Musterlösung zur Klausur Statistik

Messe-, Kongress- und Eventmanagement
W MS 09 B

Oettinger 6.9.2010

Zeit: 90Min.

Insgesamt erreichbare Punktzahl: 55, 100%: 50 Punkte.

Aufgabe 1

(a) Der Median entspricht dem 50%-Quantil - richtig.

(b) Für eine unimodale, symmetrische Verteilung gilt stets, dass der Median und der Modus denselben Wert annehmen - richtig.

(c) Die Varianz kann nur positive Werte annehmen - richtig.

(d) Ein Merkmal ist entweder metrisch oder stetig, d.h. es gibt kein Merkmal, das gleichzeitig metrisch und stetig ist - falsch.

(e) Keine der Aussagen (a)-(d) ist richtig - falsch.

Aufgabe 2

An einer neu gegründeten Universität sollen 6-stellige Matrikelnummern vergeben werden.

(a) Anzahl A aller möglichen Kombinationen für 6-stellige Matrikelnummern: Jede Stelle kann mit $0 \ldots 9$ besetzt werden. $A = 10^6$ (unter der Annahme, dass auch 000000 als Matrikelnummer zählt).

(b) Soll keine Nummer mit einer Null beginnen, gibt es für die erste Stelle nur 9 Möglichkeiten $(1 \dots 9)$. $A = 9 \cdot 10^5$.

(c) Keine Ziffer soll zweimal vorkommen: für die erste Stelle gibt es 10 Möglichkeiten,
für die zweite 9,
für die dritte 8 usw.
$A = 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 = 151200$

Aufgabe 3

Redezeiten von Politikern, alle Angaben und Ergebnisse in Minuten:

Politiker	A	B	C	D	E	F
Redezeit	6	8	10	12	20	10

(a) Geordneter Vektor der Merkmalswerte: (6, 8, 10, 10, 12, 20). Median ist

$$\bar{x}_Z = \frac{10+10}{2} = 10.$$

(b) $\bar{x} = \frac{1}{6} \cdot (6+8+10+10+12+20) = 11.$

(c) $s_w =$max$(x_i)-$min$(x_i) = 20 - 6 = 14.$

(d) mittlere absolute Abweichung

$$d_{\bar{x}} = \frac{1}{6} \sum_{i=1}^{n} \mid x_i - \bar{x} \mid$$

$$= \frac{1}{6} (\mid 6-11 \mid + \mid 8-11 \mid +2 \cdot \mid 10-11 \mid + \mid 12-11 \mid + \mid 20-11 \mid) = 10/3$$

Aufgabe 4

Vollständige gemeinsame Wahrscheinlichkeitstabelle der beiden Zufallsvariablen X und Y und relative Werte:

Y / X	1	2	3	$f_X(X)$
1	1/16	1/8	1/16	1/4
2	1/8	1/4	1/8	1/2
3	1/16	1/8	1/16	1/4
$f_Y(Y)$	1/4	1/2	1/4	1

Y / X	1	2	3	$f_X(X)$
1	1/4	1/4	1/4	1/4
2	1/2	1/2	1/2	1/2
3	1/4	1/4	1/4	1/4
$f_Y(Y)$	1	1	1	1

- Die Spalten sowie die Randverteilung der rechten Tabelle sind identisch $\Longrightarrow$ die Zufallsvariablen X und Y sind stochastisch unabhängig.
- $f(x_i|Y=2) = (1/8, 1/4, 1/8)$
- Varianz $s^2(X|Y=2) = \sum_{i=1}^{n}(x_i - \bar{x})^2 \cdot f_i$. Benötigt wird das arithmetische Mittel

$$\bar{x} = \frac{1}{1/2}(1/8 \cdot 1 + 1/4 \cdot 2 + 1/8 \cdot 3) = 2 \cdot (\frac{1+4+3}{8}) = 2$$

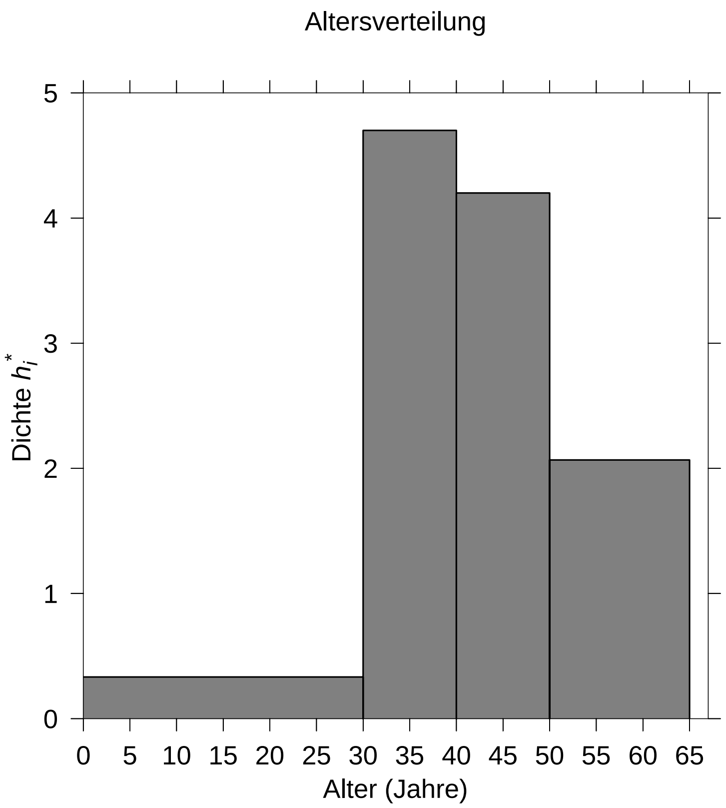
$$s^2(X|Y=2) = \sum_{i=1}^{3}(x_i - \bar{x})^2 \cdot f_i = (1-2)^2 \cdot \frac{1}{4} + (2-2)^2 \cdot \frac{1}{2} + (3-2)^2 \cdot \frac{1}{4}$$
$$= \frac{1}{4} + 0 + \frac{1}{4} = \frac{1}{2}$$

Aufgabe 5

Daten zur Altersverteilung:

Alter in Jahren von ... bis unter ...	Absolute Häufigkeit	relative Häufigkeit f_i	kumulierte rel. Häufigkeit F_i	Klassenbreite Δ_i	Dichte $h_i^* = h_i/\Delta_i$
bis 30	10	0,076	0,076	30	0,333
30 - 40	47	0,362	0,438	10	4,7
40 - 50	42	0,323	0,761	10	4,2
50 -65	31	0,238	1	15	2,067

(a) Histogramm der Altersverteilung:



Berechnung des Durchschnittsalters (in Jahren):

$$\bar{x} = \frac{1}{130}(10 \cdot 15 + 47 \cdot 35 + 42 \cdot 45 + 31 \cdot 57,5) = 42,06$$

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

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

$$40 + 10 \cdot \frac{0,5 - 57/130}{(99 - 57)/130} = 40 + 10 \cdot \frac{8}{42} = 41,9$$

(b) Durchschnittseinkommen $\bar{y}$ in Tausend Euro:

$$\bar{y} = \frac{1}{130} \cdot (10 \cdot 2,5 + 47 \cdot 4,2 + 42 \cdot 5 + 31 \cdot 4,9) = 4,49$$

(c) Gesamtvarianz des Einkommens: die Varianz s^2 ist die mittlere quadratische Abweichung $\Longrightarrow$ die Summe der quadratischen Abweichungen ist gleich $n \cdot s^2$. Die Gesamtvarianz ergibt sich also zu

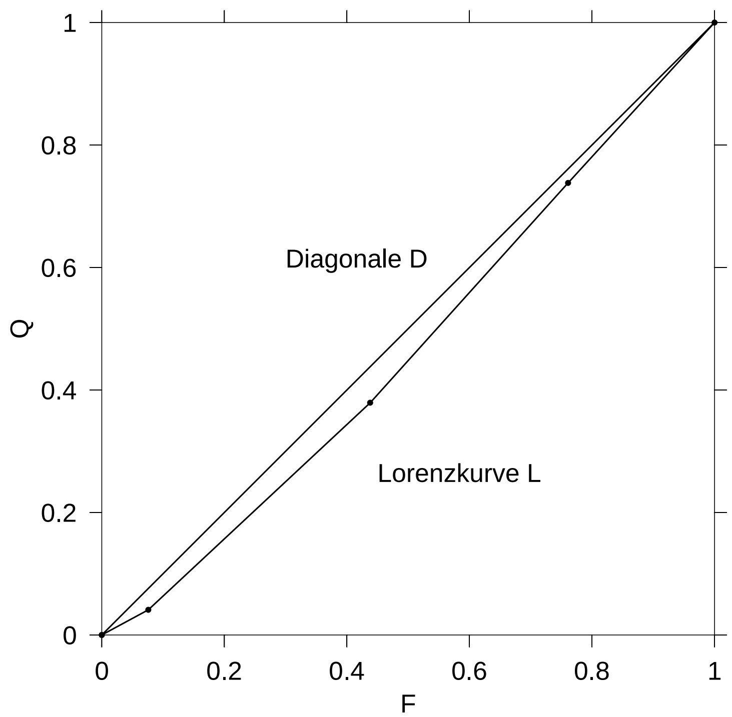
$$s^2 = \frac{1}{n}\sum_{i=1}^{4} h_i \cdot s^2_{Y,i} = \frac{1}{130}(10 \cdot 1,8 + 47 \cdot 2,9 + 42 \cdot 3,4 + 31 \cdot 3,6) = 3,14$$

(d) Für die Darstellung der Lorenzkurve werden die relativen Anteile q_i an der Gesamtsumme der Merkmalswerte bzw deren kumulierte Werte Q_i benötigt:

Alter in Jahren von ... bis unter ...	Absolute Häufigkeit	$\bar{y}$	relativer Anteil q_i	kumulierte rel. Anteile Q_i
bis 30	10	2,5	0,0411	0,0411
30 - 40	47	4,2	0,3479	0,389
40 - 50	42	5,0	0,3662	0,7552
50 -65	31	4,9	0,2448	1

Der Wert des Gini-Koeffizienten entspricht der doppelten Fläche zwischen der Diagonalen und der Lorenzkurve. Der Gini-Koeffizient ist sehr klein, damit ist die Disparität (die relative Konzentration) klein, die Einkommen sind relativ gleichmäßig verteilt.

Lorenzkurve des Einkommens



Aufgabe 6

Geeignete Mittelwerte.

1. Eine Stunde 50 km/h, 1 Stunde und 15 Minuten 40 km/h.
   Die Gesamtzeit sind 2 Stunden und 15 Minuten, die zurückgelegte Strecke $s = 1$h$\cdot 50$km/h $+1,25$h$\cdot 40$km/h = 100km.
   Durchschnittsgeschwindigkeit in km/h:
   $$\bar{v} = \frac{100}{2,25} = 44,\bar{4}.$$
   Das ist das harmonische Mittel der Geschwindigkeiten (in km/h):
   $$\bar{v} = \frac{1}{1/2(\frac{1}{50} + \frac{1}{40})} = 44,\bar{4}$$
2. Arithmetisches Mittel: $\bar{x} = \frac{1}{11}(5 \cdot 3 + 2 \cdot 2 + 3 \cdot 1) = 2$ Sterne
3. Geometrisches Mittel: $\bar{x}_G = \sqrt{(1 + 0,2) \cdot (1 + 0,3875)} - 1 = 29,03\%$
4. Es werden Durchschnittsgeschwindigkeiten gemittelt $\Longrightarrow$ harmonisches Mittel.
5. Insgesamt befragte Personen: 100 + 1000 = 1100. Für die Abschaffung sind 60 + 380 = 440. Also sind 440/1100 = 40% dafür.

Aufgabe 7

Nominale/ordinale/kardinale Merkmale:

(a) Körpergröße: kardinal

(b) Farbe: nominal

(c) Krawattenlänge: kardinal

(d) Qualität von Vorlesungen: ordinal