

Lösung zur VWA-Klausur im Fach Statistik vom 24. Mai 2011

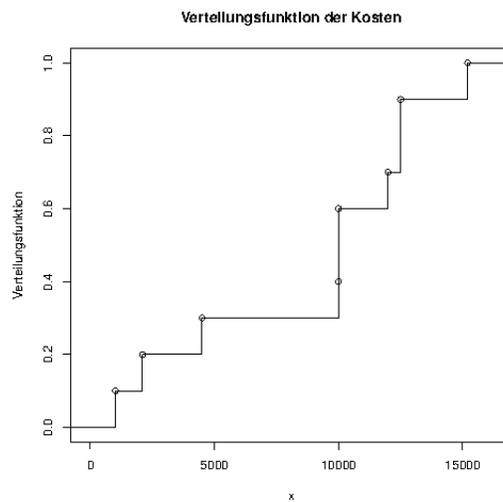
Prof. Dr. Ingo Klein

Aufgabe 1

1. Verteilungsfunktion $F_X(x)$ der Kosten:

$$F_X(x) = \begin{cases} 0 & \text{für } x < 1010 \\ 1/10 & \text{für } 1010 \leq x < 2100 \\ 2/10 & \text{für } 2100 \leq x < 4500 \\ 3/10 & \text{für } 4500 \leq x < 9999 \\ 4/10 & \text{für } 9999 \leq x < 10000 \\ 6/10 & \text{für } 10000 \leq x < 12000 \\ 7/10 & \text{für } 12000 \leq x < 12500 \\ 9/10 & \text{für } 12500 \leq x < 15200 \\ 1 & \text{für } x \geq 15200 \end{cases}$$

Grafik:



2. 0,3% der Unternehmen haben Kosten kleiner oder gleich 9.000 EUR.
(Unternehmen 2 (1.010), 4 (2.100) und 8 (4.500)).

3. Median = 10000
Unteres Quartil = 4500
Oberes Quartil = 12500
Spannweite = 14190
Interquartilsabstand = 8000

4. Varianz = 20.663.445,29
wegen

$$\begin{aligned}\text{Varianz} &= \frac{1}{10} (12500^2 + 1010^2 + 12500^2 + 2100^2 + 10000^2 + 9999^2 + 15200^2 + 4500^2 + 10000^2 + 12000^2) - 8980,9^2 \\ &= \frac{1}{10} (156250000 + 1020100 + 156250000 + 4410000 + 100000000 + 99980001 + 231040000 + 20250000 \\ &\quad + 100000000 + 1440000) - 8980,9^2 \\ &= \frac{1}{10} (1013200.101) - 8980,9^2 = 20663445,29\end{aligned}$$

5. Vor- und Nachteile:

- Spannweite:
Nachteil: Reagiert empfindlich auf Ausreißer. Hängt nur von extremen Beobachtungen ab.
Vorteil: Leicht zu berechnen.
- Interquartilsabstand:
Nachteil: Hängt nicht von allen Daten ab.
Vorteile: Nicht mehr empfindlich gegenüber Ausreißern.
- Varianz:
Nachteil: Empfindlich gegenüber Ausreißern. Verändert sich quadratisch, wenn Dimensionseinheit geändert wird.
Vorteil: Berücksichtigt alle Daten.

Aufgabe 2

1. $-10 + 10 - 50 + 100 + x = 0 \implies x = -50$

2. $1000 \cdot 1,1^4 = 2593,74$

3. $\sqrt[10]{\frac{2.500}{1.000}} = 1,09596$

4. Harmonisches Mittel:

$$\left(\frac{1}{100} + \frac{1}{80} \right)^{-1} = 88,89 \text{ kmh}$$

Aufgabe 3

1. Paasche-Index = $\frac{247}{228} = 1,083$ wegen

- Hypothetische Ausgaben für den Warenkorb der Berichtsperiode zu Preisen der Basisperiode:

				pro kg			
Lebkuchen	100 g	2,10	=	21	· 2 kg	=	42
Schokolade	100 g	0,85	=	8,50	· 10 kg	=	85
Pralinen	100 g	2,70	=	27	· 3 kg	=	81
Trüffel	100 g	4,00	=	40	· 0,5 kg	=	20
Summe							228

- Tatsächliche Ausgaben für den Warenkorb der Berichtsperiode zu Preisen der Berichtsperiode:

				pro kg			
Lebkuchen	100 g	2,30	=	23	· 2 kg	=	46
Schokolade	100 g	1,05	=	10,5	· 10 kg	=	105
Pralinen	100 g	2,50	=	25	· 3 kg	=	75
Trüffel	100 g	4,20	=	42	· 0,5 kg	=	21
Summe							247

D.h. der Warenkorb ist in dem Vergleichszeitraum um 0,83 % teurer geworden.

2. Der Warenkorb von Paasche ist aktuell mit dem Problem, das es aktuelle Produkte in der Basisperiode noch nicht gegeben haben muss. Der Warenkorb des Laspeyres enthält dagegen veraltete Produkte, die es heute nicht mehr gibt. Die Erhebung der jeweils aktuellen Mengen für den Paasche-Index ist aufwändig.

Aufgabe 4

- Arbeitstabelle:

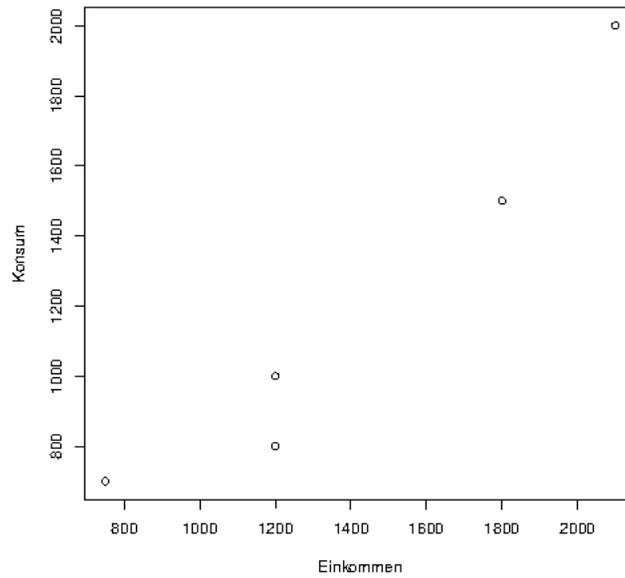
Jahr	Quartal	Kosten	gld. Durchschnitt	Differenz	
1	I	95			103,625
	II	105			99,25
	III	90	100,25	-10,25	100,625
	IV	110	101,75	8,25	97,75
2	I	97	103,25	-6,25	105,625
	II	115	105,25	9,75	109,25
	III	92	107,375	-15,375	102,625
	IV	124	107,375	16,625	111,75
3	I	100	107,875	-7,875	108,625
	II	112	107,625	4,375	106,25
	III	99	105,25	-6,25	109,625
	IV	115	103,125	11,875	102,75
4	I	90	101,75	-11,75	98,625
	II	105	101,875	3,125	99,25
	III	95			105,625
	IV	120			107,75

- Saisonkoeffizienten:

$$\begin{array}{l|l}
 \text{I} & = (-6,25 + (-7,875) + (-11,75))/3 = -8,625 \\
 \text{II} & = (9,75 + 4,375 + 3,125)/3 = 5,75 \\
 \text{III} & = (-10,25 + (-15,375) + (-6,25))/3 = -10,625 \\
 \text{IV} & = (8,25 + 16,625 + 11,875)/3 = 12,25
 \end{array}$$

Aufgabe 5

1. Scatterplot:



2. Mittelwert des Konsums: 1200 EUR.

Steigung:

$$\hat{a}_1 = \frac{\text{Kovarianz}}{\text{Varianz von Einkommen}} = \frac{225000}{230400} = 0,9766.$$

Achsenabschnitt = Mittelwert von Konsum - Steigung · Mittelwert von Einkommen:

$$\hat{a}_0 = 1.200 - 0,9766 \cdot 1.410 = 177,006.$$

Regressionsgerade: Konsum=177,006 + 0,9766 · Einkommen.

3. Regressionsgerade:

