# Musterlösung zur Nachklausur Statistik

## TIT11

## Oettinger 10.2013

**Zeit: 60Min.**

Insgesamt erreichbare Punktzahl: 107, 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) Ordinale Merkmale besitzen eine Rangfolge - Richtig.

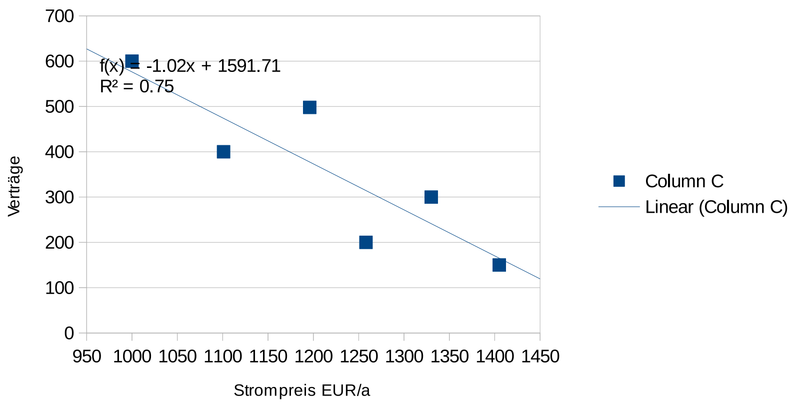
(d) Der Gini-Koeffizzient bewegt sich als anständiger Koeffizient zwischen 0 und 1. Richtig.

(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

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

Die mittlere quadratische Abweichung (die Varianz) hat den Wert

$$s^2 = \frac{1}{8}\left((2 - 9,5)^2 + (6 - 9,5)^2 + 2 \cdot (8 - 9,5)^2 + 2 \cdot (10 - 9,5)^2 + (12 - 9,5)^2 \right.$$
$$\left. + (20 - 9,5)^2\right) - 9,5^2 = 14,25$$

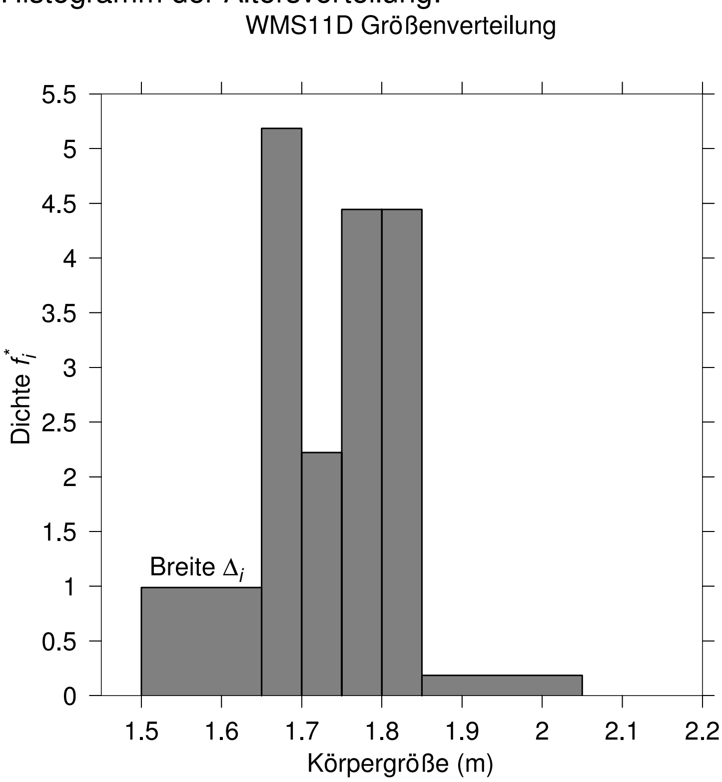
## Aufgabe 4

Daten zur Altersverteilung:

| Größenklasse | absolute Häufigkeit $h_k$ | relative Häufigkeit $f_k$ | Breite $\Delta x_k$ | Dichte $f_k^*$ |
|---|---|---|---|---|
| [1,50 ; 1,65] | 4 | 0.148 | 0.15 | 0.988 |
| ]1,65 ; 1,70] | 7 | 0.259 | 0.05 | 5.185 |
| ]1,70 ; 1,75] | 3 | 0.111 | 0.05 | 2.222 |
| ]1,75 ; 1,80] | 6 | 0.222 | 0.05 | 4.444 |
| ]1,80 ; 1,85] | 6 | 0.222 | 0.05 | 4.444 |
| ]1,85 ; 2,05] | 1 | 0.037 | 0.2 | 0.185 |

**Tabelle 1:** die Größenverteilung im Kurs in klassierter Form. Die Dichte $f_k^*$ ist der Quotient $f_k/\Delta_k$.

(a) Histogramm der Altersverteilung:



(b) Berechnung des Durchschnittsalters (in Jahren):

$$\bar{x} = \frac{1}{27}(4 \cdot 1,575 + 7 \cdot 1,675 + 3 \cdot 1,725 + 6 \cdot 1,775 + 6 \cdot 1,825 + 1,95) = 1,731$$

(c) Berechnung des Medians:
50% der befragten Personen werden in der 3.Klasse erreicht. Der Median lässt sich also wie folgt berechnen:

$$F(\bar{x}_Z) = x_3^u + (x_3^o - x_3^u)\frac{F(\bar{x}_Z) - F(x_3^u)}{F(x_3^o) - F(x_3^u)} = x_3^u + (x_3^o - x_3^u)\frac{F(0,5) - F(x_3^u)}{F(x_3^o) - F(x_3^u)}$$

$$1,675 + (1,725 - 1,675) \cdot \frac{0,5 - 11/27}{(14 - 11)/27} = 1,717$$

## 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.