

Musterlösung zur Nachklausur Statistik

TIT13

Oettinger 10.2015

Zeit: 90Min.

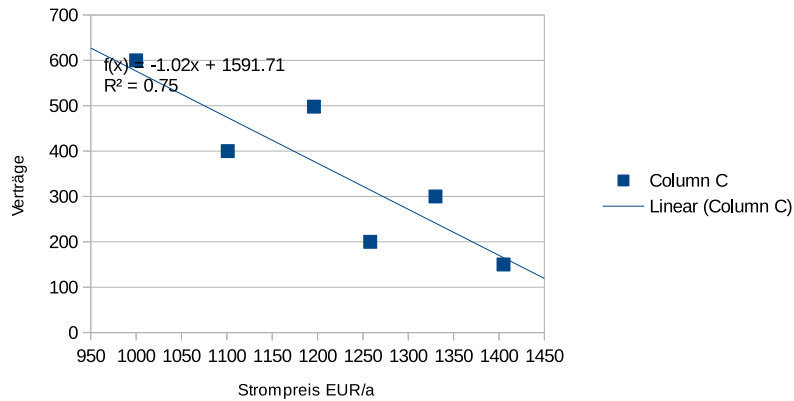
Insgesamt erreichbare Punktzahl: 105, 100%: 100 Punkte.

Aufgabe 1

- (a) Im Allgemeinen sind Modus und Median verschieden - falsch.
- (b) Bei einer symmetrischen Verteilung sind Modus und Median gleich - falsch.
- (c) Die Varianz ist eine Summe quadratischer Größen - richtig.
- (d) Der Gini-Koeffizient bewegt sich als anständiger Koeffizient natürlich zwischen 0 und 1. Richtig.
- (e) Der Quartilsabstand ist ein Maß für die Streuung - richtig.
- (f) Maschinenausfälle besitzen keine Lebensdauer - es handelt sich um eine Bewegungsmasse - falsch.

Aufgabe 2

a) Daten mit Ausgleichsgerade



b) Anpassung einer Geraden $y = a \cdot x + b$ über lineare Regression: Tabelle benötigter Daten

Preis	Verträge	$(x_i - \bar{x})$	$(y_i - \bar{y})$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
1000	600	-215	242	-52030	46225	58564
1101	400	-114	42	-4788	12996	1764
1196	498	-19	140	-2660	361	19600
1258	200	43	-158	-6794	1849	24964
1330	300	115	-58	-6670	13225	3364
1405	150	190	-208	-39520	36100	43264

Die Steigung der Geraden ist

$$a = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = -1,02$$

der Achsenabschnitt ist

$$b = \bar{y} - a\bar{x} = 1591,71$$

c) der Pearson-Koeffizient ist

$$r_{xy} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

$$= \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} = -0,87 \quad (2)$$

Der Wert nahe eins legt einen relativ guten linearen Zusammenhang zwischen den Daten nahe.

Aufgabe 3

Alle Angaben in min:

(a) Geordneter Vektor der Verspätungen: (2, 6, 8, 8, 10, 10, 12, 20). Der Umfang der Stichprobe ist $n = 8$, das untere Quartil (das 0,25-Quantil) ist

$$x_{0,25} = \frac{x_{n/4} + x_{n/4+1}}{2} = \frac{6 + 8}{2} = 7.$$

(b) Mittelwert $\bar{x} = \frac{1}{8}(2 + 6 + 2 \cdot 8 + 2 \cdot 10 + 12 + 20) = 9,5$.

(c) Die Spannweite ist $\max(\{x_i\}) - \min(\{x_i\}) = 20 - 2 = 18$.

(d) Die mittlere absolute Abweichung vom arithmetischen Mittel ist

$$d_{\bar{x}} = \frac{1}{8} (|2 - 9,5| + |6 - 9,5| + 2 \cdot |8 - 9,5| + 2 \cdot |10 - 9,5| + |12 - 9,5| + |20 - 9,5|) = 3,5$$

Die mittlere quadratische Abweichung (die Varianz) hat den Wert

$$s^2 = \frac{1}{8} ((2 - 9,5)^2 + (6 - 9,5)^2 + 2 \cdot (8 - 9,5)^2 + 2 \cdot (10 - 9,5)^2 + (12 - 9,5)^2 + (20 - 9,5)^2) = 23,75$$

Aufgabe 4

Ergänzte Tabelle der absoluten Häufigkeiten:

Y X	1	2	3	4	Summe
1	2	4	1	9	16
2	17	34	10	85	146
3	3	6	2	15	26
Summe	22	44	13	109	188

Die beiden Merkmale sind nicht statistisch unabhängig, da die relativen Häufigkeiten in den Spalten sowie der Randspalte unterschiedlich sind.

Tabelle der relativen Häufigkeiten:

Y X	1	2	3	4	Summe
1	1/11	1/11	1/13	9/109	4/47
2	17/22	17/22	10/13	85/109	73/94
3	3/22	3/22	2/13	15/109	13/94
Summe	1	1	1	1	1

Die bedingte Verteilung ist

$$f(x_i|Y = 2) = 1/11; 17/22; 3/22$$

Für die Varianz wird die zweite Zeile der Tabelle benutzt, das arithmetische Mittel ist

$$\frac{1}{n} \sum_{i=1}^n y_i = \frac{1}{146} (17 \cdot 1 + 34 \cdot 2 + 10 \cdot 3 + 85 \cdot 4) = 3,12$$

Die Varianz

$$s^2(Y|X = 2) = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$$
$$= \frac{1}{146} (17 \cdot (1 - 3,12)^2 + 34 \cdot (2 - 3,12)^2 + 10 \cdot (3 - 3,12)^2 + 85 \cdot (4 - 3,12)^2) = 1,27$$

Aufgabe 5

Geeignete Mittelwerte.

1. Geometrisches Mittel:

$$\bar{x}_G = \sqrt[3]{(1 + 0,1) \cdot (1 + 0,15) \cdot (1 - 0,0005)} - 1 = 8,13\%$$

2. Der Mittelwert berechnet sich nach

$$\bar{x} = \frac{1}{20}(5 \cdot 1 + 11 \cdot 2 + 3 \cdot 3 + 4) = 2,0$$

Für die Kandidaten, die nicht bestanden haben, kann keine Note angegeben werden, sie werden zur Berechnung *nicht* herangezogen.

3. Insgesamt befragte Personen: $100 + 1000 = 1100$. Für die Abschaffung sind $60 + 380 = 440$. Also sind $440/1100 = 40\%$ dafür.