

# Klausur Statistik

## Lösungshinweise

Prüfungsdatum: 4. Juli 2018 – Prüfer: Etschberger, Henle, Jansen, Wesp  
 Studiengang: BW, IM  
 Punkte: 24, 22, 12, 12, 10, 10 ; Summe der Punkte: 90

### Aufgabe 1

24 Punkte

#### Aufgabenteil I

In der Augsburger Innenstadt werden 25 Personen nach der Anzahl ihrer Haustiere (Merkmal  $X$ ) befragt. Daraus ergibt sich die folgende Urliste:

$$x = (2, 0, 4, 1, 1, 1, 0, 0, 4, 2, 1, 0, 1, 1, 2, 1, 0, 1, 0, 0, 2, 2, 1, 1, 1)$$

- Erstellen Sie eine Tabelle mit den absoluten sowie den kumulierten relativen Häufigkeiten.
- Mit welchen R-Befehlen kann die Teilaufgabe a) gelöst werden?
- Geben Sie zu  $X$  das empirische Quantil  $\tilde{x}_{0,92}$  an.
- Geben Sie den R-Befehl zur Berechnung von  $\tilde{x}_{0,92}$  an.
- Bestimmen Sie den Modus, den Median, das arithmetische Mittel sowie die mittlere quadratische Abweichung von  $X$ .

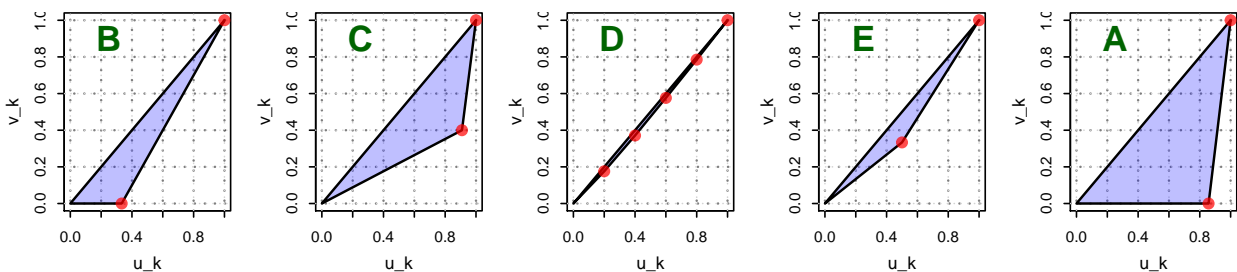
#### Aufgabenteil II

Toni hat zu den metrischen Merkmalen  $A, B, C, D, E$  jeweils eine Urliste ermittelt (siehe Tabelle rechts). Für jedes Merkmal hat er auch den zugehörigen (nicht normierten) Ginikoeffizienten berechnet und eine Lorenzkurve gezeichnet.

Leider sind seine Unterlagen völlig durcheinandergekommen und er weiß nicht mehr, welche Urliste zu welchem Koeffizienten und zu welcher Kurve gehört.

Merkmal	Urliste
$A$	0, 0, 0, 0, 5, 0, 0
$B$	1, 1, 0
$C$	1, 1, 1, 1, 1, 1, 1, 15, 1, 1, 1
$D$	100, 107, 110, 90, 105
$E$	5000, 2500

- Schreiben Sie in die Grafik jeder Lorenzkurve den Bezeichner ( $A - E$ ) des jeweiligen zugehörigen Merkmals.



- Schreiben Sie unter jeden der Ginikoeffizienten  $G$  den Bezeichner ( $A - E$ ) des jeweiligen zugehörigen Merkmals.

Wert des Ginikoeffizienten	0.509	0.857	0.037	0.167	0.333
Buchstabe des Merkmals	<b>C</b>	<b>A</b>	<b>D</b>	<b>E</b>	<b>B</b>

## Lösungshinweis:

	Anzahl Haustiere $a_j$	0	1	2	4
a)	absolute Häufigkeit $h_j$	7	11	5	2
	kumulierte relative Häufigkeit $F_j$	0.28	0.72	0.92	1.00

R

b)

```
xI = c(2, 0, 4, 1, 1, 1, 0, 0, 4, 2, 1, 0, 1, 1, 2, 1, 0, 1, 0, 0, 2, 2, 1, 1, 1)
table(xI)
cumsum(table(xI)) / length(xI)
```

c) `quantile(xI, probs = alpha, type=2)`

```
## 92%
## 3
```

R

d) s.o.

$$e) x_{\text{Mod}} = 1 \quad x_{\text{Med}} = x_{\frac{n+1}{2}} = x_{13} = 1$$

$$\bar{x} = 1.16 \quad s^2 = 1.1744$$

f) siehe oben

g) siehe oben

Bäckermeister Boris hat in seinem Sortiment auch Torten. Da er keinen Konditor angestellt hat, lässt er die Torten von einer Großkonditorei liefern. Boris hat sich in den letzten 4 Jahren zu den Einkaufspreisen ( $p$ ) in Euro sowie zu den bestellten Mengen ( $q$ ) der beiden Güter Sachertorte und Schwarzwälder Kirschtorte folgende Notizen gemacht:

Jahr	1		2		3		4	
	$p_1$	$q_1$	$p_2$	$q_2$	$p_3$	$q_3$	$p_4$	$q_4$
Sacher	10	150	11	120	12	130	15	100
Schwarzwälder Kirsch	12	170	11	150	13	100	16	120

- Berechnen Sie den Preisindex von Fisher und Lowe zur Berichtszeit 4 und Basiszeit 1.
  - Der Inhaber will nun wissen, ob zwischen den Preisen der beiden Torten ein Zusammenhang besteht. Berechnen Sie ein geeignetes Zusammenhangsmaß und interpretieren Sie das Ergebnis.
  - Bestimmen Sie den Funktionsterm eines linearen Regressionsmodells, indem der Preis der Sachertorte in Abhängigkeit vom Preis der Schwarzwälder Kirschtorte beschrieben wird.
- R**
- Geben Sie R-Befehle an, mit dem man dieses Regressionsmodell erzeugen kann.
  - Wie hoch wäre nach dem Modell aus Teilaufgabe c) der Preis einer Sachertorte, wenn eine Schwarzwälder Kirschtorte 15 Euro kostet?

Lösungshinweis:

a)  $P_{1,4}^L = \frac{15 \cdot 150 + 16 \cdot 170}{10 \cdot 150 + 12 \cdot 170} = \frac{497}{354} \approx 1,404$

$P_{1,4}^P = \frac{15 \cdot 100 + 16 \cdot 120}{10 \cdot 100 + 12 \cdot 120} = \frac{171}{122} \approx 1,402$

$P_{1,4}^F = \sqrt{\frac{497}{354} \cdot \frac{171}{122}} \approx 1,403$

$\bar{q}^{\text{Sacher}} = 125, \bar{q}^{\text{Schwarzwälder}} = 135 \Rightarrow P_{1,4}^{LO} = \frac{15 \cdot 125 + 16 \cdot 135}{10 \cdot 125 + 12 \cdot 135} = \frac{807}{574} \approx 1,406$

b)  $r = \frac{637 - 4 \cdot 12 \cdot 13}{\sqrt{(590 - 4 \cdot 12^2)(690 - 4 \cdot 13^2)}} = \frac{13}{14} \approx 0,92857$

→ starker positiver Zusammenhang (je größer der Preis der Sachertorte desto höher der Preis der Schwarzwälder Kirschtorte)

c)  $\hat{b} = \frac{637 - 4 \cdot 12 \cdot 13}{690 - 4 \cdot 13^2} = \frac{13}{14} \approx 0,92857$

$\hat{a} = 12 - \frac{13}{14} \cdot 13 \approx -0,0714$

$\Rightarrow \hat{y}(x) = -0,0714 + 0,92857 \cdot x$

**R** d)

```
Sacher.p = c(10, 11, 12, 15)
Schwarz.p = c(12, 11, 13, 16)
lm(Sacher.p ~ Schwarz.p)
```

```
##  
## Call:  
## lm(formula = Sacher.p ~ Schwarz.p)  
##  
## Coefficients:  
## (Intercept)    Schwarz.p  
##   -0.07143      0.92857  
  
cor(Sacher.p, Schwarz.p)  
## [1] 0.9285714
```

e)  $\hat{y}(15) = -0,0714 + 0,92857 \cdot 15 \approx 13,86$

### Aufgabe 3

12 Punkte

Schneewittchen und ihre Schwester Rosenrot bekommen je eine Schale mit je 5 Äpfeln geschenkt. In Schneewittchens Schale sind 5 grüne Äpfel von denen 3 giftig und 2 ungiftig sind. In Rosenrots Schale sind 5 rote Äpfel von denen 2 giftig und 3 ungiftig sind.

Beim gemeinsamen Picknick isst Schneewittchen 2 grüne Äpfel und Rosenrot 3 rote Äpfel (das Gift wirkt sehr langsam, so dass beide genug Zeit zum Essen haben). Welche der beiden Schwestern hat eine besser Überlebenschance, wenn

- ein vergifteter Apfel für eine tödliche Vergiftung ausreicht?
- zwei vergiftete Äpfel gegessen werden müssen um eine tödliche Wirkung zu erzielen?

Geben Sie jeweils die Wahrscheinlichkeiten an, mit denen die beiden Schwestern in den jeweiligen Szenarien sterben.

#### Lösungshinweis:

- Man rechnet wie wahrscheinlich es ist, dass Sie nicht vergiftet werden und nimmt dann die Gegenwahrscheinlichkeit:

$$P_{\text{Schneewittchen}}(\text{tot}) = 1 - \frac{2}{5} \cdot \frac{1}{4} = \frac{9}{10} = 0,9$$
$$P_{\text{Rosenrot}}(\text{tot}) = 1 - \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{1}{3} = \frac{9}{10} = 0,9$$

In diesem Fall sterben beide mit einer Wahrscheinlichkeit von 90 %.

- Dieser Fall ist ein wenig schwieriger, denn man muss sich jeweils mehrere Szenarien ansehen. Das einfachste ist es, jeweils einen Baum zu zeichnen ( $G$  steht für einen giftigen Apfel). Schneewittchen stirbt nur im Fall  $GG$ , wohingegen Rosenrot in den Fällen  $GG\bar{G}$ ,  $G\bar{G}G$  und  $\bar{G}GG$  stirbt, wobei  $G$  für giftigen Apfel steht und  $\bar{G}$  für einen ungiftigen Apfel ( $GGG$  kann es nicht geben, da Rosenrot nur zwei vergiftete Äpfel hat). Damit ergibt sich:

$$P_{\text{Schneewittchen}}(\text{tot}) = \frac{3}{5} \cdot \frac{2}{4} = \frac{6}{20} = 0,3$$
$$P_{\text{Rosenrot}}(\text{tot}) = \frac{2}{5} \cdot \frac{1}{4} \cdot 1 + \frac{2}{5} \cdot \frac{3}{4} \cdot \frac{1}{3} + \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{1}{3} = 0,3.$$

Auch in diesem Fall sind die beiden Wahrscheinlichkeiten gleich groß.

## Aufgabe 4

12 Punkte

Gegeben sind Funktion  $f, g: \mathbb{R} \rightarrow \mathbb{R}$  sowie ihre Graphen in den folgenden Abbildungen:

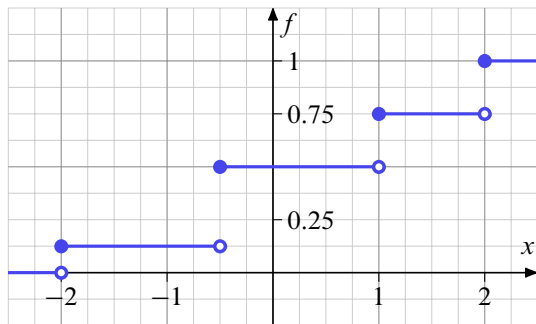


Abbildung 1: Graph von  $f$

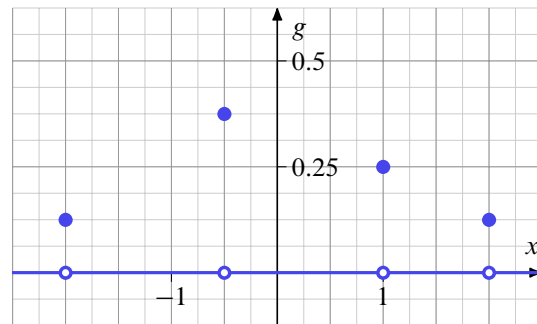


Abbildung 2: Graph von  $g$

- a) Entscheiden Sie jeweils, ob es sich bei  $f$  bzw.  $g$  um eine Wahrscheinlichkeitsfunktion oder um eine Verteilungsfunktion handelt (Begründung ist nötig).

Die Verteilung der Zufallsvariable  $X$  sei durch den Graphen von  $f$  beschrieben. Ermitteln Sie, falls möglich, die nachfolgenden Wahrscheinlichkeiten.

- |                            |                                |
|----------------------------|--------------------------------|
| b) $P(X \leq 0)$           | f) $P(X = 1)$                  |
| c) $P(X \geq 2)$           | g) $P(X \leq 0 \cap X \geq 1)$ |
| d) $P(-0.5 \leq X \leq 1)$ | h) $P(X \leq 0   X \geq 1)$    |
| e) $P(X = 0)$              | i) $P(X \geq -1   X \geq 2)$   |

### Lösungshinweis:

- a)  $f$  ist eine Verteilungsfunktion (mon. wachsend,  $\lim_{x \rightarrow -\infty} f(x) = 0$ ,  $\lim_{x \rightarrow \infty} f(x) = 1$ ),  $g$  wäre eine Wahrscheinlichkeitsfunktion, wenn die Summe der Wahrscheinlichkeiten 1 ergäbe.
- b)  $P(X \leq 0) = P(X = -2) + P(X = -0.5) = 0.125 + 0.375 = 0.5$
- c)  $P(X \geq 2) = P(X = 2) = 0.25$
- d)  $P(-0.5 \leq X \leq 1) = P(X = -0.5) + P(X = 1) = 3/8 + 1/4 = 5/8$
- e)  $P(X = 0) = 0$
- f)  $P(X = 1) = 1/4$
- g)  $P(X \leq 0 \cap X \geq 1) = 0$
- h)  $P(X \leq 0 | X \geq 1) = \frac{P(X \leq 0 \cap X \geq 1)}{P(X \geq 1)} = 0$
- i)  $P(X \geq -1 | X \geq 2) = \frac{P(X \geq 1 \cap X \geq 2)}{P(X \geq 2)} = \frac{P(X \geq 2)}{P(X \geq 2)} = 1$

## Aufgabe 5

10 Punkte

Gegeben ist eine einfache Stichprobe  $X_1, X_2, X_3$ . Betrachtet werden im Folgenden die Stichprobenfunktionen

$$A = 0.5 \cdot X_1 + 0.9 \cdot X_2 - 0.4 \cdot X_3, \quad B = X_1$$

- Sind  $A, B$  erwartungstreue Schätzfunktionen für das arithmetische Mittel  $\mu$  der Grundgesamtheit?
- Zeigen Sie, dass  $A$  als Schätzer nicht so wirksam ist wie  $B$ .

### Lösungshinweis:

- $E[A] = (0.5 + 0.9 - 0.4) \cdot \mu = \mu$   
 $E[B] = \mu$   
beide erwartungstreu.
- $\text{Var}[A] = (0.5^2 + 0.9^2 + 0.4^2) \cdot \sigma = 1.22 > \text{Var}[B] = \sigma$ .

## Aufgabe 6

10 Punkte

Gegeben ist eine Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  mit

$$f(x, y) = x^2 - x - 2 + g(y).$$

Dabei ist  $g(y)$  für die Teilaufgaben a), b) gegeben mit

$$g(y) = \frac{1}{3}y^3 - \frac{5}{2}y^2 + 6y + 3.$$

- Bestimmen Sie den Gradienten  $\nabla f(x, y)$  sowie die Hessematrix  $H_f(x, y)$ .
- Ermitteln Sie alle lokalen Extrema von  $f$ .

Für die Teilaufgaben c), d), e) sei  $g$  ein reelles Polynom 3. Grades in  $y$ , also von der Form

$$g(y) = ay^3 + by^2 + cy + d$$

mit  $a, b, c, d \in \mathbb{R}, a \neq 0$ .

- Begründen Sie, wie viele kritische Punkte  $f$  maximal besitzen kann.
- Geben Sie eine Funktion  $g(y)$  an, so dass  $f$  keinen kritischen Punkt hat.
- Angenommen in Abbildung 3 ist der Graph der Ableitung von  $g(y)$  nach  $y$  dargestellt. Bestimmen Sie anhand dieses Graphen alle kritischen Punkte der Funktion  $f$ .

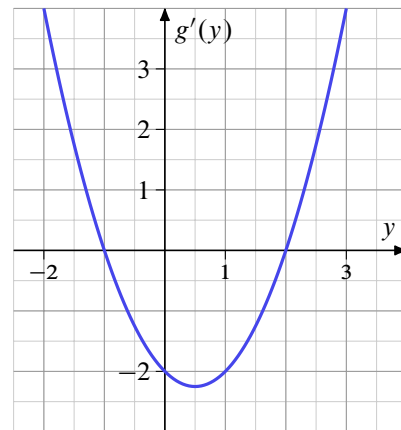


Abbildung 3: Graph von  $g'(y)$

### Lösungshinweis:

- $\nabla f(x, y) = \begin{pmatrix} 2x - 1 \\ y^2 - 5y + 6 \end{pmatrix}$ ,  $H_f(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 2y - 5 \end{pmatrix}$
- $\nabla f(x, y) = (0, 0)^T \Leftrightarrow (x, y) = (0.5, 2), (0.5, 3)$ .  
 $H_f(0.5, 2) = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \Leftrightarrow (0.5, 2)$  ist Sattelpunkt.  
 $H_f(0.5, 3) = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \Leftrightarrow (0.5, 3)$  ist lokales Minimum.
- $f_y$  hat entweder 2, eine doppelte oder keine reelle Nullstelle. Da  $f_x = 0 \Leftrightarrow x = 0.5$  eindeutig, gibt es 2, 1 oder 0 kritische Punkte.
- Es sei  $f_y = y^2 + 1$ , dann hat  $f_y$  keine Nullstelle und es gibt damit keinen kritischen Punkt. Dazu ist zum Beispiel  $g(y) = \frac{1}{3}y^3 + y$ .
- $f_x = 0 \Leftrightarrow x = 1/2$ . Nach der Zeichnung  $f_y = 0 \Leftrightarrow y = -1$  oder  $y = 2$ . Kritische Punkte sind  $(x, y) = (0.5, -1), (0.5, 2)$ .