

# Aufgabensammlung Statistik (Anders)

## VL\_0

## VL\_1&2

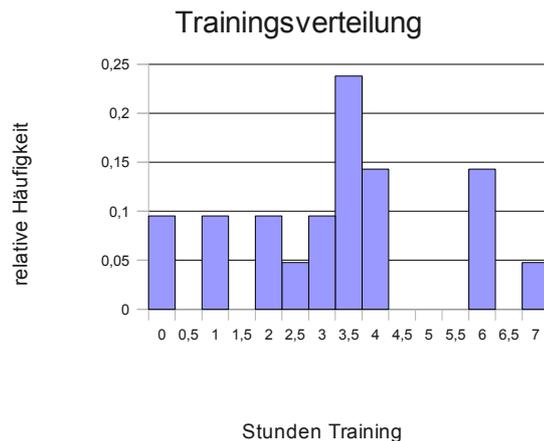
### Aufgabe 1.1:

Die unten angegebene Tabelle zeigt, wie lange Ihr Kommilitone „Hercules“ in den letzten drei Januarwochen dieses Jahres jeweils im Fitnesscenter war:

Wie viele Stunden hat Hercules in den letzten drei Wochen im „Joey´s Bench Press“ trainiert ?	Mo	Di	Mi	Do	Fr	Sa	So
	2	3,5	0	3	2,5	6	6
	1	3,5	1	4	6	3,5	4
	3	3,5	0	2	4	3,5	7

- Fertigen Sie ein Histogramm und eine Tabelle der Häufigkeit, der relativen Häufigkeit und der kumulierten relativen Häufigkeit der Verteilung an. **(4P)**
- Hercules, Arnold, Popeye und Sylvester treffen sich zu einem Vitamindrink an der Bar. Sie wollen ihr Trainingspensum für diese drei Wochen vergleichen. Welcher Lageparameter (arithmetisches Mittel, geometrisches Mittel, Median oder Modalwert) ist Ihrer Meinung nach am geeignetsten und warum? **(2P)**
- Geben Sie Hercules´s arithmetisches Mittel, geometrisches Mittel, Median und Modalwert an. **(2P)**
- Fertigen Sie eine „Five Point Summary“ von Hercules´s Trainingspensum an und zeichnen Sie sie als Box-Whisker Plot. **(3P)**
- Nennen Sie zwei Ihnen bekannte Streuparameter von Hercules´s Trainingspensum und ermitteln Sie die Zahlenwerte für die angegebene Tabelle. **(3P)**

Stunden	H(x)	h(x)	kum h(x)	x*h(x)	quad_abw
0	2	0,10	0,0952	0,00	21,59
0,5	0	0,00	0,0952	0,00	0
1	2	0,10	0,1905	0,10	10,45
1,5	0	0,00	0,1905	0,00	0
2	2	0,10	0,2857	0,19	3,31
2,5	1	0,05	0,3333	0,12	0,62
3	2	0,10	0,4286	0,29	0,16
3,5	5	0,24	0,6667	0,83	0,23
4	3	0,14	0,8095	0,57	1,53
4,5	0	0,00	0,8095	0,00	0
5	0	0,00	0,8095	0,00	0
5,5	0	0,00	0,8095	0,00	0
6	3	0,14	0,9524	0,86	22,1
6,5	0	0,00	0,9524	0,00	0
7	1	0,05	1,0000	0,33	13,8
	<b>21</b>	<b>1</b>	<b>3,29</b>	<b>73,79</b>	



Min:	0
Q25:	2
Median:	3,5
Q75:	4
Max:	7

geom_mittel	2,88
Modalwert:	3,500
arith. Mittelwert	3,286
s:	1,921
emp. Varianz:	3,689

**Aufgabe 1.2: (11 P)**

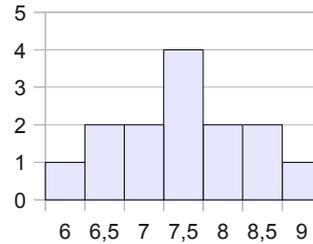
In der ersten Vorlesungseinheit wurde das folgende Beispiel eines Schlafstagebuches der Studentin "Susanne" als Beispiel einer Datenerhebung vorgestellt.

Mo	Di	Mi	Do	Fr	Sa	So
8	6,5	7	7,5	7,5	8,5	8,5
7,5	7	6,5	6	8	9	7,5

Wie viele Stunden hat Susanne in den letzten zwei Wochen geschlafen ?

a) Erstellen Sie eine Häufigkeitstabelle (3-spaltig: absolute, relative, und kumulierte relative Häufigkeiten) und zeichnen Sie ein Histogramm der Verteilung.

Merkmal[Std]	Häufigk.	rel. Hfgk.	kum.rel.Hfgk.
6	1	0,071	0,071
6,5	2	0,143	0,214
7	2	0,143	0,357
7,5	4	0,286	0,643
8	2	0,143	0,786
8,5	2	0,143	0,929
9	1	0,071	1,000
Summen:	14	1	



b) Erstellen Sie die "five point summary" (Fünf-Zahlen-Zusammenfassung) der Daten und zeichnen Sie einen "Box-Whisker-Plot".



Min: 6, 25erPerzentil: 7, Median: 7,5, 75er Perzentil: 8, Max: 9 -> (6,7,7,5,8 9).

c) Berechnen Sie den Modalwert, das arithmetische Mittel, empirische Standardabweichung und Varianz.

Modalwert: 7,5 arithm. Mittel: 7,5 emp. Stddev: 0,855 emp.Var.: 0,731

**Aufgabe 1.3: (8 P)**

Erläutern Sie die Begriffe evtl. mit einem Beispiel:

- a) opportune Stichprobe,
- b) systematische Stichprobe,
- c) geschichtete Stichprobe,
- e) Zufallsstichprobe.

Mo	Di	Mi	Do	Fr	Sa	So
8	6,5	7	7,5	7,5	8,5	8,5
7,5	7	6,5	6	8	9	7,5

**Aufgabe 1.4:**

Die unten angegebene Tabelle zeigt, wie lange Ihr Kommilitone „JunkieSurf“ in den letzten drei Juliwochen dieses Jahres jeweils im Internet "Blutige Orgie" gespielt hat:

Wie viele Stunden hat JunkieSurf in den letzten drei Juliwochen „Blutige Orgie“ gespielt ?	Mo	Di	Mi	Do	Fr	Sa	So
	24	22	1	3	5	24	0
	0	3.5	7	2	2	3.5	4
	3	3.5	0	0	24	3.5	7

a) Fertigen Sie einen Dot Plot und eine Tabelle der Häufigkeit, der relativen Häufigkeit und der kumulierten relativen Häufigkeit der Verteilung an. (4P)

- b) JunkieSurf, SquareEyes, DigiBrute und BodySlasher treffen sich zu einem Chat in einem einschlägigen Chatroom. Sie wollen ihr Spielpensum für diese drei Wochen vergleichen. Welcher Lageparameter (arithmetisches Mittel, Median oder Modalwert) ist Ihrer Meinung nach am geeignetsten und warum?(2P)
- c) Geben Sie JunkieSurf's arithmetisches Mittel, Median und Modalwert an.(2P)
- d) Fertigen Sie eine „Five Point Summary“ von JunkieSurf's Trainingspensum an und zeichnen Sie sie als Box-Whisker Plot.(3P)
- e) Nennen Sie zwei Ihnen bekannte Streuparameter von JunkieSurf's Trainingspensum und ermitteln Sie die Zahlenwerte für die angegebene Tabelle.(3P)

### Aufgabe 1.5:

Ihr Kommilitone Peter "The Chimney" Stuyvesant möchte aus finanziellen Gründen seinen Zigarettenverbrauch etwas einschränken und beginnt zunächst mit einem Protokoll seines gegenwärtigen Verbrauchs.

Die unten angegebene Tabelle zeigt, wie viele Zigaretten Ihr Kommilitone Peter "The Chimney" Stuyvesant in den letzten zwei Wochen dieses Jahres jeweils täglich geraucht hat:

Wie viele Zigaretten hat Peter Stuyvesant in den letzten zwei Wochen täglich jeweils geraucht ?	Mo	Di	Mi	Do	Fr	Sa	So
	23	17	15	17	22	31	15
	22	17	16	14	19	27	19

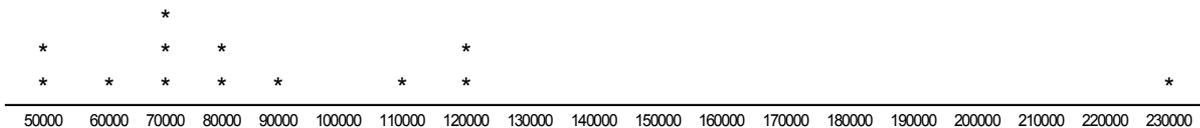
- a) Fertigen Sie einen Dot Plot und eine Tabelle der Häufigkeit, der relativen Häufigkeit und der kumulierten relativen Häufigkeit der Verteilung an.(4P)
- b) Peter "The Chimney" Stuyvesant, John "Clouds" Player, Nicole "la brume" Gauloise, Lisa "The Fog" Dunhill und Smokey Sven treffen sich in einer Vorlesungspause vor dem Fh-Eingang am Aschenbecher zu einer Zigarette. Sie wollen ihre Nikotindosis für diese zwei Wochen vergleichen. Welcher Lageparameter (arithmetisches Mittel, Median oder Modalwert) des Zigarettenkonsums ist Ihrer Meinung nach am geeignetsten und warum?(1P)
- c) Geben Sie Peter Stuyvesant's arithmetisches Mittel, Median und Modalwert an.(2P)
- d) Fertigen Sie eine „Five Point Summary“ von Peter Stuyvesant's Zigarettenkonsum an und zeichnen Sie sie als Box-Whisker Plot.(3P)
- e) Nennen Sie zwei Ihnen bekannte Streuparameter von Peter Stuyvesant's Konsum und ermitteln Sie die Zahlenwerte für die angegebene Tabelle.(2P)

**Aufgabe 1.6:**

Sie fragen auf der Mönckebergstraße in Hamburg 13 Männer, die Krawatte, Anzug und geputzte Schuhe tragen (MKAgs) nach dem aktuellen Jahreseinkommen. Weil das auf den Euro genau niemand weiß, lassen Sie sich das Resultat als den nächstliegenden glatten 10 000er-Euro-Betrag angeben. Sie erhalten folgende Urliste:

MKAgs1: 90000, MKAgs2: 50000, MKAgs3: 110000, MKAgs4: 120000, MKAgs5: 80000, MKAgs6: 70000, MKAgs7: 80000, MKAgs8: 50000, MKAgs9: 60000, MKAgs10: 70000, MKAgs11: 230000, MKAgs12:70000, MKAgs13: 120000

a) Fertigen Sie einen Dot Plot und eine Tabelle der Häufigkeit, der relativen Häufigkeit und der kumulierten relativen Häufigkeit der Verteilung an. **(4P)**



Einkommen in 10 TSD	Häufigk	rel. Häuf.	kum. rel. Häuf.
5	2	0,15	0,15
6	1	0,08	0,23
7	3	0,23	0,46
8	2	0,15	0,62
9	1	0,08	0,69
10	0	0,00	0,69
11	1	0,08	0,77
12	2	0,15	0,92
13	0	0,00	0,92
14	0	0,00	0,92
15	0	0,00	0,92
16	0	0,00	0,92
17	0	0,00	0,92
18	0	0,00	0,92
19	0	0,00	0,92
20	0	0,00	0,92
21	0	0,00	0,92
22	0	0,00	0,92
23	1	0,08	1,00
	<b>13</b>	<b>1</b>	

b) Ein Herrenausstatter hat Interesse an Ihren Daten und möchte sein Sortiment auf MKAgs-Kunden maßschneidern. Welcher Lageparameter (arithmetisches Mittel, Median oder Modalwert) des Jahreseinkommens ist Ihrer Meinung nach am geeignetsten und warum? **(1P)**

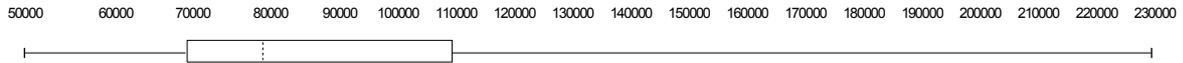
Median, weil der 230000 Ausreißer im arithmetischen Mittel eine falsche Lage vortäuschen würde. (Der Modalwert wäre auch geeignet.)

c) Geben Sie arithmetisches Mittel, Median und Modalwert des Einkommens Ihrer MKAgs-Stichprobe an. **(2P)**

*arithm. Mittel: 92 300, Median: 80 000, Modalwert: 70 000*

d) Fertigen Sie eine „Five Point Summary“ der MKAgS-Einkommen an und zeichnen Sie sie als Box-Whisker Plot. **(3P)**

Min: 50 000, 25er Perzentil: 70 000, Median: 80 000,  
75er Perzentil: 110 000, Max: 230 000



e) Nennen Sie zwei Ihnen bekannte Streuparameter und ermitteln Sie die Zahlenwerte für die angegebenen Einkommensdaten. **(2P)**

Varianz : 2 269 000 000 €<sup>2</sup>,  
Standardabweichung : 47 600 €

Einkommen in 10 TSD	Häufigk	# * Ek	#*(x- mitt)^2
5	2	10	35,8
6	1	6	10,44
7	3	21	14,93
8	2	16	3,03
9	1	9	0,05
10	0	0	0
11	1	11	3,13
12	2	24	15,34
13	0	0	0
14	0	0	0
15	0	0	0
16	0	0	0
17	0	0	0
18	0	0	0
19	0	0	0
20	0	0	0
21	0	0	0
22	0	0	0
23	1	23	189,59
	13	120	272,31

Mittel            9,23  
 Varianz:        22,69 in 10<sup>8</sup> Euro<sup>2</sup>  
 Std.abw.        4,76 in 10<sup>4</sup> Euro

### Aufgabe 1.7

Wegen der im Zuge der Wirtschaftskrise klamm gewordenen Kassen entwickelt die Führungsspitze des Bankhauses "Reibach, Saus & Braus" ein neues Geschäftsfeld, um weiterhin hohe Boni ausschütten zu können.

Die Anlageberater werden in Altersheime geschickt, um die Bewohner der Heime zu Warentermin-geschäften mit Schweinehälften und Zuckerrüben zu animieren, die dann über "Reibach, Saus & Braus" zu satten Provisionen abgewickelt werden sollen.

Der Berater Friedbert Lästig kann im Verlauf von zwei Arbeitswochen folgende Vertragsabschluss-zahlen vorweisen:

	Mo	Di	Mi	Do	Fr
1.Woche	3	2	3	4	5
2.Woche	2	1	7	2	6
3.Woche	0	2	3	6	2

- Erstellen Sie eine Häufigkeitstabelle (3-spaltig: absolute, relative, und kumulierte relative Häufigkeiten) und zeichnen Sie ein Histogramm der Verteilung.
- Erstellen Sie die "five point summary" (Fünf-Zahlen-Zusammenfassung) der Daten und zeichnen Sie einen "Box-Whisker-Plot".
- Berechnen Sie den Modalwert, das arithmetische Mittel, empirische Standardabweichung und empirische Varianz.

## VL\_3

### Aufgabe 3.1:

- a) Wann sind zwei Ereignisse A und B unverträglich, wann sind sie unabhängig ? **(4P)**

(Geben Sie Ihre Antwort auch in mathematischer Form mit den in der Vorlesung benutzten Symbolen an.)

Sind die folgenden Ereignispaare A und B jeweils unverträglich, sind sie unabhängig ?(nicht zutreffendes bitte durchstreichen)

Sie ziehen eine Karte aus einem Skatblatt (7,8,9,10,As,Bube,Dame,König, jeweils in Pique, Kreuz, Herz und Karo)

b) **(1P)**

A: Die Karte ist schwarz (Pique oder Kreuz)

B: Die Karte ist eine Dame

unverträglich

unabhängig

c) **(1P)**

A: Die Karte ist eine 7 oder das Karo-As

B: Die Karte ist die Herz-10

unverträglich

unabhängig

d) **(1P)**

A: Die Karte ist eine Ziffernkarte(7,8,9,10), der Kreuz-Bube oder das Herz-As

B: Die Karte ist rot

unverträglich

unabhängig

e) **(1P)**

A: Die Karte ist eine Ziffernkarte(7,8,9,10), der Kreuz-Bube oder die Kreuz-Dame

B: Die Karte trägt das Symbol Karo

unverträglich

unabhängig

f) (1P)

A: Die Karte ist eine 7 oder das Karo-As

B: Die Karte ist rot

unverträglich

unabhängig

### Aufgabe 3.2:

Im 17. Jahrhundert lebt der junge Adelige Siegfried Freiherr von Süßholz. Weil er schöne Uniformen liebt, wird er Leutnant bei den kaiserlichen Kürassieren. Er ist gutaussehend, liebt die Literatur, die Kunst, edle Pferde und vor allem schöne Frauen. Militärische Übungen, Fechten und Schießen verabscheut er.

Aus dieser Kombination ergibt sich für ihn ein Problem. Wegen seiner notorischen Schürzenjägerei besteht ständig das Risiko, von gehörnten Ehemännern und gekränkten Verlobten zum Duell gefordert zu werden. Leider ist für einen solchen Fall wegen mangelnder Fähigkeiten seinerseits die Wahrscheinlichkeit, ein Duell siegreich zu überstehen nur etwa 0,2. Neben der Schande zieht eine Niederlage (es gibt kein Unentschieden beim Duell) immer auch den Tod oder (selten) nur eine so schwere Verletzung nach sich, dass damit auch die Schürzenjägerei endgültig beendet ist.

Durchschnittlich alle vier Monate macht er eine neue Eroberung in der Damenwelt. Er ist nicht dumm und meidet deshalb Damen, deren offizielle Verehrer besonders kampfeswillig und misstrauisch sind. Dadurch erreicht er, dass bei einer neuen Eroberung die Wahrscheinlichkeit, entdeckt und zum Duell gefordert zu werden nur 0,47 beträgt.

**a) (2P)** Wie groß ist die Wahrscheinlichkeit, dass Freiherr von Süßholz bei diesem Lebenswandel 3 Jahre lebendig oder zumindest unverletzt zu übersteht ?

Für eine Eroberung:  $p(\text{kein Duell}) = (1 - 0,47) = 0,53$

$$p(\text{Duell gewonnen} | \text{Duell}) = 0,47 * 0,2 = 0,094$$

$$p(\text{überlebt}) = 0,53 + 0,094 = 0,624$$

$$\text{oder } 1 - p(\text{Duell verloren} | \text{Duell}) = 1 - 0,47 * 0,8 = 0,624$$

3 Jahre oder 9 Eroberungen unverletzt =  $0,624^9 = 0,01434$

**b) (2P)** Wenn er trotz lausiger Fähigkeiten aus purem Glück seine ersten 5 Duelle gewänne, eilte ihm ein Ruf voraus, der potenzielle Duellgegner derartig abschrecken würde, dass er fortan ungestraft den Damen nachstellen könnte. Wie groß ist die Wahrscheinlichkeit, dass er diesen paradisiatischen Zustand irgendwann erreicht ?

$p(5 \text{ Duelle, alle gewonnen}) = 0,2^5 = 0,00032$

## VL\_4

### Aufgabe 4.1:

Vielleicht kennen Sie das Würfelspiel „Lügenmäxchen“. Dabei wird mit einem (fairen) Würfelpaar gewürfelt und das höchste erreichbare Ergebnis ist eine Kombination einer 1 mit einer 2 (das sogenannte Mäxchen). Wie groß ist die Wahrscheinlichkeit, bei sieben Würfeln mit dem Würfelpaar genau 2 Mäxchen zu erhalten? (Erläutern Sie Ihren ausführlich dargestellten Rechenweg) **(8P)**

### Aufgabe 4.2: (9 P)

Ein im siebzehnten Jahrhundert verbreitetes Spiel bestand darin, ein Würfelpaar 24 mal zu werfen. Das Problem war, ob man auf ein Auftreten mindestens einer Doppelsechs eine Wette 1:2 abschließen sollte, oder ob das auf Dauer ein Verlustgeschäft wäre.

a) Wie groß ist die Wahrscheinlichkeit, mindestens eine Doppelsechs zu erhalten ?

b) Wie groß ist die Wahrscheinlichkeit, bei 24 Würfeln genau zwei mal eine Doppelsechs zu erhalten ?

#### **Aufgabe 4.3(9P):**

*Hinweis: Offiziell sind die folgenden Aufgaben durch Anwendung einer Verteilung, die wir in der Vorlesung nicht behandelt haben, zu lösen. Mit gesundem Menschenverstand und unseren Kenntnissen aus Kombinatorik und Wahrscheinlichkeitsrechnung lassen sich die Lösungen auch ohne auswendig gelernte Verfahren finden.*

a) Wie groß ist die Wahrscheinlichkeit, beim normalen Lotto „6 aus 49“ (6 Kugeln werden aus einem Behälter mit 49 nummerierten Kugeln ohne zurücklegen gezogen) mit einem Tip von 6 Zahlen sechs richtige zu erhalten?

b) wie groß ist die Wahrscheinlichkeit, keine der gezogenen Zahlen im Tip zu haben (0 richtige)?

c) Wie groß ist die Wahrscheinlichkeit, genau 3 richtige zu tippen?

*(teil c ist schwierig, deshalb dazu ein Lösungshinweis: Wie groß ist  $p$ , zunächst 3 richtige und dann 3 falsche zu ziehen? Wie groß sind die Wahrscheinlichkeiten für die anderen falsch/richtig Anordnungen? Wie viele solche falsch/richtig Anordnungen gibt es? Es gibt viele weitere Wege zur richtigen Lösung.)*

#### **Aufgabe 4.4:**

Zur Statistiklausur erscheinen im Durchschnitt 95 von 100 angemeldeten Studenten. Wie groß ist die Wahrscheinlichkeit, dass in der ersten Hörsaalreihe nur drei von 8 Studenten erscheinen? (Sie dürfen die Poisson-Verteilung als Näherung verwenden)

(Kommentieren Sie Ihren Rechenweg) **(6P)**

## **VL\_5**

#### **Aufgabe 5.1:**

Zu Semesterbeginn mussten bis vor kurzem alle Wedeler Studenten zur Einschreibung persönlich im SR1 erscheinen und sich in eine Schlange einreihen. Die Wartezeit sei normalverteilt mit  $\mu = 5$  Minuten und  $\sigma = 2$  Minuten.

Wie groß ist die Wahrscheinlichkeit, dass ein zufällig gewählter Student weniger als 3 Minuten warten muss? (Kommentieren Sie Ihren Rechenweg) **(6P)**

## **VL\_6, VL\_7 und VL\_8**

#### **Aufgabe 8.1:(3P)**

a) Welche Verteilung benutzen Sie für ein Konfidenzintervall, wenn Sie von einer Normalverteilung in der Grundgesamtheit ausgehen können und die Standardabweichung bekannt ist ?

b) Wenn wir den Stichprobenumfang vergrößern und alle anderen Zahlen konstant halten, wird dann das Konfidenzintervall größer oder kleiner ?

c) Wenn wir die Standardabweichung der Grundgesamtheit  $\sigma_n$  kennen und den Wert der Stichprobenstandardabweichung  $s_n$  berechnet haben, welchen Wert würden Sie zur Berechnung des Konfidenzintervalls verwenden?

#### **Aufgabe 8.2:(5P)**

6 Wirtschaftsinformatiker, die für start-up Firmen arbeiten, wurden nach ihrer wöchentlichen Arbeitszeit gefragt.

Die Daten (in Stunden) sind: 45, 55, 60, 55, 55, 53.

Berechnen Sie ein 90% Konfidenzintervall für die durchschnittliche Arbeitszeit der GG(= alle Wirtschaftsingenieure, die für start-up Firmen arbeiten)!

Kommentieren und erläutern Sie Ihre Rechnung!

### **Aufgabe 8.3: (10 P)**

Das Gewicht von Meerschweinchen sei normalverteilt mit einem Mittelwert von 1017 g und einer Standardabweichung von 89 g.

a) Wie groß ist die Wahrscheinlichkeit, dass ein beliebig herausgegriffenes Meerschweinchen zwischen 999 und 1009 g wiegt?

b) 7 beliebig ausgewählte Meerschweinchen sollen mit einem kleinen Heißluftballon fliegen. Der Heißluftballon hat eine für 7,400 kg ausreichende Auftriebskraft. Wie groß ist die Wahrscheinlichkeit, dass die 7 Meerschweinchen für den Flug insgesamt zu schwer sind und trotz Vorfreude am Boden bleiben?

### **Aufgabe 8.4:(3P)**

Ein Jungmanager in einer Firma, die Drahtseile für Aufzüge herstellt, hat in seinem Studium das "Null Fehler Prinzip" verinnerlicht und ordnet als erste Maßnahme in der neu angetretenen Führungsposition für die Produktion 0-Toleranz und 0-Fehler an. Für eine maximale Traglast von 800 kg werden nunmehr nur noch (vermeintlich) exakt bei 800 kg Traglast versagende Stahlseile ausgeliefert.

Tatsächlich raufen sich die Fertigungsingenieure die Haare und stellen heimlich die Maschinen so ein, dass im Mittel Seile einer maximalen Traglast von 840 kg produziert werden. Die Standardabweichung bei den Produktionsmaschinen beträgt für die Reißfestigkeit 30 kg Last.

Wie groß ist die Wahrscheinlichkeit, dass ein unter diesen Bedingungen produziertes Seil bereits bei geringerer Traglast als 800 kg versagt ?

## VL\_9

### Aufgabe 9.1: (10 P)

Sie finden einen ehemals für das Spiel aus Aufgabe 4.2 verwendeten Würfel aus dem siebzehnten Jahrhundert. Sie würfeln 999 mal und erhalten 133 mal eine sechs. Ist der Würfel fair? Eine Signifikanzzahl von 0.01 soll bei Ihrer Rechnung zugrunde gelegt werden.

### Aufgabe 9.2:(6P)

Nach einer Vorlesung zum Testen von Hypothesen wird einem Statistikdozenten nach unvorsichtig gemachten diesbezüglichen Bemerkungen ein angeknickter Kronkorken zugespielt.

Der Dozent geht davon aus, dass er den Kronkorken wie ein faire Münze (gleiche Wahrscheinlichkeit von 0,5 für Kopf und Zahl) als Entscheidungshilfe verwenden kann. Um diese Hypothese zu testen, wirft er den Kronkorken 40 mal. Dabei erhält er 27 mal die Lage "Krone" (pikst beim drauftreten) und 23 mal die Lage "Tisch" (der Markenaufdruck ist lesbar).

a) Lässt sich bei dieser Datenlage die Hypothese, es handle sich um eine "faire Münze" aufrecht erhalten? (Rechnen Sie und legen Sie eine Signifikanzzahl von 0,05 zugrunde.)

b) Sie haben den Test entweder zweiseitig oder einseitig ausgeführt. Begründen Sie Ihre Wahl.

### Aufgabe 9.3 (4 P):

Jemand behauptet Männer mit Krawatten, Anzügen und geputzten Schuhen, die auf der Mönckebergstraße herumlaufen, verdienen durchschnittlich 130000,-€ . Testen Sie diese Hypothese anhand Ihrer Stichprobe (nehmen Sie dabei an:  $\alpha = 0,05$ ).

Sie erhalten folgende Urliste:

MKAgs1: 90000, MKAgs2: 50000, MKAgs3: 110000, MKAgs4: 120000, MKAgs5: 80000, MKAgs6: 70000, MKAgs7: 80000, MKAgs8: 50000, MKAgs9: 60000, MKAgs10: 70000, MKAgs11: 230000, MKAgs12: 70000, MKAgs13: 120000

Test bei unbekannter Standardabw. -> Student-T Verteilung verwenden bei  $s_1 = 47600$  Euro,  $s_{13} = 47600/\sqrt{13} = 13202$  Euro,  $\mu = 92300$  Euro.

Standardisieren:  $t = (130000 - 92300) / 13202 = 2,8556$

Zweiseitiger Test für  $\alpha = 0,05$  -> T-Tabelle für 12 DOF, welcher t-Wert für  $p = 0,025$  ?

->  $t_{schwelle} = 2,1788$

Ist geringer als 2,8556, die wahre Abweichung ist also größer als die für Beibehaltung der Hypothese gerade noch zu tolerierende Abweichung.

Wir müssen die Hypothese verwerfen!

## VL\_10

### Aufgabe 10.1

Johann Gregor Mendel erhielt bei einem seiner Kreuzungsversuche von Erbsenpflanzen folgende Häufigkeiten:

315 runde gelbe Erbsen,

108 runde grüne Erbsen

101 kantige gelbe Erbsen

32 kantige grüne Erbsen

Sprechen die Ergebnisse für oder gegen die Theorie, dass das Verhältnis der Häufigkeiten 9:3:3:1 sein sollte ( $\alpha = 0.05$ ) ?

(Hinweis: Führen Sie einen Chi-Quadrat-Anpassungstest durch, wie viele Freiheitsgrade liegen vor, was sind die Werte der erwarteten Häufigkeiten, ...?)



#### Aufgabe 10.4:

Der Sohn eines Wedeler Statistikdozenten formte im Sommerurlaub aus Lehm Würfel für Spiele um hohe Einsätze (Wer spült Geschirr etc...). Wegen der etwas unsymmetrischen Form eines Würfels und unerfreulicher Spielausgänge äußerte der Dozent Zweifel an der "Fairness" des Würfels d.h. der Gleichverteilung der Ergebnisse 1-6.

Ein Experiment sollte unter der Annahme  $\alpha = 0,05$  Aufschluss über die Fairness des Würfels geben. Bei 205 Würfeln wurde folgendes Ergebnis erzielt:

Ergebnis	1	2	3	4	5	6
Häufigkeit	34	43	22	40	24	42



Kann der Würfel als fair betrachtet werden ( $\alpha = 0,05$ ) ?

	1	2	3	4	5	6	
erhalten	<b>34</b>	<b>43</b>	<b>22</b>	<b>40</b>	<b>24</b>	<b>42</b>	Würfe: 205
erwartet	34,17	34,17	34,17	34,17	34,17	34,17	ChiQuad:
Chiquad	0,001	2,284	4,333	0,996	3,025	1,796	<b>12,43</b>

DOF : **5**

Alternativ:

chisqr(alfa): **11,07**      CHITEST: **0,0293**

p-value: **0,0293**

Der zum  $\alpha$ -Wert gehörende chiquadrat-Wert von 11,07 ist kleiner als der tatsächliche Wert. Alternativ: Der p-Wert von 0,0293 ist kleiner als der  $\alpha$ -Wert.

Die Hypothese muss also verworfen werden: Der Würfel ist nicht fair!

#### Aufgabe 10.5:

Ein großer Telekommunikationsprovider, die „Knebel Deutschland AG“ beschließt, höhere Kundenzahlen dadurch zu generieren, dass sie alle Kündigungen von Verträgen wegschmeißt und ignoriert. Damit das nicht sofort auffliegt, stellt sie den Schriftverkehr mit den Kunden ein, versucht aber trotzdem, Monate später noch Abbuchungen von deren Konten vorzunehmen und führt stillschweigend die Verträge weiter.

Wenn der Verjährungstermin der daraus resultierenden Ansprüche gegenüber den nichtsahnenden Ex-Kunden näher rückt, gibt sie diese Forderungen an ein Inkassounternehmen weiter, weil die Beschwerdeabteilung keine Lust auf Schriftverkehr mit geprellten und erbosten Ex-Kunden hat.

Geschäftspartner der *Knebel Deutschland AG* in dieser Sache ist das Inkassounternehmen der Anwaltssozietät „*Schneider,Aufschneider&Abschneider*“, das nun versucht, mit Drohbrieffen und Belästigungen durch Inkassogebühren aufgeblähte Geldbeträge einzutreiben.

Die *Knebel Deutschland AG* gibt in der ersten Tranche 23939 Forderungen dieser Art an „*Schneider,Aufschneider&Abschneider*“ ab und versichert dem Anwalt Aufschneider schriftlich, sie habe aus langjähriger Erfahrung sichere Kenntnisse über die Erfolgswahrscheinlichkeit von Inkasso-Drohmaßnahmen: Es zahlen in der Altersgruppe unter 25J 20% der Betroffenen, in der Gruppe 25J-45J 11%, in der Gruppe 45J-65J 4% und in der Gruppe über 65J sogar 35% .

Die folgende Tabelle gibt die Struktur der Tranche nach Altersgruppen, die nach erfolgter Inkasso Prozedur tatsächliche Anzahl der Zahlungen von Ex-Kunden an *Schneider,Aufschneider&Abschneider* und die Behauptung der *Knebel Deutschland AG* wieder:

unter 25	25-45	45-65	über 65
7312	7212	6398	3017
1533	678	210	990
0,2000	0,1100	0,0400	0,3500

Stimmt die Behauptung der *Knebel Deutschland AG* über die Zahlwahrscheinlichkeiten in den Altersgruppen?

(Weil die Messlatte für einen Betrugsvorwurf vor Gericht hoch hängen sollte, nehmen Sie eine Konfidenzzahl von  $\alpha=0,0001$  beim Chiquadrat Anpassungstest an. Wegen des extremen Wertes von  $\alpha$  finden Sie ein erweitertes Tabellenblatt zur Chi<sup>2</sup>-Verteilung in der Klausur.)

a) Stellen Sie eine vollständige Tabelle zur Ermittlung des Chi<sup>2</sup> Wertes zusammen und bestimmen Sie Chi<sup>2</sup>.

Altersgruppe	unter 25	25-45	45-65	über 65
Anzahl	7312	7212	6398	3017
tatsächl. Zahlungen	1533	678	210	990
angegebene Zahlwahrsch.	0,2000	0,1100	0,0400	0,3500
erw. chi <sup>2</sup>	1462,4	793,32	255,92	1055,95
	3,408	16,763	8,239	4,119

chi<sup>2</sup> 32,53  
chi<sup>2</sup>-Schwelle 21,11  
p-Wert 4,05E-007

b) Bestimmen Sie die Anzahl der Freiheitsgrade. 3

c) Testen Sie die Hypothese, dass die *Knebel Deutschland AG* bei der Angabe der Zahlwahrscheinlichkeiten gegenüber *Schneider, Aufschneider & Abschneider* die Wahrheit gesagt hat mit der Konfidenzzahl  $\alpha=0,0001$ .

Weil der p-Wert kleiner als 0,0001 ist hat die *Knebel Deutschland AG* hier offensichtlich die Unwahrheit gesagt.

# VL\_11

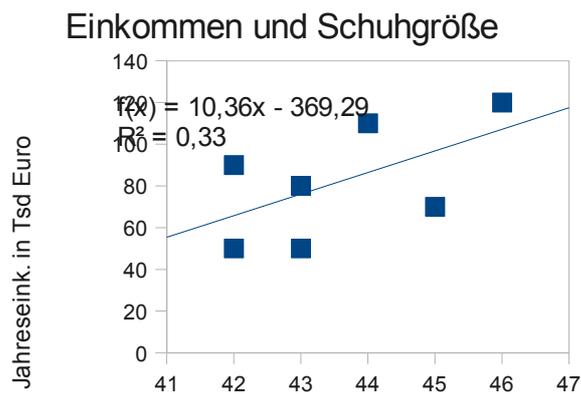
## Aufgabe 11.1:

Sie fragen auf der Mönckebergstraße in Hamburg 8 Männer, die Krawatte, Anzug und geputzte Schuhe tragen (MKAgs) nach dem aktuellen Jahreseinkommen und der Schuhgröße:

MKAgs1: (90000, 42) MKAgs2: (50000, 43), MKAgs3: (110000, 44), MKAgs4: (120000, 46)  
 MKAgs5: (80000, 43), MKAgs6: (70000, 45), MKAgs7: (80000, 43), MKAgs8: (50000, 42)

a) (2P) Zeichnen Sie einen sogenannten Scatter-Plot der Daten (mit richtigen Achsenbezeichnungen).

Schuhgr.	EK in Tsd
42	90
43	50
44	110
46	120
43	80
45	70
43	80
42	50



b) (4P) Berechnen Sie über Varianz und Kovarianz die Konstanten a und b für die Lineare Regression und den Korrelationskoeffizienten. Geben Sie außerdem die konkrete Form der ermittelten Regressionsgerade (mit physikalischen Einheiten) an.

Erw.Eink. = 10,36 Tsd€/Schuhgr \* x - 369,29 Tsd€

Schuhgr.	EK in Tsd	yvar	xvar	kovar
42	90	76,56	2,25	-13,13
43	50	976,56	0,25	15,63
44	110	826,56	0,25	14,38
46	120	1501,56	6,25	96,88
43	80	1,56	0,25	0,63
45	70	126,56	2,25	-16,88
43	80	1,56	0,25	0,63
42	50	976,56	2,25	46,88

mitt		4487,5	14,00	145,00
43,5	81,25			

xVarianz	2,00000	<b>b:</b>	<b>10,36</b>
kovarianz	20,71	<b>a:</b>	<b>-369,29</b>
yVarianz	641,07	<b>r:</b>	<b>0,58</b>
		r^2:	0,33

c) (2P) Kommentieren Sie den Wert des Korrelationskoeffizienten. Besteht eine starke Korrelation? Wie ändert sich der Korrelationskoeffizient, wenn Sie das Einkommen statt in Euro in DM angeben würden? (Hinweis 1,95 DM ≈ 1€)

Keine Überzeugende Korrelation, aber nicht "NICHTS". In DM gerechnet würde der gleiche Korrelationskoeffizient herauskommen, da er dimensionslos ist. (a und b hätten natürlich andere Zahlenwerte und andere Einheiten.)