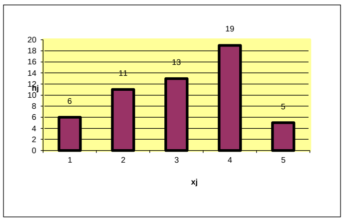
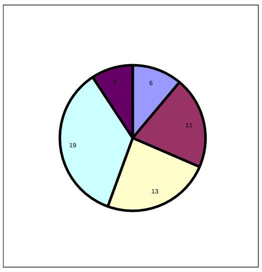
1. ---

2. Nominal skaliert, da die zugeordneten Zahlen willkürlich sind.
Modus M = 4, da 4 am häufigsten auftritt.
$f_j = h_j / 54$.





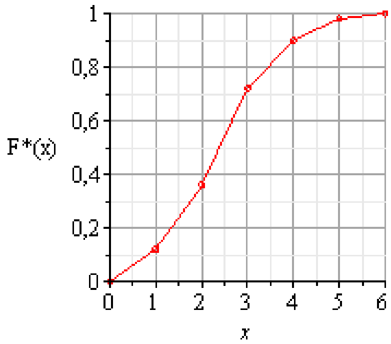
3. Ordinal skaliert, da die Zahlen verglichen werden können.

1. Quartil: $n \cdot p = n \cdot \frac{1}{4} = 12{,}5$ nicht ganzzahlig. Also $Q_1 = a_{[n \cdot p]+1} = a_{13} = 2$, $\bar{x} = \frac{146}{50} = 2{,}92$.

2. Quartil =Median = Zentralwert Z: $n \cdot p = n \cdot \frac{1}{2} = 25$ ganzzahlig.. Also $Z = \frac{a_{n \cdot p} + a_{n \cdot p+1}}{2} = \frac{a_{25} + a_{26}}{2} = 3$

3. Quartil: $n \cdot \frac{3}{4} = 37{,}5$. Also $Q_3 = a_{[37{,}5]+1} = a_{38} = 4$.

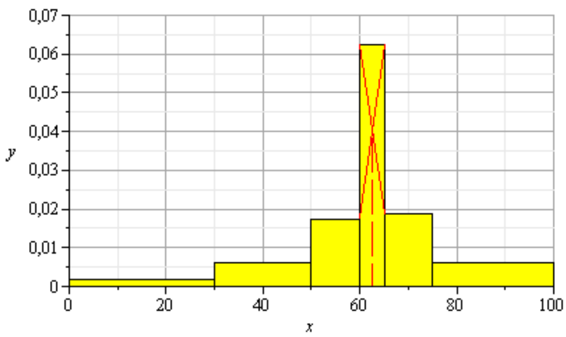
j	x_j	h_j	f_j	H_j	F_j
1	1	6	0,12	6	0,12
2	2	12	0,24	18	0,36
3	3	18	0,36	36	0,72
4	4	9	0,18	45	0,90
5	5	4	0,08	49	0,98
6	6	1	0,02	50	1,00
	$\sum$	50	1		

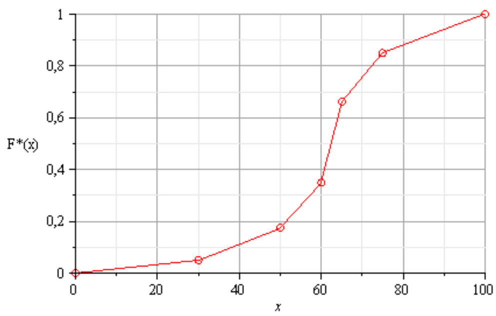


4. Beispiel einer Verhältnisskala

Aus einer Sendung wurden 80 Tomaten herausgegriffen und ihre Masse in Gramm bestimmt.

Von .... bis unter ...	h_j	w_j	x_j	$h_j \cdot x_j$	$f_j = \frac{h_j}{n}$	$f_j^* = \frac{f_j}{w_j}$	F_j
0 – 30	4	30	15	60	0,05	0,001667	0,05
30 – 50	10	20	40	400	0,125	0,00625	0,175
50 – 60	14	10	55	770	0,175	0,0175	0,35
60 – 65	25	5	62,5	1562,5	0,3125	0,0625	0,6625
65 – 75	15	10	70	1050	0,1875	0,01875	0,85
75 – 100	12	25	87,5	1050	0,15	0,006	1
$\sum$	n = 80	100	–	4892,5	1	–	–





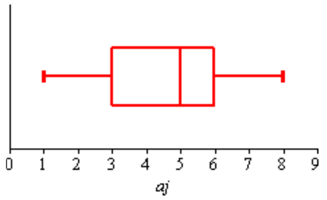
Aus dem Histogramm folgt Modus $D \approx 62{,}6$. Aus dem rechten Diagramm $Q_1 \approx 52$, $Z \approx 62$, $Q_3 \approx 70$.

Bestimmen Sie den Mittelwert $\bar{x} = \frac{1}{n}\sum_{j=1}^{m} h_j \cdot x_j = 61{,}15625 \approx 61{,}2$.

5. Die Diagonale bedeutet, dass alle Personen das gleiche Einkommen haben.
Die mittlere Kurve zeigt eine leichte Konzentration des Einkommens bei Personen mit höherem Einkommen.
Die untere Kurve zeigt eine starke Konzentration des Einkommens bei Personen mit höherem Einkommen.

6. Gegeben sind die 25 Daten 1, 1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 6, 6, 6, 6, 7, 7, 7, 8.
Bestimmen Sie den Modus, die drei Quartile und das arithmetische Mittel und zeichnen Sie einen Box-Plot.

$Q_1 = a_7 = 3$, $Q_2 = Z = a_{13} = 5$, $Q_3 = a_{19} = 6$, Modus $= (4+5)/2 = 4{,}5$. $\bar{x} = 113/25 = 4{,}52$.

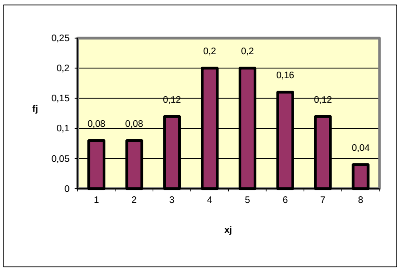


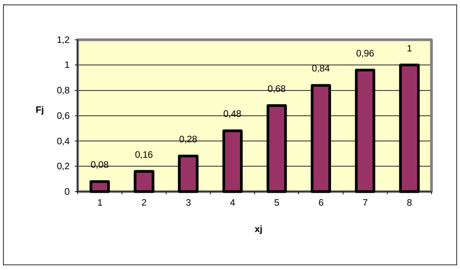
$$Q = \begin{cases} a_{[n \cdot p]+1} & \text{falls} \quad n \cdot p \quad \text{nicht ganzzahlig} \\ \dfrac{a_{n \cdot p} + a_{n \cdot p+1}}{2} & \text{falls} \quad n \cdot p \quad \text{ganzzahlig} \end{cases}$$

Erstellen Sie eine Tabelle wie auf Seite 8 des Skripts.

j	x_j	h_j	$h_j \cdot x_j$	f_j	$f_j \cdot x_j$	F_j	$f_j \cdot x_j^2$
1	1	2	2	0,08	0,08	0,08	0,08
2	2	2	4	0,08	0,16	0,16	0,32
3	3	3	9	0,12	0,36	0,28	1,08
4	4	5	20	0,2	0,8	0,48	3,2
5	5	5	25	0,2	1	0,68	5
6	6	4	24	0,16	0,96	0,84	5,76
7	7	3	21	0,12	0,84	0,96	5,88
8	8	1	8	0,04	0,32	1	2,56
$\sum$	–	25	113	1	$\bar{x}$ **= 4,52**	–	23,88

Zeichnen Sie die Diagramme für die relative Häufigkeit und die kumulierte relative Häufigkeit.





Bestimmen Sie die Varianz s^2 und die Standardabweichung s.

$$s^2 = \overline{x^2} - \bar{x}^2 = \sum_{j=1}^{8} f_j \cdot x_j^2 - \bar{x}^2 = 23{,}88 - 4{,}52^2 = 3{,}4496, \text{ also } s = \sqrt{3{,}4496} \approx 1{,}86 .$$

7. Es sei $x_1, x_2 > 0$. Aus $\dfrac{1}{\bar{x}_{harm}} = \dfrac{1}{2} \cdot \left(\dfrac{1}{x_1} + \dfrac{1}{x_2} \right)$ folgt $\dfrac{1}{\bar{x}_{harm}} = \dfrac{1}{2} \cdot \dfrac{x_1 + x_2}{x_1 \cdot x_2}$, also $\bar{x}_{harm} = 2 \cdot \dfrac{x_1 \cdot x_2}{x_1 + x_2}$.

Andererseits ist $\bar{x}_{geom} = \sqrt{x_1 \cdot x_2}$.

① Aus dem Ansatz $2 \cdot \dfrac{x_1 \cdot x_2}{x_1 + x_2} \le \sqrt{x_1 \cdot x_2}$ folgt nach Kürzen mit $\sqrt{x_1 \cdot x_2}$ die Ungleichung $2 \cdot \dfrac{\sqrt{x_1 \cdot x_2}}{x_1 + x_2} \le 1$

und folglich, $2 \cdot \sqrt{x_1 \cdot x_2} \le x_1 + x_2$, bzw. $0 \le x_1 - 2 \cdot \sqrt{x_1 \cdot x_2} + x_2$. Das bedeutet $0 \le \left(\sqrt{x_1} - \sqrt{x_2}\right)^2$.

② Aus dem Ansatz $\sqrt{x_1 \cdot x_2} \le \dfrac{x_1 + x_2}{2}$ folgt $2 \cdot \sqrt{x_1 \cdot x_2} \le x_1 + x_2$, d.h. $0 \le \left(\sqrt{x_1} - \sqrt{x_2}\right)^2$ wie bei ①.

Das Gleichheitszeichen gilt im Fall $0 = \left(\sqrt{x_1} - \sqrt{x_2}\right)^2$, d.h. $x_1 = x_2$.

8. Nach diesen n Jahren beträgt das Kapital $K_n = K_0 \cdot \prod_{j=1}^{n}\left(1+\frac{p_j}{100}\right)$. Und das soll $K_0 \cdot \left(1+\frac{\bar{p}}{100}\right)^n$ sein.

Daraus folgt $K_0 \cdot \left(1+\frac{\bar{p}}{100}\right)^n = K_0 \cdot \prod_{j=1}^{n}\left(1+\frac{p_j}{100}\right)$, bzw. $1+\frac{\bar{p}}{100} = \sqrt[n]{\prod_{j=1}^{n}\left(1+\frac{p_j}{100}\right)} = \sqrt[n]{\prod_{j=1}^{n} q_j} = \bar{q}_{geom}$, so dass $\bar{p} = \left(\bar{q}_{geom} - 1\right)\cdot 100$.

Im Beispiel gilt $\bar{q}_{geom} = \sqrt[4]{1{,}02\cdot 1{,}03\cdot 1{,}04\cdot 1{,}05} \approx 1{,}035$, so dass $\bar{p} \approx (1{,}035-1)\cdot 100 = 3{,}5$, also im Mittel 3,5% Zins.

9. In der Zeit t füllt das j-te Rohr den Bruchteil $\frac{t}{t_j}$ des Beckens, d.h. dass alle n Rohre zusammen in der Zeit t den Bruchteil $\sum_{j=1}^{n}\frac{t}{t_j} = t\cdot\sum_{i=1}^{n}\frac{1}{t_j}$ des Beckens füllen.

Das Becken ist voll, wenn $t\cdot\sum_{i=1}^{n}\frac{1}{t_j} = 1$, d.h. $t = \frac{1}{\sum_{i=1}^{n}\frac{1}{t_j}} = \frac{1}{n}\cdot\frac{n}{\sum_{i=1}^{n}\frac{1}{t_j}} = \frac{1}{n}\cdot\bar{t}_{harm}$.

Beispiel: Es gilt $t = \frac{1}{\frac{1}{20}+\frac{1}{25}+\frac{1}{30}} = \frac{300}{37} \approx 8{,}108$ Stunden.

Außerdem ist $\bar{t}_{harm} = \frac{3}{\frac{1}{20}+\frac{1}{25}+\frac{1}{30}} = \frac{900}{37} \approx 24{,}32$ Stunden.

Das arithmetische Mittel ist aber $\bar{t}_{arith} = \frac{20+25+30}{3} = 25$.

10. Zusammensetzen von drei Messreihen

1. Messreihe	1,1	1,2	–	–	–
2. Messreihe	0,9	1,0	0,8	–	–
3. Messreihe	1,0	0,8	1,1	1,0	1,1

Bestimmen Sie die drei Mittelwerte $\bar{x}_1$, $\bar{x}_2$ und $\bar{x}_3$ und die drei Varianzen s_1^2, s_2^2 und s_3^2.

$\bar{x}_1 = 1{,}15$, $\bar{x}_2 = 0{,}90$, $\bar{x}_3 = 1{,}00$.

$s_1^2 = 0{,}0025$, $s_2^2 = \frac{1}{150} \approx 0{,}00667$, $s_3^2 = 0{,}012$.

Bestimmen Sie $\bar{x}$ und s^2 für die drei Messreihen zusammen mit Hilfe der Formeln $\bar{x} = \frac{1}{n}\cdot\sum_{j=1}^{m} n_j\cdot\bar{x}_j$ und $s^2 = s_{int}^2 + s_{ext}^2$ mit $s_{int}^2 = \frac{1}{n}\cdot\sum_{j=1}^{m} n_j\cdot s_j^2$ und $s_{ext}^2 = \frac{1}{n}\cdot\sum_{j=1}^{m} n_j\cdot\left(\bar{x}_j - \bar{x}\right)^2$.

$\bar{x} = \frac{1}{10}\cdot(2\cdot 1{,}15 + 3\cdot 0{,}9 + 5\cdot 1) = 1$.

$s_{int}^2 = \frac{1}{10}\cdot\left(2\cdot 0{,}0025 + 3\cdot\frac{1}{150} + 5\cdot 0{,}012\right) = 0{,}0085$.

$s_{ext}^2 = \frac{1}{10}\cdot\left(2\cdot(1{,}15-1)^2 + 3\cdot(0{,}9-1)^2 + 5\cdot(1-1)^2\right) = 0{,}0075$.

Also $s^2 = 0{,}0085 + 0{,}0075 = 0{,}016$.

Bestimmen Sie $\bar{x}$ und s^2 für die drei Messreihen zusammen mit Hilfe der 10 Daten.

$\bar{x} = \frac{1}{10} \cdot 10 = 1$ und $s^2 = \frac{1}{10} \cdot \left(0{,}1^2 + 0{,}2^2 + 0{,}1^2 + 0^2 + 0{,}2^2 + 0^2 + 0{,}2^2 + 0{,}1^2 + 0^2 + 0{,}1^2\right) = 0{,}016$, oder

$s^2 = \overline{x^2} - \bar{x}^2 = \frac{1}{10}\left(1{,}1^2 + 1{,}2^2 + 0{,}9^2 + 1{,}0^2 + 0{,}8^2 + 1{,}0^2 + 0{,}8^2 + 1{,}1^2 + 1{,}0^2 + 1{,}1^2\right) - 1 = 1{,}016 - 1 = 0{,}016$.

11. a. Aus $20 \cdot 85 + 10 \cdot x = 30 \cdot 80$ folgt $x = 70$.

b. Aus $70p + 92n = 86(p+n)$ folgt $70p + 92n = 86p + 86n$, also $6n = 16p$, so dass $\frac{p}{n} = \frac{3}{8}$.

12. Nochmals zu den 80 Tomaten aus Aufgabe 4

Von .... bis unter ...	w_j	h_j	f_j	$f_j \cdot w_j^2$	x_j	$f_j \cdot x_j$	$(x_j - \bar{x})^2$	$f_j \cdot (x_j - \bar{x})^2$
0 – 30	30	4	0,05	45	15	0,75	2130,39941	106,519971
30 – 50	20	10	0,125	50	40	5	447,586914	55,9483643
50 – 60	10	14	0,175	17,5	55	9,625	37,8994141	6,63239746
60 – 65	5	25	0,3125	7,8125	62,5	19,53125	1,80566406	0,56427002
65 – 75	10	15	0,1875	18,75	70	13,125	78,2119141	14,6647339
75 – 100	25	12	0,15	93,75	87,5	13,125	693,993164	104,098975
$\sum$	100	80	1	232,8125	–	$\bar{x} = 61{,}15625$	–	288,42871

$\bar{x} = \sum_{j=1}^{m} f_j \cdot \bar{x}_j = 61{,}16$. $s_{int}^2 = \sum_{j=1}^{m} f_j \cdot \frac{w_j^2}{12} = \frac{232{,}8125}{12} = 19{,}4010$, $s_{ext}^2 = \sum_{j=1}^{m} f_j \cdot \left(\bar{x}_j - \bar{x}\right)^2 = 288{,}42871$, also

$s^2 = s_{int}^2 + s_{ext}^2 \approx 307{,}8297$ und $s \approx 17{,}545$.

13. Zehn Firmen wurden nach der Höhe ihres Umsatzes in Mio. Euro befragt. Es ergab sich die Tabelle:

j	Umsatz x_j	absolute Häufigkeit h_j	relative Häufigkeit f_j	kumulierte relative Häufigk. F_j	relative Merkmalssumme MS. $\ell_j = \frac{h_j \cdot x_j}{\sum_{k=1}^{4} h_k \cdot x_k}$	kumulierte rel. MS. $L_j = \sum_{k=1}^{j} \ell_k$
1	5	2	0,2	0,2	0,05	0,05
2	10	1	0,1	0,3	0,05	0,10
3	15	2				
4	30	5				
$\sum$		10	1	–	1	–

Tragen Sie L_j über F_j auf und zeichnen Sie die Lorenzkurve.

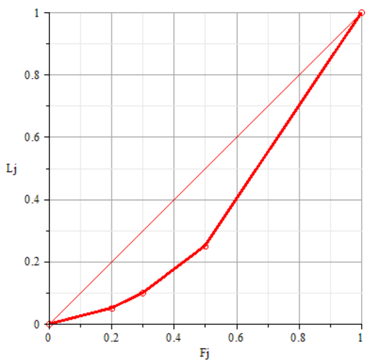
Dabei ist F_j der Prozentsatz von der Befragten, deren Umsatz bis zu x_j beträgt.

L_j ist der Prozentsatz des Gesamtumsatzes $\sum_{k=1}^{4} h_k \cdot x_k = 200$ bis zum Umsatz x_j .

Geben Sie Beispiele für die Zuordnungen $L_j \rightarrow F_j$ und umgekehrt.

(Bestimmen Sie den Gini-Koeffizienten G.)

j	Umsatz x_j	absolute Häufigkeit h_j	relative Häufigkeit f_j	kumulierte relative Häufigk. F_j	relative Merkmalssumme MS. $\ell_j = \frac{h_j \cdot x_j}{\sum_{k=1}^{4} h_k \cdot x_k}$	kumulierte rel. MS. $L_j = \sum_{k=1}^{j} \ell_k$
1	5	2	0,2	0,2	0,05	0,05
2	10	1	0,1	0,3	0,05	0,10
3	15	2	0,2	0,5	0,15	0,25
4	30	5	0,5	1,0	0,75	1,00
$\sum$		10	1	–	1	–



Der Flächeninhalt des ersten Dreiecks ($0 \le F_j \le 0,2$) beträgt

$\frac{1}{2} \cdot 0,2 \cdot 0,05 = 0,005$.

Der Flächeninhalt des anschließenden Trapezes ($0,2 \le F_j \le 0,3$) beträgt

$\frac{1}{2} \cdot (0,05 + 0,1) \cdot 0,1 = 0,0075$.

Der Flächeninhalt des anschließenden Trapezes ($0,3 \le F_j \le 0,5$) beträgt

$\frac{1}{2} \cdot (0,1 + 0,25) \cdot 0,2 = 0,035$.

Der Flächeninhalt des letzten Trapezes ($0,5 \le F_j \le 1,0$) beträgt

$\frac{1}{2} \cdot (0,25 + 1,0) \cdot 0,5 = 0,3125$.

Die Summe dieser Flächeninhalte ist $0,36$.
Und damit folgt der Flächeninhalt zwischen der Diagonale und der Lorenzkurve zu $0,5 - 0,36 = 0,14$ und der Gini-Koeffizient zu $G = \frac{0,14}{0,5} = 0,28$.

14. 100 Leute wurden nach der Höhe ihres wöchentlichen Einkommens befragt. Es ergab sich die Tabelle:

j	Einkommen von ... bis unter ...	absolute Häufigkeit h_j	Klassenbreite w_j	Klassenmitte m_j	relative Häufigkeit f_j	Relative Dichte f_j^*	kumulierte relative Häufigk. F_j	relative Merkmalssumme MS. $\ell_j = \frac{h_j \cdot m_j}{\sum_{k=1}^{m} h_k \cdot m_k}$	kumulierte rel. MS. $L_j = \sum_{k=1}^{j} \ell_k$
1	0 – 200	15	200	100	0,15	0,00075	0,15	0,040	0,040
2	200 – 400	50	200	300	0,50	0,0025	0,65	0,395	0,435
3	400 – 600	20							
4	600 – 800	10							
5	800 – 1000	5							
$\sum$		100	1000	–	1	–	–	1	–

Tragen Sie L_j über F_j auf und zeichnen Sie die Lorenzkurve.

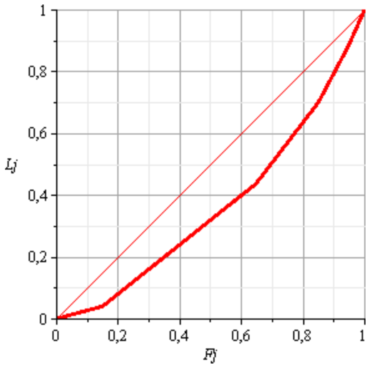
Dabei ist F_j der Prozentsatz von der Befragten, deren Einkommen bis zur Klasse j gehört.

L_j ist der Prozentsatz des Gesamteinkommens $\sum_{k=1}^{m} h_k \cdot m_k$ bis zum Einkommen der Klasse j.

Geben Sie Beispiele für die Zuordnungen $L_j \to F_j$ und umgekehrt.

j	Einkommen von ... bis unter ...	absolute Häufigkeit h_j	Klassenbreite w_j	Klassenmitte m_j	relative Häufigkeit f_j	Relative Dichte f_j^*	kumulierte relative Häufigk. F_j	relative Merkmalssumme MS. $\ell_j = \frac{h_j \cdot m_j}{\sum_{k=1}^{m} h_k \cdot m_k}$	kumulierte rel. MS. $L_j = \sum_{k=1}^{j} \ell_k$
1	0 – 200	15	200	100	0,15	0,00075	0,15	0,040	0,040
2	200 – 400	50	200	300	0,50	0,0025	0,65	0,395	0,435
3	400 – 600	20	200	500	0,20	0,001	0,85	0,263	0,698
4	600 – 800	10	200	700	0,10	0,0005	0,95	0,184	0,882
5	800 – 1000	5	200	900	0,05	0,00025	1,00	0,118	1,000
$\sum$		100	1000	–	1	–	–	1	–

Das Gesamteinkommen beträgt $\sum_{k=1}^{m} h_k \cdot m_k = 38000$.



Beispiele:
40% der Befragten verdienen etwa 24% des Gesamteinkommens,
50% des Gesamteinkommens wird von 70% der Befragten verdient.