diagonalen und der Lorenzkurve. (Bei einer maximalen Konzentration besteht die Fläche aus einem Dreieck mit der Fläche $\frac{1}{2}$.)

$LKM = 2F$

Die Fläche berechnet man am einfachsten als Dreiecksfläche minus die Summe der Flächen zwischen Lorenzkurve und x-Achse (bestehend aus einem Dreieck und mehreren Trapezen).

Für das Lorenzsche Konzentrationsmaß, bekannt als Gini-Koeffizient, gilt

$0 \leq LKM \leq \frac{n-1}{n}$, d.h. die maximale Konzentration ist kleiner als 1 und hängt von der

Anzahl der Merkmale ab

Aus diesem Grund wird der Gini-Koeffizient normiert

$$LKM_{nor} = \frac{n}{n-1} LKM$$

Daraus ergibt sich die Formel

$$LKM_{nor} = 2 \cdot \frac{n}{n-1} \left(\frac{1}{2} - \sum_{i=1}^n \left(\frac{l_{i-1} + l_i}{2} \right) \cdot (k_{i-1} - k_i) \right)$$

und für das Beispiel: $LKM = 0,6408$ $LKM_{nor} = 0,801$

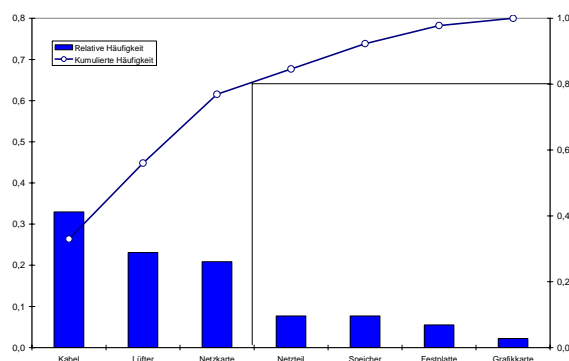
Pareto-Diagramm

Dieses Diagramm beruht auf dem Paretoprinzip, welches besagt, dass die meisten Auswirkungen eines Problems nur auf eine kleine Anzahl von Ursachen zurückzuführen sind. Als idealtypisch gilt dabei die so genannte 80/20-Regel; 20 % der Kunden sorgen für 80 % des Umsatzes. Das Verfahren kann gleichermaßen für Kunden- und Lieferantenbewertung, oder Fehleranalysen (beruhend auf eine Fehlersammelkarte) eingesetzt werden. Die ABC-Analyse ist eine Modifikation dieses Verfahrens; dort wird etwa im Fall der Kundenbewertung (Kundenwert) eine Kategorisierung in drei Klassen durchgeführt: mit lediglich 20 % der Kunden werden bereits 80 % des Umsatzes erzielt (A-Kunden = hohe Bedeutung), 30 % der Kunden bringen 15 % des Umsatzes (B-Kunden = mittlere Bedeutung) und von 50 % der Kunden kommen nur 5 % des Umsatzes (C-Kunden = geringe Bedeutung).

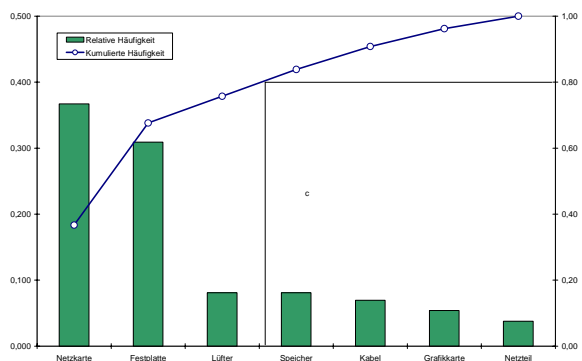
Aus jeder Kategorie wird die relative Häufigkeit ermittelt. Die Kategorien werden absteigend sortiert; eine Summenkurve (=kumulierte Häufigkeiten) wird erstellt. Bei der klassischen Fehleranalyse werden ein Stabdiagramm oder Histogramm (abhängig von der Datenskala) und eine Summenkurve (entspricht in etwa einer approximierenden Verteilungsfunktion) in einem Diagramm kombiniert.

Beispiel: In einem Betrieb werden Fehlerart und Kosten pro Fehler registriert:

Fehlerart	Kosten pro Fehler	Anzahl der Fehler	rel. Häufigkeit	Gesamtkosten	Summenhäufigkeit
Kabel	3	30	0,330	90	0,330
Lüfter	5	21	0,231	105	0,560
Netzkarte	25	19	0,209	475	0,769
Netzteil	7	7	0,077	49	0,846
Speicher	15	7	0,077	105	0,923
Festplatte	80	5	0,055	400	0,978
Grafikkarte	35	2	0,022	70	1,000
Summe:		91	1	1294	



Kriterium: Fehlerhäufigkeit



Kriterium: Gesamtkosten