Aufgabe: Berechnen Sie den Gini-Koeffizienten und skizzieren Sie den Lösungsweg anhand eines aussagekräftigen Beispiels.

Lösung: Für eine aufsteigend sortierte, diskret verteilte Größe $x=\left(x_1,\ldots,x_n\right)$ berechnet sich die inverse Lorenzkurve F_i als Summenquotient

$$F_i=\frac{\sum_{j=1}^{i}x_j}{\sum_{j=1}^{n}x_j}=\frac{1}{n\overline{x}}\sum_{j=1}^{i}x_j=\sum_{j=1}^{i}f_j,$$

wobei der Mittelwert gegeben ist durch

$$\overline{x}=\frac{1}{n}\sum_{j=1}^{n}x_j\quad\text{mit}\quad f_j=\frac{x_j}{n\overline{x}}.$$

Entsprechend verteilt sich die zugehörige Größe $y=\left(y_1,\ldots,y_n\right)$ auf die Ordinate, womit die Lorenzkurve L_i gegeben durch

$$L_i=\frac{\sum_{j=1}^{i}y_j}{\sum_{j=1}^{n}y_j}=\frac{1}{n\overline{y}}\sum_{j=1}^{i}y_j=\sum_{j=1}^{i}l_j,$$

mit dem Mittelwert

$$\overline{y}=\frac{1}{n}\sum_{j=1}^{n}y_j\quad\text{und}\quad l_j=\frac{y_j}{n\overline{y}}.$$

Es zeigt sich, daß für die Differenz der Koordinaten zweier Punkte auf der Lorenzkurve gilt:

$$F_i-F_{i-1}=\sum_{j=1}^{i}f_j-\sum_{j=1}^{i-1}f_j=f_i+\sum_{j=1}^{i-1}f_j-\sum_{j=1}^{i-1}f_j=f_i$$

und

$$L_i-L_{i-1}=\sum_{j=1}^{i}l_j-\sum_{j=1}^{i-1}l_j=l_i+\sum_{j=1}^{i-1}l_j-\sum_{j=1}^{i-1}l_j=l_i.$$

Mit den Maximalwerten $F_n=L_n=1$ ergibt sich die Darstellung in Abb. 1. Dabei ist die Fläche *A* gegeben durch das Dreieck zwischen der 45°-Linie und der *F*-Achse und die Fläche *B* durch die Polygonfläche unter der Lorenzkurve. Die Differenz aus beiden Flächen ergibt normiert auf die Fläche *A* den Gini-Koeffizienten.

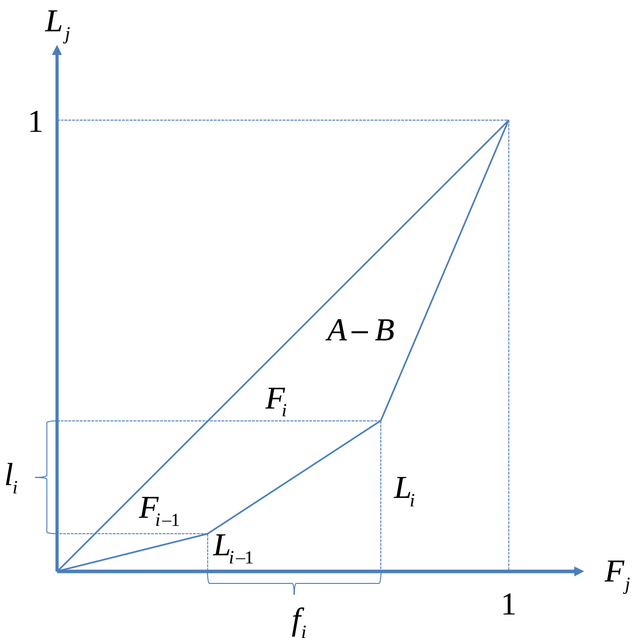

Abbildung 1: Der Gini-Koeffizient für eine diskrete Variable

Nun ist die Fläche eines Trapezes mit zwei äquidistanten Stützstellen $F_i = L_i$ und $F_{i-1} = L_{i-1}$ gegeben durch

$$A_i = \frac{1}{2}\left(L_i + L_{i-1}\right)\left(F_i - F_{i-1}\right) = \left(L_i - \frac{l_i}{2}\right) f_i.$$

Damit berechnet sich die Summe A aller Flächenelemente unter der 45°-Linie wie in Abb. 1 angegeben zu

$$A = \sum_{i=1}^{n} A_i = \sum_{i=1}^{n}\left(L_i - \frac{l_i}{2}\right) f_i = \sum_{i=1}^{n} f_i L_i - \frac{1}{2}\sum_{i=1}^{n} f_i l_i = 1 - \frac{1}{2} = \frac{1}{2}.$$

Die Polygonflächen unter der Lorenzkurve ergeben sich mit $L_i < F_i$ und $L_{i-1} < F_{i-1}$ analog aus

$$B_i = \left(L_i + L_{i-1}\right)\frac{F_i - F_{i-1}}{2} = \left(L_i - \frac{l_i}{2}\right) f_i.$$

Mithin ist

$$B = \sum_{i=1}^{n} B_i = \sum_{i=1}^{n}\left(L_i - \frac{l_i}{2}\right) f_i.$$

Abschließend ein Beispiel aus einer Unternehmensstruktur.[1]

[1] Die Zahlenwerte sind der Wikipedia-Seite https://de.wikipedia.org/wiki/Gini-Koeffizient entnommen.

50 %	der Belegschaft	(f_1)	leisten	2,5 %	der Arbeit	(l_1).
40 %	der Belegschaft	(f_2)	leisten	47,5 %	der Arbeit	(l_2).
9 %	der Belegschaft	(f_3)	leisten	27,0 %	der Arbeit	(l_3).
1 %	der Belegschaft	(f_4)	leistet	23,0 %	der Arbeit	(l_4).

50 Prozent der Belegschaft sind sogenannte Minderleister, weitere 40 Prozent erfüllen die Erwartungen, 9 Prozent übertreffen sie und nur 1 % sind Spitzenleister. Gemäß folgender Tabelle ergeben sich daraus die normierten Größen l_i / f_i :

$$\begin{aligned}
&f_1 = 0,50, \quad l_1 = 0,025, \qquad l_1 / f_1 = 0,05,\\
&f_2 = 0,40, \quad l_2 = 0,475, \qquad l_2 / f_2 = 1,188,\\
&f_3 = 0,09, \quad l_3 = 0,270, \qquad l_3 / f_3 = 3,\\
&f_4 = 0,01, \quad l_4 = 0,230, \qquad l_4 / f_4 = 23.
\end{aligned}$$

Damit nun eine Lorenzkurve zustande kommt, müssen diese Werte so vorsortiert werden, daß gilt:

$$\frac{l_i}{f_i} \geq \frac{l_{i-1}}{f_{i-1}}.$$

Das ist in dem angegebenen Beispiel bereits der Fall, es braucht also nicht mehr vorsortiert zu werden. Die (F_i, L_i)-Paare entstehen dann aus den (f_j, l_j)-Paaren nach der Rechenvorschrift

$$F_i = \sum_{j=1}^{i} f_j \quad \text{und} \quad L_i = \sum_{j=1}^{i} l_j .$$

Im nächsten Schritt werden aus den Differenzen des vorhergehenden Schritts durch Aufsummierung die zusätzlich benötigten Funktionswerte ermittelt, wobei der Nullpunkt noch mit dazukommt, d.h.

$$\begin{aligned}
&F_0 = 0, && L_0 = 0,\\
&F_1 = f_1 = 0,5, && L_1 = l_1 = 0,025,\\
&F_2 = f_1 + f_2 = 0,9, && L_2 = l_1 + l_2 = 0,5,\\
&F_3 = f_1 + f_2 + f_3 = 0,99, && L_3 = l_1 + l_2 + l_3 = 0,77,\\
&F_4 = f_1 + f_2 + f_3 + f_4 = 1, && L_4 = l_1 + l_2 + l_3 + l_4 = 1.
\end{aligned}$$

Die einzelnen Flächenelemente haben damit folgende Werte

$$B_1 = \left(L_1 - \frac{l_1}{2} \right) f_1 = 0,00625, \quad B_2 = \left(L_2 - \frac{l_2}{2} \right) f_2 = 0,105,$$

$$B_3 = \left(L_3 - \frac{l_3}{2} \right) f_3 = 0,05715, \quad B_4 = \left(L_4 - \frac{l_4}{2} \right) f_4 = 0,00885.$$

Die Gesamtfläche ergibt sich wieder durch Summation,

$$B = B_1 + B_2 + B_3 + B_4 = 0{,}17725.$$

Mit $A = 1/2$ erhalten wir schließlich den Gini-Koeffizienten

$$\text{GUK} = \frac{A-B}{A} = 1 - 2B = 0{,}6455.$$

Anmerkung

Nur im Falle $F_i = i/n$, d.h. mit äquidistanten Stützstellen, führt auch der Ausdruck aus [1],

$$A - B = \sum_{i=1}^{n} \left(L_i - L_{i-1}\right) \frac{F_i + F_{i-1}}{2} - \frac{1}{2} = \frac{1}{2} \sum_{i=1}^{n} l_i \left(\frac{i}{n} + \frac{i-1}{n} \right) - \frac{1}{2} \sum_{j=1}^{n} l_j = \frac{1}{2} \sum_{i=1}^{n} l_i \left(\frac{2i-n-1}{n} \right),$$

zu einem eher theoretisch verwertbaren Ergebnis:

$$\begin{aligned} I_G &= \frac{A-B}{A} = \sum_{i=1}^{n} l_i \frac{2i-n-1}{n} = \sum_{i=1}^{n} \frac{x_i}{n\overline{x}} \frac{2i-n-1}{n} \\ &= \sum_{i=1}^{n} \frac{x_i}{n\overline{x}} \frac{2i}{n} - \frac{n+1}{n} \sum_{i=1}^{n} \frac{x_i}{n\overline{x}} = \frac{2}{n} \frac{\sum_{i=1}^{n} i x_i}{\sum_{i=1}^{n} x_i} - \frac{n+1}{n}. \end{aligned}$$

Im Falle maximaler Ungleichheit, d.h. für $x_1 = x_2 = \ldots = x_{n-1} = 0$ und $x_n = 1$ ist

$$I_G = 2 - \frac{n+1}{n} = 1 - \frac{1}{n}.$$

Wenn alle Werte gleich sind, d.h. für $x_1 = x_2 = \ldots = x_n \neq 0$, folgt

$$I_G = \frac{2}{n} \frac{x_1 \sum_{i=1}^{n} i}{n x_1} - \frac{n+1}{n} = \frac{2}{n} \frac{n(n+1)}{2n} - \frac{n+1}{n} = 0.$$

Der Standard-Ginikoeffizient ist damit definiert durch

$$\text{GUK} = \frac{I_G}{1 - \frac{1}{n}} = \frac{n}{n-1} \sum_{i=1}^{n} l_i \frac{2i-n-1}{n} = \frac{1}{n(n-1)\overline{x}} \sum_{i=1}^{n} x_i \left(2i - n - 1\right).$$

Dieser ist auch im Falle maximaler Ungleichheit gleich 1,

$$\text{GUK} = \frac{1}{n(n-1)\overline{x}} x_n \left(2n - n - 1\right) = \frac{1}{n\overline{x}} x_n = 1.$$

Anhang

```
% Programm Gini-Koeffizient

f1 = 0.50; l1 = 0.025;
f2 = 0.40; l2 = 0.475;
f3 = 0.09; l3 = 0.270;
f4 = 0.01; l4 = 0.230;

lf1 = l1/f1;
X = ['lf1 = ',num2str(lf1)];
       disp(X)
lf2 = l2/f2;
X = ['lf2 = ',num2str(lf2)];
       disp(X)
lf3 = l3/f3;
X = ['lf3 = ',num2str(lf3)];
       disp(X)
lf4 = l4/f4;
X = ['lf4 = ',num2str(lf4)];
       disp(X)

disp(' ');

F0 = 0;
X = ['F0 = ',num2str(F0)];
       disp(X)
F1 = f1;
X = ['F1 = ',num2str(F1)];
       disp(X)
F2 = f1 + f2;
X = ['F2 = ',num2str(F2)];
       disp(X)
F3 = f1 + f2 + f3;
X = ['F3 = ',num2str(F3)];
       disp(X)
F4 = f1 + f2 + f3 + f4;
X = ['F4 = ',num2str(F4)];
       disp(X)

disp(' ');

L0 = 0;
X = ['L0 = ',num2str(L0)];
       disp(X)
L1 = l1;
X = ['L1 = ',num2str(L1)];
       disp(X)
L2 = l1 + l2;
X = ['L2 = ',num2str(L2)];
       disp(X)
L3 = l1 + l2 + l3;
X = ['L3 = ',num2str(L3)];
       disp(X)
L4 = l1 + l2 + l3 + l4;
X = ['L4 = ',num2str(L4)];
       disp(X)
```



```
disp(' ');

B1 = (L1 - 0.5*l1)*f1;
X = ['B1 = ',num2str(B1)];
       disp(X)
B2 = (L2 - 0.5*l2)*f2;
X = ['B2 = ',num2str(B2)];
       disp(X)
B3 = (L3 - 0.5*l3)*f3;
X = ['B3 = ',num2str(B3)];
       disp(X)
B4 = (L4 - 0.5*l4)*f4;
X = ['B4 = ',num2str(B4)];
       disp(X)

disp(' ');

B = B1 + B2 + B3 + B4;
X = ['B = ',num2str(B)];
       disp(X)
GUK = 1 - 2*B;
X = ['GUK = ',num2str(GUK)];
       disp(X)
```

```
>> gini
lf1 = 0.05
lf2 = 1.1875
lf3 = 3
lf4 = 23

F0 = 0
F1 = 0.5
F2 = 0.9
F3 = 0.99
F4 = 1

L0 = 0
L1 = 0.025
L2 = 0.5
L3 = 0.77
L4 = 1

B1 = 0.00625
B2 = 0.105
B3 = 0.05715
B4 = 0.00885

B = 0.17725
GUK = 0.6455
```

