

Formelsammlung Mathematik für Klausuren

Nur für den Gebrauch zur Vorlesung in QR-Methoden und
für die Klausuren bestimmt

© Jürgen Meisel (2024)

<http://www.juergenmeisel.de>

e-Mail: info@juergenmeisel.de

Funktionen mit mehreren unabhängigen Variablen

Funktion mit mehreren Variablen:

Eine reelle Funktion mit mehreren Variablen $(x_1, x_2, x_3, \dots, x_n)$ ist definiert als eine

$$f: \mathbb{R}^n \rightarrow \mathbb{R} \quad \text{mit } n \geq 2 \quad \text{und } n \in \mathbb{N}$$

Abbildung $(x_1, x_2, x_3, \dots, x_n) \rightarrow x_{n+1} = f(x_1, x_2, x_3, \dots, x_n)$

Partielle Ableitung:

Unter der partiellen Ableitung (1. Ordnung) der Funktion $f(x_1, x_2, x_3, \dots, x_n)$ nach der Variablen x_i versteht man die Ableitung von $f(x_1, x_2, x_3, \dots, x_n)$ nach x_i unter Konstanthaltung aller übrigen Variablen.

Schreibweisen: $\frac{\partial f(x_1, x_2, x_3, \dots, x_n)}{\partial x_i}$ oder $\frac{\partial f}{\partial x_i}$ oder f_{x_i}

Ökonomische Interpretation: Der Wert der partiellen Ableitung gibt an, um wie viel Einheiten sich der Funktionswert $f(x_1, x_2, x_3, \dots, x_n)$ ändert, wenn sich x_i um eine Einheit ändert, während alle übrigen Variablen konstant bzw. unverändert bleiben (ceteris paribus-Bedingung).

Entsprechende Ableitungen höherer Ordnung sind ebenfalls möglich und besitzen folgen-

de Schreibweisen: $\frac{\partial^2 f(x_1, x_2, x_3, \dots, x_n)}{\partial x_i \partial x_k}$ oder $\frac{\partial^2 f}{\partial x_i \partial x_k}$ oder $f_{x_i x_k}$

Satz von Schwarz: Sind für die Funktion $f(x_1, x_2, x_3, \dots, x_n)$ sämtliche zweite Ableitungen stetig, so sind diese unabhängig von der Differentiationsreihenfolge und es

gilt:
$$\frac{\partial^2 f}{\partial x_i \partial x_k} = \frac{\partial^2 f}{\partial x_k \partial x_i} \quad \text{oder} \quad f_{x_i x_k} = f_{x_k x_i}$$

Gradient: Vektor der ersten partiellen Ableitungen

$$\text{grad}(f) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right)$$

Hesse-Matrix: Matrix der zweiten partiellen Ableitungen

$$H(f) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix} = \begin{pmatrix} f_{x_1 x_1} & f_{x_1 x_2} & \dots & f_{x_1 x_n} \\ f_{x_2 x_1} & f_{x_2 x_2} & \dots & f_{x_2 x_n} \\ \dots & \dots & \dots & \dots \\ f_{x_n x_1} & f_{x_n x_2} & \dots & f_{x_n x_n} \end{pmatrix}$$

Fall $n = 2$:

$$H(f) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} \end{pmatrix} = \begin{pmatrix} f_{x_1 x_1} & f_{x_1 x_2} \\ f_{x_2 x_1} & f_{x_2 x_2} \end{pmatrix}$$

Extrema bei Funktionen mit mehreren Variablen ohne Nebenbedingungen:

(1) Stationäre Stelle(n) ermitteln:

$$\text{grad}(f) \stackrel{!}{=} \vec{0} = (0, 0, \dots, 0) \Rightarrow P(x_{0_1}, x_{0_2}, \dots, x_{0_n}, f)$$

(2) Prüfen der stationären Stellen mittels Hesse-Matrix

Hinreichende Bedingung für ein relatives Extremum:

Für den Fall $n = 2$ gilt daher:

$$\text{Maximum} \Leftrightarrow \begin{array}{l} \text{Stationäre Stelle existiert} \\ \text{negativ definit} \end{array} \Leftrightarrow \left\{ \begin{array}{l} f_{x_1 x_1} < 0 \\ \det \begin{pmatrix} f_{x_1 x_1} & f_{x_1 x_2} \\ f_{x_2 x_1} & f_{x_2 x_2} \end{pmatrix} > 0 \end{array} \right.$$

$$\text{Minimum} \Leftrightarrow \begin{array}{l} \text{Stationäre Stelle existiert} \\ \text{positiv definit} \end{array} \Leftrightarrow \left\{ \begin{array}{l} f_{x_1 x_1} > 0 \\ \det \begin{pmatrix} f_{x_1 x_1} & f_{x_1 x_2} \\ f_{x_2 x_1} & f_{x_2 x_2} \end{pmatrix} > 0 \end{array} \right.$$

Allgemein gilt folgendes:

Die Funktion $f(x_1, x_2, x_3, \dots, x_n)$ besitze in $P(x_{0_1}, x_{0_2}, \dots, x_{0_n}, f)$ eine stationäre Stelle d.h. $\text{grad}(f) \stackrel{!}{=} \vec{0} = (0, 0, \dots, 0)$, und habe in $P(x_{0_1}, x_{0_2}, \dots, x_{0_n}, f)$ stetige partielle Ableitungen zweiter Ordnung.

Aus der Definitheit der Hesse-Matrix $H(f)$ der Funktion $f(x_1, x_2, x_3, \dots, x_n)$ folgt:

Ist die Hesse-Matrix $H(f)$ in $P(x_{0_1}, x_{0_2}, \dots, x_{0_n}, f)$

- positiv definit**, dann hat $f(\dots)$ in $P(x_{0_1}, x_{0_2}, \dots, x_{0_n}, f)$ ein rel. Minimum.
- negativ definit**, dann hat $f(\dots)$ in $P(x_{0_1}, x_{0_2}, \dots, x_{0_n}, f)$ ein rel. Maximum.
- indefinit**, dann hat $f(\dots)$ in $P(x_{0_1}, x_{0_2}, \dots, x_{0_n}, f)$ einen Sattelpunkt.

Definition zur Definitheit der Hesse-Matrix $H(f)$:

$$(i) \quad H(f) \text{ positiv definit} \Leftrightarrow \det(H_i(f)) > 0$$

$$(ii) \quad H(f) \text{ negativ definit} \Leftrightarrow (-1)^i \det(H_i(f)) > 0$$

$H_i(f)$ sind die jeweiligen Teil- bzw. Untermatrizen der Hesse-Matrix; von diesen werden die sogenannten Hauptminoren bzw. Hauptabschnittsdeterminanten berechnet:

Beispiel für den Fall $n=3$:

$$\begin{aligned} |H_1(f)| &= |f_{x_1 x_1}| \wedge |H_2(f)| = \begin{vmatrix} f_{x_1 x_1} & f_{x_1 x_2} \\ f_{x_2 x_1} & f_{x_2 x_2} \end{vmatrix} \\ \wedge |H_3(f)| &= \begin{vmatrix} f_{x_1 x_1} & f_{x_1 x_2} & f_{x_1 x_3} \\ f_{x_2 x_1} & f_{x_2 x_2} & f_{x_2 x_3} \\ f_{x_3 x_1} & f_{x_3 x_2} & f_{x_3 x_3} \end{vmatrix} \end{aligned}$$

Extrema bei Funktionen mit mehreren Variablen mit Nebenbedingungen:

- Lösungsverfahren: (1) Variablensubstitution
(2) Lagrangeansatz

Lagrangeansatz für den Fall $n=2$ und eine Nebenbedingung:

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda \cdot g(x_1, x_2)$$

$f(x_1, x_2)$ ist die Zielfunktion $g(x_1, x_2)$ ist die Nebenbedingung

Vorgehensweise:

Stationäre Stelle ermitteln \Rightarrow Austauschverhältnis ermitteln \Rightarrow Berechnen der Extrema unter Verwendung der Nebenbedingung

Anmerkung: notwendige Bedingung genügt bei konvexen Zielfunktionen bzw. Nebenbedingungen.

Lineare Optimierung

Lösung eines Maximierungsproblems mittels regulärer Simplexmethode bzw. mit dem Simplexalgorithmus

- (1) Aufstellen des linearen Ungleichungssystems aus m Ungleichungen mit n Problemvariablen (= Nebenbedingungen) und der Zielfunktion
- (2) Umformung des linearen Ungleichungssystems in ein lineares Gleichungssystem durch Einführung je einer Schlupfvariablen (Basisvariablen) pro Ungleichung
- (3) Erstellung eines Tableaus (LGS aus Nebenbedingungen und Zielfunktion) mit Hilfe der Koeffizienten der Variablen
- (4) Pivot-Spalte mittels größtem negativem Koeffizient in der Zielfunktion ermitteln, dann mit der Spalte die Zeile mit dem größten Engpass (d.h. kleinster positiver Quotient aus Ergebnisspalte und Koeffizienten der Pivot-Spalte) anhand der Restriktionen berechnen (\Rightarrow Pivot-Zeile)
Anmerkung: Sind mehrere Koeffizienten gleich groß, dann kann einer frei gewählt werden.
- (5) Pivot-Element bestimmen (Kombination aus Pivot-Spalte und Pivot-Zeile)
- (6) Pivot-Element mittels Gauß-Verfahren zu 1 umformen
- (7) Alle anderen Elemente in der Pivot-Spalte mittels Gauß-Verfahren zu 0 umformen
- (8) Mit dem nächsten größtem negativem Koeffizient der Zielfunktion fortsetzen und das Verfahren wie bisher wiederholen
- (9) Ist kein Wert der Zielfunktion mehr positiv, dann ist die Lösung optimal

Sollte ein Minimierungsproblem vorliegen, dann muss dies per Dualisierung in ein Maximierungsproblem umgewandelt werden, damit das Verfahren anwendbar ist.

Formelsammlung Statistik

Mittelwerte und Streuungsmaße bei Einzelwerten, Klassen und stetigen Werten

	Einzelwerte (vorwiegend diskret)	Klassierte Werte
Arithmetisches Mittel	<p>Einzelwerte: $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$</p> <p>Gewichtet auf Basis relativer Häufigkeit:</p> $\bar{x} = \sum_{i=1}^k \left(x_i \cdot \frac{n_i}{n} \right) = \sum_{i=1}^k (x_i \cdot p_i)$ <p>mit $p_i = \frac{n_i}{n} =$ relative Häufigkeit</p>	$\bar{x} = \sum_{i=1}^k \left[(x_i)_m \cdot \frac{n_i}{n} \right] = \sum_{i=1}^k \left[(x_i)_m \cdot p_i \right]$ <p>mit $(x_i)_m$ als Klassenmitte der Klasse i</p> <p>und $p_i = \frac{n_i}{n} =$ relative Häufigkeit der Klasse i</p>
Median / Zentralwert	$\bar{x}_M = \begin{cases} x_{\frac{n+1}{2}} & \text{für } n \text{ ungerade} \\ \frac{1}{2} \left(x_{\frac{n}{2}} + x_{\frac{n}{2}+1} \right) & \text{für } n \text{ gerade} \end{cases}$	$\bar{x}_M \Rightarrow \bar{x}_{0,5} = [a; b] \rightarrow \bar{x}_{0,5} = a + \frac{\Delta_i \cdot [0,5 - F(a)]}{p_i}$ <p>$\Delta_i =$ Klassenbreite $p_i =$ rel. Häufigkeit</p> <p>$F(a) =$ kum. rel. Häufigkeit bis Medianintervall $[a; b]$</p>
Modus / Modalwert	<i>Häufigster Wert einer Verteilung.</i>	<p>Schritt 1: Modale Klasse bestimmen</p> <p>⇒ Klasse mit max. Häufigkeitsdichte</p> <p>Schritt 2:</p> <p>⇒ Wert der Klassenmitte der modalen Klasse</p>
Spannweite	$w = \max(x_i) - \min(x_i)$	$w = \text{Klassengrenze}_{\max} - \text{Klassengrenze}_{\min}$ <p>Differenz zwischen größter und kleinster Klassengrenze</p>
Interquartil-abstand	IQA = Q3 - Q1	IQA = Q3 - Q1

	Einzelwerte (vorwiegend diskret)	Klassierte Werte
Mittlere absolute Abweichung (vom Median)	$d = \frac{1}{n} \sum_{i=1}^n x_i - \bar{x}_M $	$d = \sum_{i=1}^k \left[(x_i)_m - \bar{x}_M \cdot \frac{n_i}{n} \right] = \sum_{i=1}^k \left[(x_i)_m - \bar{x}_M \cdot p_i \right]$ <p>mit $(x_i)_m$ als Klassenmitte der Klasse i</p> <p>und $p_i = \frac{n_i}{n}$ = relative Häufigkeit der Klasse i</p>
Varianz und Standardabweichung (Basis: arithmetisches Mittel)	$V(X) \quad \left. \begin{array}{l} \text{Definitions} \\ = \\ \text{formel} \end{array} \right\} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ $V(X) \quad \left. \begin{array}{l} \text{Rechen} \\ = \\ \text{formel} \end{array} \right\} \left[\frac{1}{n} \sum_{i=1}^n (x_i)^2 \right] - \bar{x}^2$ $\Rightarrow S(X) = \sqrt{V(X)}$	$V(X) \quad \left. \begin{array}{l} \text{Definitions} \\ = \\ \text{formel} \end{array} \right\} \sum_{i=1}^n \left[(x_i)_m - \bar{x} \right]^2 \cdot p_i$ $V(X) \quad \left. \begin{array}{l} \text{Rechen} \\ = \\ \text{formel} \end{array} \right\} \left[\sum_{i=1}^n (x_i)_m^2 \cdot p_i \right] - \bar{x}^2$ $\Rightarrow S(X) = \sqrt{V(X)}$ <p>mit $(x_i)_m$ als Klassenmitte der Klasse i</p> <p>und p_i als relative Häufigkeit der Klasse i</p>
Quartile Q1 / Q3 (Basis: Median)	$\bar{x}_p = \begin{cases} x_{[n \cdot p]+1} & \text{für } (n \cdot p) \\ & \text{nicht ganzzahlig} \\ \frac{1}{2} (x_{n \cdot p} + x_{n \cdot p + 1}) & \text{für } (n \cdot p) \text{ ganzzahlig} \end{cases}$ <p>$p = 0,25 \rightarrow q1$ und $p = 0,75 \rightarrow q3$</p>	$q_1 \Rightarrow \bar{x}_{0,25} = [a; b] \rightarrow \bar{x}_{0,25} = a + \frac{\Delta_i \cdot [0,25 - F(a)]}{p_i}$ $q_3 \Rightarrow \bar{x}_{0,75} = [a; b] \rightarrow \bar{x}_{0,75} = a + \frac{\Delta_i \cdot [0,75 - F(a)]}{p_i}$ <p>Δ_i = Klassenbreite p_i = rel. Häufigkeit</p> <p>$F(a)$ = kum. rel. Häufigkeit bis q_1/q_3 - Intervall $[a; b]$</p>
Variationskoeffizient	$v_{arith_Mittel} = \frac{S(X)}{\bar{x}}$	<p>oder</p> $v_{Median} = \frac{S(X)}{\bar{x}_M}$

Allgemeine Darstellung für n Summanden:

$$\begin{aligned} V\left(\sum_{k=1}^n a_k X_k\right) &= V(a_1 X_1 + \dots + a_n X_n) \\ &= a_1^2 \cdot V(X_1) + \dots + a_n^2 \cdot V(X_n) + 2a_1 a_2 \cdot \text{Cov}(X_1; X_2) + 2a_1 a_3 \cdot \text{Cov}(X_1; X_3) + \dots + 2a_{n-1} a_n \cdot \text{Cov}(X_{n-1}; X_n) \\ &= \sum_{k=1}^n a_k^2 \cdot V(X_k) + 2 \sum_{\substack{k=1; \\ k < l}}^n a_k a_l \cdot \text{Cov}(X_k; X_l) \end{aligned}$$

Allgemeine Darstellung für n Summanden:

$\text{Cov}(X; Y) = 0 \Leftrightarrow X; Y$ unabhängig :

$$V\left(\sum_{k=1}^n a_k X_k\right) = V(a_1 X_1 + \dots + a_n X_n) = a_1^2 \cdot V(X_1) + \dots + a_n^2 \cdot V(X_n) = a_k^2 \cdot V\left(\sum_{k=1}^n X_k\right)$$

Diskrete Zufallsvariable - **Varianz** (einer Grundgesamtheit):

⇒ Zusammenhang zwischen Definition und Berechnung

$$V(X) \stackrel{\text{Def.}}{=} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \stackrel{\text{Berechnung}}{=} \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2$$

Begriff der **Kovarianz / Co-Varianz** (einer Grundgesamtheit)

$$\text{Cov}(X, Y) \stackrel{\text{Def.}}{=} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \stackrel{\text{Berechnung}}{=} \frac{1}{n} \sum_{i=1}^n (x_i \cdot y_i) - \bar{x} \cdot \bar{y}$$

Gini-Koeffizient: Berechnung der Flächen mit relativen kumulierten Häufigkeiten:

$$\text{Gini}(GK) = \frac{\text{Konzentrationsfläche } K}{\text{maximale Konzentrationsfläche } K_{\max}}$$

$$\text{Gini}(GK) = \frac{\text{Fläche zw. Gleichvert. und Lorenzkurve}}{\text{Fläche unter Gleichverteilung}}$$

Anmerkung: Als „Konzentrationsfläche“ K bezeichnet man die Fläche, zwischen der Gleichverteilungsdiagonalen und der LORENZ-Kurve: $0 \leq K \leq 1/2$.

Wegen $K_{\max} = \frac{1}{2} - \frac{1}{2n}$ lautet der normierte Gini – Koeffizient :

$$\text{norm. Gini} = K \cdot \frac{2n}{n-1}$$

Der Wertebereich des Gini-Koeffizienten liegt zwischen 0 (= Gleichverteilung) und 1 (= vollständige Konzentration auf einen Merkmalsträger).

Erläuterung zur Berechnung des Gini-Koeffizienten

Fläche unter der Gleichverteilungsgeraden: $A_{\Delta} = \frac{1}{2} \cdot g \cdot h = \frac{1}{2} \cdot 1 \cdot 1 = \frac{1}{2}$

Fläche unterhalb der Lorenzkurve: Dreiecke $A_{Dreieck} = \frac{1}{2} \cdot g \cdot h$

und Trapeze $A_{Trapez} = \frac{1}{2} \cdot (h_i + h_{i+1}) \cdot \Delta x_i$

Berechnung des Gini-Koeffizienten

Maximale Konzentrationsfläche => Fläche unterhalb der Gleichverteilungsgerade

$$A_{\Delta} = \frac{1}{2} \cdot g \cdot h = \frac{1}{2} \cdot 1 \cdot 1 = \frac{1}{2}$$

Fläche unterhalb der Lorenzkurve:

$$A_1 = \frac{1}{2} \cdot g \cdot h \rightarrow \frac{1}{2} \cdot 0,5 \cdot 0,025 = 0,00625$$

$$A_2 = \frac{1}{2} \cdot (h_1 + h_2) \cdot g \rightarrow \frac{1}{2} \cdot (0,025 + 0,5) \cdot 0,4 = 0,105$$

$$A_3 = \frac{1}{2} \cdot (h_1 + h_2) \cdot g \rightarrow \frac{1}{2} \cdot (0,5 + 0,77) \cdot 0,09 = 0,05715$$

$$A_4 = \frac{1}{2} \cdot (h_1 + h_2) \cdot g \rightarrow \frac{1}{2} \cdot (0,77 + 1) \cdot 0,01 = 0,00885$$

Fläche unterhalb der Lorenzkurve:

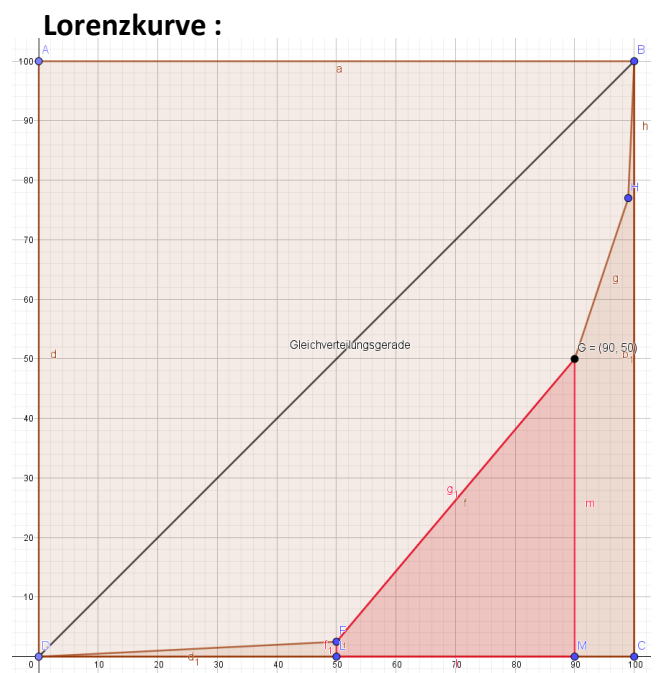
$$A_1 + A_2 + A_3 + A_4 = 0,17725$$

Fläche zwischen Gleichverteilung und Lorenzkurve:

$$0,5 - 0,17725 = 0,32275$$

$$GK = \frac{\text{Fläche zw. Gleichvert. und Lorenzkurve}}{\text{Fläche unter Gleichverteilung}}$$

$$GK = \frac{0,5 - 0,17725}{0,5} = \frac{0,32275}{0,5} = 0,6455$$



Auswertung:

keine Konzentration $\Leftrightarrow GK \in [0; 0,3]$

(mäßige) Konzentration $\Leftrightarrow GK \in]0,3; 0,7[$

vollständige Konzentration $\Leftrightarrow GK \in [0,7; 1]$

Grundlage: Verteilungstabelle

Bevölkerung			Vermögen		
Prozentanteil pro Klasse	rel. Anteil	kumulierte Prozentwerte	Prozentanteil pro Klasse	rel. Anteil	kumulierte Prozentwerte
0,00	0,00	0,00	0,00	0,00	0,000
50,00	0,50	0,50	2,50	0,03	0,025
40,00	0,40	0,90	47,50	0,48	0,500
9,00	0,09	0,99	27,00	0,27	0,770
1,00	0,01	1,00	23,00	0,23	1,000
100,00	1,00	x-Achse	100,00	1,00	y-Achse

$\frac{\Delta y}{\Delta x} = \text{Steigung}$
 $\rightarrow \text{Sekantensteigung}$
 $\rightarrow \text{Differenzenquotient}$

0,05
1,1875
3
23

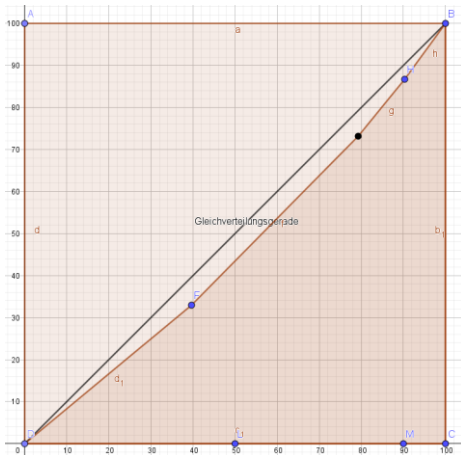
Ordnen der Werte: (gemäß) Differenzenquotient

$\frac{\Delta y}{\Delta x} = (\text{durchschnittliche}) \text{Steigung} = \text{Sekantensteigung}$

"von klein nach groß"

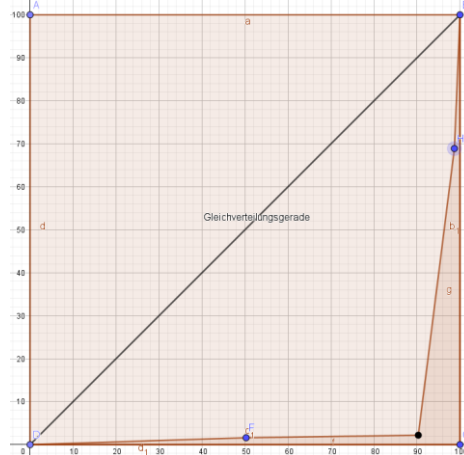
richtige
Reihenfolge der
Anordnung

Sonderformen:



$$GK = \frac{\text{Fläche zw. Gleichvert. und Lorenzkurve}}{0,5} = \frac{\rightarrow 0}{0,5} \rightarrow 0$$

$\rightarrow GK \rightarrow 0 \rightarrow \text{nahezu Gleichverteilung}$



$$GK = \frac{\text{Fläche zw. Gleichvert. und Lorenzkurve}}{0,5} = \frac{\rightarrow 0,5}{0,5} \rightarrow 1$$

$\rightarrow GK \rightarrow 1 \rightarrow \text{vollständige Konzentration}$

Preisindizes

nach Laspeyres:
$$L_P = \frac{\sum p_{1i} \cdot q_{0i}}{\sum p_{0i} \cdot q_{0i}} \quad \text{mit } i \in N$$

nach Paasche:
$$P_P = \frac{\sum p_{1i} \cdot q_{1i}}{\sum p_{0i} \cdot q_{1i}} \quad \text{mit } i \in N$$

nach Fisher:
$$F_P = \sqrt{L_P \cdot P_P} \quad \text{geometr. Mittel}$$

Anmerkung: Geometrisches Mittel / Geometrischer Mittelwert

Geometr. Mittel: Mittelwert einer Produktfolge

$$g_{MW} = \sqrt[n]{\prod_{i=1}^n q_i} = \sqrt[n]{q_1 \cdot q_2 \cdot \dots \cdot q_n} \quad \text{mit } q_i = 1 + \frac{p_i}{100}$$

Effektive Verzinsung einer tesaurierenden

(Zinsen werden kapitalisiert, d.h. dem Kapital hinzugerechnet) Anlage

=> Effektive Verzinsung bei Zinseszinsseffekt

$$i_{eff} = \sqrt[n]{\prod_{i=1}^n q_i} - 1 = \sqrt[n]{q_1 \cdot q_2 \cdot \dots \cdot q_n} - 1 = g_{MW} - 1$$

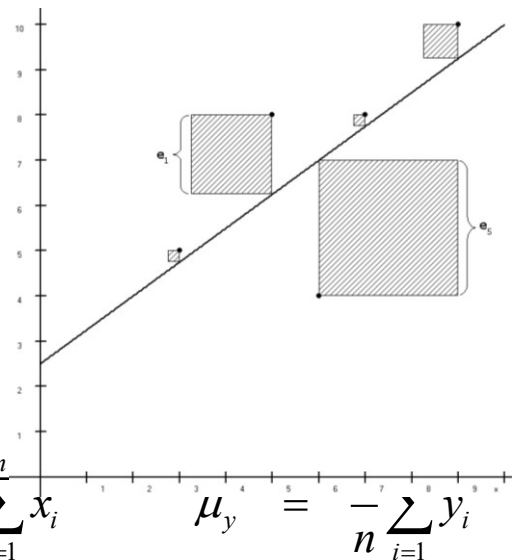
$$p_{eff} = (g_{MW} - 1) \cdot 100 = i_{eff} \cdot 100$$

Lineare Regression und Korrelation

Anwendung der partiellen Differentiation in der Statistik: Regressionsgerade

Gesucht: *Eine Gerade, die zu allen Punkten den kleinsten Abstand hat; d.h. die Differenzen e_i müssen insgesamt möglichst klein sein (Optimierungsproblem)*

Ansatz:
$$\sum_{i=1}^n (e_i)^2 \rightarrow \min.$$



Lineare Regression und Korrelation

Ansatz:
$$y = b_0 + b_1 x \quad \mu_x = \frac{1}{n} \sum_{i=1}^n x_i \quad \mu_y = \frac{1}{n} \sum_{i=1}^n y_i$$

$$b_1 = \frac{\sigma_{xy}}{\sigma_x^2} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i y_i) - \mu_x \cdot \mu_y}{\frac{1}{n} \sum_{i=1}^n (x_i^2) - \mu_x^2} = \frac{Kov(X, Y)}{V(X)} \quad \text{und} \quad b_0 = \mu_y - b_1 \cdot \mu_x$$

Korrelation nach Brevais-Pearson:

$$r \stackrel{\text{Definition}}{=} \frac{\sum_{i=1}^n (x_i - \mu_x) \cdot (y_i - \mu_y)}{\sqrt{\sum_{i=1}^n (x_i - \mu_x)^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \mu_y)^2}} \stackrel{\text{Berechnung}}{=} \frac{\frac{1}{n} \cdot \sum_{i=1}^n (x_i y_i) - \mu_x \cdot \mu_y}{\sqrt{\frac{1}{n} \cdot \sum_{i=1}^n (x_i^2) - \mu_x^2} \cdot \sqrt{\frac{1}{n} \cdot \sum_{i=1}^n (y_i^2) - \mu_y^2}}$$
$$r = \frac{\text{Kov}(X, Y)}{S(X) \cdot S(Y)} = \frac{\sigma(X, Y)}{\sigma(X) \cdot \sigma(Y)}$$

$$\text{Rangkorrelation nach Spearman} \quad r_s = 1 - \frac{6 \sum_{i=1}^n (d_i)^2}{n(n^2 - 1)}$$

$$r = [-1; 1] \quad \text{möglicher Wertebereich}$$

$$r = [-1; -0,7[\quad \text{sehr starke negative Korrelation}$$

$$r = [-0,7; -0,3[\quad \text{starke negative Korrelation}$$

$$r = [-0,3; 0,3[\quad \text{keine Korrelation (Punktwolke)}$$

$$r = [0,3; 0,7[\quad \text{starke positive Korrelation}$$

$$r = [0,7; 1] \quad \text{sehr starke positive Korrelation}$$
