Klausur: **Statistik**

Jürgen Meisel

Zugelassene Hilfsmittel: nicht progr. Taschenrechner
Bearbeitungszeit: **120 Minuten**

Statistik

1.) Mittelwerte und Streumaße: Taschengeld

Eine Umfrage unter 240 Schülern ergab folgende Verteilung des monatlichen Taschengeldes in €:

Klasse	$0 \le x < 5$	$5 \le x < 10$	$10 \le x < 20$	$20 \le x < 30$	$30 \le x < 50$	$50 \le x < 75$
rel. H.	10,8 %	15 %	21,8 %	15,8 %	19,1 %	17,5 %

a) Bestimmen Sie die absoluten Häufigkeiten.

b) Ermitteln Sie das arithmetische Mittel.

c) Berechnen Sie die Varianz und die Standardabweichung.

d) Erstellen Sie ein Histogramm.

e) Bestimmen Sie nun noch den Median und den Modus.

f) Erstellen/zeichnen Sie nun einen kompletten Boxplot mit einer Berechnung der Quartile 1 und 3.

g) Wie viele der befragten Schüler erhalten mind. 20 € Taschengeld?

h) Wie viel Taschengeld erhält ein Kind, das zum „ärmsten" Viertel gerechnet wird im Durchschnitt?

i) Wie viel Taschengeld geben die Eltern insgesamt monatlich aus?

Lösung:

Klasse	$0 \le x < 5$	$5 \le x < 10$	$10 \le x < 20$	$20 \le x < 30$	$30 \le x < 50$	$50 \le x < 75$
rel. H.	10,8 %	15 %	21,8 %	15,8 %	19,1 %	17,5 %
r. SH	0,108	0,258	0,476	0,634	0,825	1,000
r. HD	0,0216	0,03	0,0218	0,0158	0,00955	0,007
abs. H	26	36	52	38	46	42
Mitte	2,5	7,5	15	25	40	62,5

Mittelwert: $\mu = \dfrac{6.530}{240} = 27{,}21$

Varianz: $\sigma^2 = \dfrac{275.300}{240} - 27{,}21^2 = 406{,}79 \quad \rightarrow \quad \sigma = 20{,}17$

Median:

$$\overline{x_M} = x_{k-1} + \frac{x_k - x_{k-1}}{f(x_k)} \cdot \left[\frac{1}{2} - F(x_{k-1})\right]$$

$$\overline{x_M} = 20 + \frac{30-20}{0{,}158} \cdot \left[\frac{1}{2} - 0{,}476\right] = 21{,}52$$

Modus: Klassenmitte von Klasse 3: $\overline{x_{Modus}} = 15$

Quartile 1 und 3:

$$\overline{x_{q_1}} = x_{k-1} + \frac{x_k - x_{k-1}}{f(x_k)} \cdot \left[\frac{1}{2} - F(x_{k-1})\right]$$

$$\overline{x_{q_1}} = 5 + \frac{10-5}{0{,}15} \cdot \left[\frac{1}{4} - 0{,}108\right] = 9{,}73$$

$$\overline{x_{q_3}} = 30 + \frac{50-30}{0{,}191} \cdot \left[\frac{3}{4} - 0{,}634\right] = 42{,}15$$

126 Schüler erhalten mindestens 20,00 € Taschengeld.

Durchschnitt „ärmstes Viertel“: $\mu = \dfrac{2{,}5 \cdot 0{,}108 + 7{,}5 \cdot 0{,}15}{0{,}258} = 5{,}41$

Taschengeldsumme im Monat: 6.530,00 € (vgl. Mittelwert)

2.) Spearmanscher Rangkorrelationskoeffizient

Bei einem Turnwettkampf haben 8 Turner die folgenden Bewertungen an zwei Geräten erhalten:

Turner Nr.	1	2	3	4	5	6	7	8
Boden	9,6	9,2	9,4	9,4	9,3	9,5	9,8	9,5
Seitpferd	9,1	9,8	9,6	9,9	9,7	9,4	9,2	9,7

Berechnen Sie den Spearmanschen Rangkorrelationskoeffizienten.

Zusatzfrage: Welche weiteren 4 Geräte gibt es beim Herrenturnen noch?

Lösung:

Teilnehmer	Boden	Seitpferd	d_i	d_i^2
1	2	8	6	36,00
2	8	2	6	36,00
3	5,5	5	0,5	0,25
4	5,5	1	4,5	20,25
5	7	3,5	3,5	12,25
6	3,5	6	2,5	6,25
7	1	7	6	36,00
8	3,5	3,5	0	0,00
				147,00

$$r_S = 1 - \frac{6 \cdot 147}{8 \cdot 63} = -0,75$$

Zusatzfrage: Barren; Reck; Sprung; Ringe

3.) Lineare Regression und Korrelation

Der Ernteertrag eines landwirtschaftlichen Gutes von der Menge eines eingesetzten Düngemittels je Hektar ab.

Auf 6 Versuchsfeldern wird der Düngemitteleinsatz getestet.

Dabei wurden die folgenden Erträge je Hektar erzielt:

Versuchsfeld	1	2	3	4	5	6	
Düngemittel-einsatz [100 kg]	6	3	8	2	7	2	
Ernteertrag [t]	30	10	22	14	36	24	

a) Stellen Sie den Ernteertrag in Abhängigkeit von dem Düngemitteleinsatz durch eine Funktion $y = b_0 + b_1 x$ dar.

b) Wie hoch wäre demnach der Ernteertrag bei 1 Tonne Düngemittel?

c) Wie groß ist der Korrelationskoeffizient?

Lösung:

$$y = b_0 + b_1 x \rightarrow y = 12{,}188 + 2{,}245x$$

$$y_{10} = 12{,}188 + 2{,}245 \cdot 10 = 34{,}64$$

$$r = 0{,}616$$

4.) Preisindizes

Familie Winzig leistet sich leistet sich folgende Produkte und Ausgaben.

Für die einzelnen Jahre wurde aus dem Haushaltsbuch folgende Übersicht erstellt:

	Preis 2007	Umsatz 2007	Preis 2010	Verbrauch 2010
Chips	1,00 € / Tüte	40,00 €	1,50 € / Tüte	50 Tüten
Erdnüsse	6,00 € / kg	36,00 €	9,00 € / kg	7 kg
Bier	4,20 € / ltr.	504,00 €	5,50 € / ltr.	150 ltr.

a) Berechnen Sie den Preisindex nach Laspeyres für 2010 zur Basis 2007.

b) Wie groß ist der durchschnittliche jährliche Preisanstieg in %?

Lösung:

$$L_P = \frac{1{,}5 \cdot 40 + 9 \cdot 6 + 5{,}5 \cdot 120}{40 + 36 + 504} = 1{,}3345$$

$$g = \sqrt[3]{1{,}3345} = 1{,}10095 \rightarrow 10{,}1[\%]$$

Stochastik

1.) Hypergeometrische Verteilung

Eine Kegelrunde aus 12 Personen unternehmen eine Kaffeefahrt nach Helgoland und müssen nach der Rückfahrt durch die Zollkontrolle. Obwohl alle angeben, nur die erlaubte Menge Zigaretten und Alkohol eingekauft zu haben, haben Knut und Gustav zu viel Zigaretten mitgenommen.
Der Zollbeamte wählt drei von den 12 aus, um sie zu durchsuchen.

a) Mit welcher Wahrscheinlichkeit erwischt der Zollbeamte keinen Schmuggler?

b) Mit welcher Wahrscheinlichkeit erwischt der Zollbeamte mindestens einen der beiden Schmuggler?

c) Wie viele Personen müsste der Zollbeamte mindestens kontrollieren, um mit einer Wahrscheinlichkeit von mindestens 50 % mindestens einen Schmuggler zu überführen?

Lösung:

$$a)\quad H(X=0) \;=\; \frac{\binom{2}{0}\binom{10}{3}}{\binom{12}{3}} \;=\; 0,5454$$

$$b)\quad H(X\geq 1) \;=\; 1-H(X=0) \;=\; 0,4545$$

$$c)\quad H(X\geq 1) \;\geq\; 0,5 \;\rightarrow\; 1-H(X=0) \;\geq\; 0,5 \;\rightarrow\; H(X=0) \;\leq\; 0,5$$

$$\rightarrow\quad H_t(X=0) \;=\; \frac{\binom{2}{0}\binom{10}{t}}{\binom{12}{t}} \;\leq\; 0,5$$

$$\rightarrow\quad \frac{1\cdot\dfrac{10!}{t!\cdot(10-t)!}}{\dfrac{12!}{t!\cdot(12-t)!}} \;\leq\; 0,5 \;\rightarrow\; \frac{10!\cdot(12-t)!}{12!\cdot(10-t)!} \;\leq\; 0,5 \;\rightarrow\; \frac{(12-t)(11-t)}{11\cdot 12} \;\leq\; 0,5$$

$$\rightarrow\quad t^2-23t+66 \;\leq\; 0 \;\rightarrow\; t_1=3,36 \quad und \quad t_2=19,64\,[keine\ Lösung] \;\rightarrow\; t=4$$

2.) Binomialverteilung: Polizeikontrolle und Alkoholtest

In Zwergstadt an einer bestimmten Stelle führt die Polizei in regelmäßigen Abständen in der Nacht von Freitag auf Samstag zwischen 23.00 Uhr und 2.00 Uhr Verkehrskontrollen durch. Dabei muss der Fahrer "in die Röhre pusten", um festzustellen, ob der Alkoholgehalt im Blut im gesetzlich erlaubten Rahmen liegt oder nicht.
Aus mehrjähriger Erfahrung weiß die Polizei, dass ungefähr 15 % aller männlichen und 5 % aller weiblichen Fahrer um diese Zeit an dieser Stelle die "Promillegrenze" überschreiten.
Wir nennen diese Personen hier kurz "Alkoholsünder".
Am letzten Wochenende wurden 100 Verkehrsteilnehmer überprüft.
Darunter befanden sich 35 Frauen.

a) Stellen Sie den Sachverhalt in einem Baumdiagramm oder einer Vierfeldertafel dar.

Lösung:

	A (Alk.-Sünder)	$\overline{A}$ (kein Alk.-Sünder)	
M (männlich)	0,0975	0,5525	0,65
F (weiblich)	0,0175	0,3325	0,35
	0,115	0,885	1

Nebenrechnung:
$$P(M \cap A) = 0{,}15 \cdot 0{,}65 = 0{,}0975$$
$$P(F \cap A) = 0{,}05 \cdot 0{,}35 = 0{,}0175$$

Baumdiagramm: Stufe 1: M/F Stufe 2: $A/\overline{A}$

b) Wie viel Prozent aller Fahrer sind **keine** Alkoholsünder?

Lösung: $$P(\overline{A}) = P(M \cap \overline{A}) + P(F \cap \overline{A}) = 0{,}5525 + 0{,}3325 = 0{,}885$$

c) Berechnen Sie für die zufällige Auswahl einer Person die Wahrscheinlichkeit folgender Ereignisse:

(i) Die überprüfte Person ist weiblich und Alkoholsünderin.
(ii) Die überprüfte Person ist nüchtern.
(iii) Die überprüfte Person ist männlich.
Mit welcher Wahrscheinlichkeit liegt Trunkenheit am Steuer vor?

Lösung:

(i) $P(F \cap A) = 0,0175$

(ii) $P(\overline{A}) = P(M \cap \overline{A}) + P(F \cap \overline{A}) = 0,5525 + 0,3325 = 0,885$

(iii) $P_M(A) = \dfrac{P(M \cap A)}{P(M)} = \dfrac{0,65 \cdot 0,15}{0,65} = 0,15$

Im Folgenden wird davon ausgegangen, dass 10 % aller Verkehrsteilnehmer, die an der entsprechenden Stelle kontrolliert werden Alkoholsünder sind. Weiterhin wird angenommen, dass die Anzahl der Alkoholsünder in den Kontrollen einer Binomialverteilung genügt.

d) Mit welcher Wahrscheinlichkeit findet man in einer Zufallsstichprobe unter 100 überprüften Personen:
- (i) genau 8 Alkoholsünder?
- (ii) weniger als 10 Alkoholsünder?
- (iii) mehr als 14 Alkoholsünder?
- (iii) mind. 5 und höchstens 14 Alkoholsünder?

Lösung:

(i) $B(X=8) = B(X \leq 8) - B(X \leq 7) = 0,3209 - 0,2061 = 0,1148$

(ii) $B(X<10) = B(X \leq 9) = 0,4513$

(iii) $B(X>14) = 1 - B(X \leq 14) = 1 - 0,9274 = 0,0726$

(iv) $B(5 \leq X \leq 14) = B(X \leq 14) - B(X \leq 4) = 0,9274 - 0,0237 = 0,9037$

e) Wie viele Personen muss man mind. überprüfen, damit mit einer Wahrscheinlichkeit von mind. 99,9 % mind. ein Alkoholsünder dabei ist?

Lösung:

$$B(X \geq 1) \geq 0,999 \rightarrow 1 - B(X=0) \geq 0,999 \rightarrow B(X=0) \leq 0,001$$

$$\rightarrow \binom{n}{0} 0,1^0 0,9^n \leq 0,001 \rightarrow 0,9^n \leq 0,001 \xrightarrow{\ln} n \cdot \ln 0,9 \leq \ln 0,001$$

$$\xrightarrow{:\ln 0,9} n \geq \frac{\ln 0,001}{\ln 0,9} \geq 65,56 \rightarrow n \geq 66$$

3.) Normalverteilung und Satz von Bayes bei Autoreifen

Autoreifen haben eine mittlere Lebensdauer von 50.000 Kilometern mit einer Standardabweichung von 6.000 Kilometern.

Die Reifen werden bei der Unternehmung Goodmonth in drei Produktionshallen hergestellt, wobei 40% in Halle 1, 25 % in Halle 2 und der Rest in Halle 3 produziert werden. Dabei werden auch Mängelprodukte erzeugt. 5 %, 8 %, und 6 % entsprechend der Hallenreihenfolge.

a) Wie groß ist die Wahrscheinlichkeit, einen fehlerhaften Reifen zu produzieren?

Lösung:

$$P(\text{"}fehlerhaft\text{"}) = 0,4 \cdot 0,05 + 0,25 \cdot 0,08 + 0,35 \cdot 0,06 = 0,061 = 6,1[\%]$$

b) Bei Produkttests wurde ein fehlerhafter Reifen gefunden. Wie wahrscheinlich ist es, dass er in Halle 2 hergestellt wurde?

Lösung: Satz von Bayes

$$P_{(\text{"}fehlerhaft\text{"})}(\text{"}Halle\ 2\text{"}) = \frac{P(f \cap H2)}{P(f)}$$

$$P_{(\text{"}fehlerhaft\text{"})}(\text{"}Halle\ 2\text{"}) = \frac{0,25 \cdot 0,08}{0,061} = 0,3278 = 32,78[\%]$$

c) Wie viel % der erzeugten Reifen haben eine Lebensdauer von

(i) mindestens 60.000 Kilometern?
(ii) höchstens 48.000 Kilometern?
(iii) zwischen 45.000 und 55.000 Kilometern?

Lösung:

(i)

$$P(X \geq 60.000) \quad = \quad 1 - P(X \leq 60.000) \quad = \quad 1 - \Phi\left(\frac{5}{3}\right) \quad = \quad 1 - 0,95254 \quad = \quad 0,04746$$

$$Umrechnung: \quad x \quad = \quad \frac{60.000 - 50.000}{6.000} \quad = \quad \frac{5}{3} \quad = \quad 1,67$$

(ii)

$$P(X \leq 48.000) \quad = \quad \Phi\left(-\frac{1}{3}\right) \quad = \quad 1 - \Phi\left(\frac{1}{3}\right) \quad = \quad 1 - 0,62930 \quad = \quad 0,3707$$

$$Umrechnung: \quad x \quad = \quad \frac{48.000 - 50.000}{6.000} \quad = \quad -\frac{1}{3}$$

(iii)

$$P(45.000 \leq X \leq 55.000) \quad = \quad 2 \cdot \Phi\left(\frac{5}{6}\right) - 1 \quad = \quad 2 \cdot 0,79673 - 1 \quad = \quad 0,59346$$

$$Umrechnung: \quad x \quad = \quad \frac{55.000 - 50.000}{6.000} \quad = \quad \frac{5}{6}$$

d) Bei wie viel % der Reifen weicht die Lebensdauer um höchstens 2σ Kilometer vom erwarteten Mittelwert ab?

Lösung:

$$P(38.000 \leq X \leq 62.000) \quad = \quad 2 \cdot \Phi(2) - 1 \quad = \quad 2 \cdot 0,97725 - 1 \quad = \quad 0,9545$$

$$Umrechnung: \quad x \quad = \quad \frac{62.000 - 50.000}{6.000} \quad = \quad 2 \quad [\text{wegen } 2\sigma\text{-Umgebung}]$$

$$\rightarrow \quad 1 - 0,9545 \quad = \quad 0,0455 \quad = \quad 4,55\,[\%]$$