

Aufgabe 1

Sind die folgenden Merkmale diskret oder stetig?

- 1) Geschwindigkeit – stetig
- 2) Hörerzahl einer Vorlesung – diskret
- 3) Anzahl der Mitarbeiter – diskret
- 4) Einkommen – approximativ stetig
- 5) Zeit für die Beschleunigung – stetig
- 6) Bücherbestand – diskret
- 7) Stromverbrauch – stetig
- 8) Bierkonsum – stetig

Aufgabe 2

Welche Skalierung liegt vor?

- 1) Die Merkmalsausprägungen drücken lediglich eine Verschiedenartigkeit aus. (nominal)
- 2) Die Merkmalsausprägungen bringen neben der Verschiedenartigkeit auch eine natürliche Rangordnung zum Ausdruck. (ordinal)
- 3) Die Merkmalsausprägungen lassen auch zu, die Abstände zwischen den Merkmalsausprägungen zu vergleichen. (kardinal)

Aufgabe 3

Welches Skalenniveau besitzen die folgenden Merkmale?

- 1) Militärdienstgrad – ordinal
- 2) Alter – kardinal
- 3) Geschlecht – nominal
- 4) Nationalität – nominal
- 5) Intelligenzquotient – ordinal
- 6) Semesterzahl – kardinal
- 7) Anzahl von Verkehrsunfällen – kardinal
- 8) Tarifklassen bei der Kfz-Haftpflicht – ordinal
- 9) Studienfach – nominal
- 10) Fahrpreise – kardinal
- 11) Schulbildung – ordinal
- 12) Krankheit einer Person – nominal
- 13) Erlerner Beruf – nominal
- 14) Seitenzahl von Büchern – kardinal
- 15) Restaurantsterne – ordinal
- 16) Wahlergebnis verschiedener Parteien – kardinal

Aufgabe 4

Welche der folgenden Merkmale sind diskret oder stetig, quantitativ oder qualitativ, nominal-, ordinal- oder kardinalskaliert:

Aufgabe	d/s	n/o/k	Quan./qual.
Geschlecht	diskret	nominal	qualitativ
Autotyp	diskret	nominal	qualitativ
Zahl der Zuschauer bei den Spielen der Fußballbundesliga	diskret	kardinal	quantitativ
Einschätzung der Wirtschaftslage mit Antwortmöglichkeiten: gut, schlecht	diskret	ordinal	qualitativ
Betriebsdauer von Kaffeemaschinen	stetig	kardinal	quantitativ

Aufgabe 5

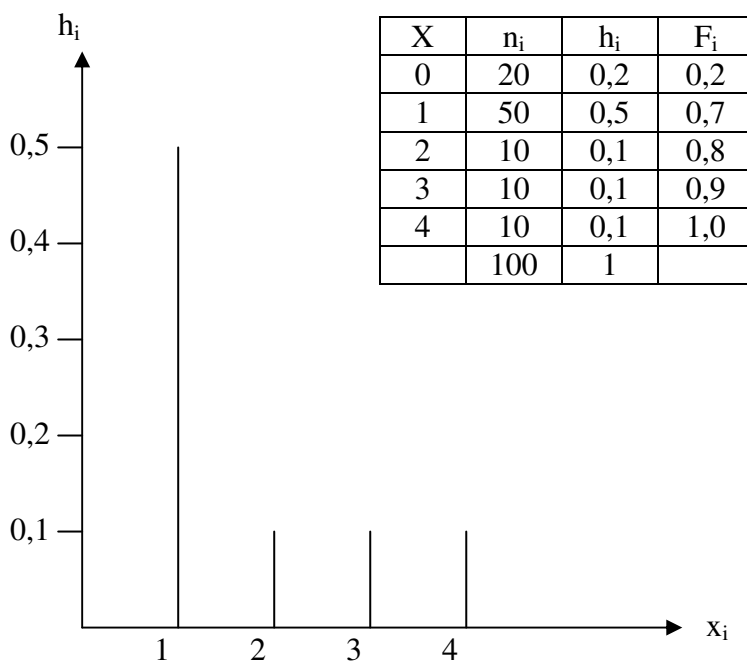
In einem Unternehmen sind 5000 Personen beschäftigt, davon sind 1500 Frauen. Stellen Sie die Häufigkeitsverteilung des Merkmales A: „Geschlecht eines Beschäftigten in diesem Unternehmen“ tabellarisch und graphisch dar.

A	n_i	$h_i = \frac{n_i}{n}$	
Frauen	1500	0,3	108°
Männer	3500	0,7	252°
	5000	1	

Aufgabe 6

Eine Umfrage ergibt, dass von 100 Personen 20 Personen keine Urlaubsreise, 50 Personen genau eine Urlaubsreise, 10 Personen zwei Urlaubsreisen, 10 Personen drei Urlaubsreisen und 10 Personen vier Urlaubsreisen im Jahr 2002 unternommen haben.

- Erstellen Sie die Häufigkeitsverteilung für diese Stichprobe
- Geben Sie eine graphische Darstellung der relativen sowie der kumulierten Häufigkeitsverteilungen.



Aufgabe 7

Folgende Daten wurden in einem IHK-Berzirk für eine bestimmte Branche ermittelt:

- Charakterisieren Sie das Merkmal X: „Umsatz“ so genau wie möglich.
- Bestimmen Sie die absoluten und relativen Häufigkeiten, sowie die Verteilungsfunktion von X.
- Stellen Sie die relative HV sowie die Verteilungsfunktion graphisch dar.

Umsatz in 100.000 € (von mehr als ... bis einschl.)	Anzahl der Unternehmungen
0 – 2	40
2 – 8	50
8 – 10	10

X	n _i	h _i	Δx _i	f _i = $\frac{h_i}{\Delta x_i}$	F _i
0-2	40	0,4	2	0,2	0,4
2-8	50	0,5	6	0,083	0,9
8-10	10	0,1	2	0,05	1
	100	1,0			

Häufigkeitspolygon
Polygonzug

Modus

- ✚ ungruppiert Daten: Ausprägung mit der max. absoluten/relativen Häufigkeit
- ✚ gruppierte Daten: Klassenmitte der Klasse mit max. Dichte

Aufgabe 8

Eine Befragung von 927 Männern nach der Anzahl ihrer Krawatten ergab folgendes Ergebnis:

Anzahl der Krawatten	1	2	3	4	5	6	7	8	9	mehr als 9
Häufigkeit	128	202	129	84	16	202	54	39	31	19

Bestimmen Sie den Modus. $M_1=2$, $M_2=6$

Median

- ✚ ungruppierte Daten

- n ungerade $Z = X * \left(\frac{n+1}{2}\right)$
- n gerade $Z = 0,5 * \left(X * \left(\frac{n}{2}\right) + X * \left(\frac{n}{2} + 1\right) \right)$

Aufgabe 9

- a) Bei einer Klassenprüfung wurden von 17 Teilnehmern folgende Punktzahlen erzielt:
14, 11, 10, 8, 9, 8, 10, 13, 12, 13, 14, 10, 11, 13, 14
Bestimmen Sie den Median
- b) Bei einem Stabhochsprungwettbewerb werden von 8 Springern folgende Höhen übersprungen:
4,90; 5,00; 4,60; 4,80; 5,00; 5,10; 4,80; 5,30
Bestimmen Sie den Median.

zu a)

8, 8, 9, 10, 10, 10, 11, 11, 12, 12, 13, 13, 13, 14, 14, 14

$$Z = X * \left(\frac{n+1}{2}\right) = X * \left(\frac{17+1}{2}\right) \quad Z = 12$$

zu b)

Zunächst ordnen

$$Z = 0,5 * (x_4 + x_5) = (4,90 + 5,00) = 4,95$$

Aufgabe 10

Der Student Paul hat an 7 Tagen einer Woche folgende Mengen Bier getrunken (Angabe in Liter): 0,7; 1,6; 2,5; 3,2; 1,6; 2,4; 2,8.

Bestimmen Sie das arithmetische Mittel.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k x_i n_i = \frac{1}{7} * (0,7*1+1,6*2+2,5*1+3,2*1+2,4*1+2,8*1) = \frac{14,8}{7} \approx 2,1143$$

$$\bar{x} = \frac{1}{n} \sum (0,7+1,6+2,5+3,2+1,6+2,4+2,8) = \frac{14,8}{7} \approx 2,1143$$

Aufgabe 11

x_i	n_i	h_i	Δx_i	$f_i \frac{h_i}{\Delta x_i}$	F_i
0-2	40	0,4	2	0,200	0,4
2-8	50	0,5	6	0,083	0,9
8-10	10	0,1	2	0,050	1
	100	1,0			

Modus bei gruppierten Daten

Gruppenmitte der Gruppe mit max. Dichte (D=1)

Median bei gruppierten Daten $F(z)=0,5$ – liegt in der zweiten Gruppe

x_i	F_i
200.000	0,4
Z	0,5
800.000	0,9

$\frac{dx}{Dx} = \frac{dF}{DF}$

$$Dx = Z - 200.000 \quad \frac{Z - 200.000}{600.000} = \frac{0,1}{0,5}$$

$$Dx = 800.000 - 200.000 = 600.000 \quad Z - 200.000 = 120.000$$

$$dF = 0,5 - 0,4 = 0,1 \quad Z = 320.000$$

$$DF = 0,9 - 0,4 = 0,5$$

Mittelwerte:

$$\bar{x} = \frac{1}{n} \sum_{j=1}^k x_j^* \cdot n_j = \frac{1}{100} (100.000 \cdot 40 + 500.000 \cdot 50 + 900.000 \cdot 10) = 380.000$$

Aufgabe 12

Spanweite: Differenz zwischen größten und kleinsten Ausprägung

a) $R = 5,0 - 1,7 = 3,3$

b) $R = 16 - 6 = 10$

Aufgabe 13

$$D = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}|$$

$$\bar{x} = \frac{1}{5} (3 + 7 + 8 + 9 + 13) = \frac{40}{5} = 8$$

$$d = \frac{1}{5} * (|3-8| + |7-8| + |8-8| + |9-8| + |13-8|) = 2,4$$

Aufgabe 14

$$s^2 = \frac{1}{n} \sum_{i=1}^k (x_i - \bar{x})^2 * n_i$$

$$\bar{x} = \frac{1}{10} (1,40 + 1,50 * 2 + 1,60 * 4 + 1,70 + 1,80 * 2) = 1,61$$

$$s^2 = \frac{1}{10} * ((1,40 - 1,61)^2 * 1 + (1,50 - 1,61)^2 * 2 + (1,60 - 1,61)^2 * 4 + (1,70 - 1,61)^2 * 1 + (1,80 - 1,61)^2 * 2)$$

$$= \frac{1}{10} (0,0441 + 0,0242 + 0,0004 + 0,0081 + 0,0722) = 0,0149$$

$$\text{Standartabweichung: } S = \sqrt{0,0149} \approx 0,1221$$

Aufgabe 15

Einkommen von – bis	n_i	x_j^*	$x_j^* * n_j$	$(x_i - \bar{x})^2 * n_i$
20 – 40	12	30	360	34604,28
40 – 60	32	50	1600	36342,80
60 – 80	70	70	4900	13138,30
80 – 120	56	100	5600	14878,64
120 – 160	28	140	3920	88751,32
160 – 200	2	180	360	18547,38
Summe	200		16740	206262

$$(30-83,7)^2 * 12$$

$$(50-83,7)^2 * 32$$

$$(70-83,7)^2 * 70$$

$$(100-83,7)^2 * 56$$

$$(140-83,7)^2 * 28$$

$$(180-83,7)^2 * 2$$

$$\bar{x} = \frac{1}{n} \sum x_j^* * n_j = \frac{1}{200} * 16740 = 83,7$$

$$s^2 = \frac{1}{200} * 206262 = 1031,31 \quad S = \sqrt{1031,31} \approx 32,1140$$

VC = Variationskoeffizient

$$VC = \frac{s}{\bar{x}} = \frac{32,1140}{83,7} = 0,3836$$

Aufgabe 16

X: Seitenlänge

$$\bar{x} = 1,5$$

$$S = 0,05$$

$$(s)^2 \Rightarrow s^2 = (0,05)^2 = 0,0025$$

$$VC = \frac{s}{\bar{x}} = \frac{0,05}{1,5} = 0,0333$$

$x^2 = \text{Fläche}$

$$\overline{x^2} \neq \bar{x}^2$$

Bsp.:

$$x_i \quad 1 \quad 3$$

$$\bar{x} = 2 \quad \bar{x}^2 = 4 \quad x_i^2 = 1 \quad 9 \quad \overline{x^2} = 5$$

Berechnung über Verschiebungssatz

$$S^2 = \frac{1}{n} \sum_{j=1}^n x_j^2 * \bar{x}^{-2} = \bar{x}^2 - \bar{x}^{-2} \quad \bar{x}^2 = s^2 + \bar{x}^{-2} = 0,0025 + 1,5^2 = 2,2525$$

Aufgabe 17

Gruppen	km/h von – bis	n _i	\bar{x}_i	S _i ²	S _i ² * n _i	($\bar{x}_i - \bar{x}$) ² * n _i
1	15 – 40	30	30	9	270	18750
2	40 – 60	90	50	16	1440	2250
3	60 – 75	80	70	25	2000	18000
		200			3710	39000

$$(30-55)^2 * 30$$

$$(50-55)^2 * 90$$

$$(70-55)^2 * 80$$

$$S^2 = \frac{1}{n} \sum_{i=1}^k s_i^2 * n_i + \frac{1}{n} \sum_{i=1}^k (\bar{x}_i - \bar{x})^2 * n_i$$

S²int S²ext

$$\bar{x} = \frac{1}{200} * (30 * 30 + 50 * 90 + 70 * 80) = 55$$

$$S^2 = \frac{1}{200} * 3710 + \frac{1}{200} * 39000 = 18,55 + 195 = 213,55$$

Aufgabe 18

a) n=4

$$Z = \frac{1}{2} (x_{\binom{n}{2}} + x_{\binom{n}{2+1}}) = \frac{1}{2} (x_{\binom{4}{2}} + x_{\binom{4}{2+1}}) = \frac{1}{2} (x_2 + x_3) = \frac{1}{2} (9,8 + 10) = 9,9g$$

$$\bar{x}_B = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{4} (9,7 + 8 + 10 + 10,5) = 10g$$

b)

$$\frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{10} (55,2 + 9,7 + 9,8 + 10 + 10,5) = \frac{1}{10} (55,2 + 10 * 4) = 9,52g$$

$$s^2 = \frac{1}{n} \sum_{i=1}^k s_i^2 * n_i + \frac{1}{n} \sum_{i=1}^k (\bar{x} - \tilde{x})^2 * n_i = \frac{1}{10} (s_1^2 * 6 + s_n^2 * 4) + \frac{1}{10} ((\bar{x}_1 - \bar{x})^2 * 6 + (\bar{x}_B - \bar{x})^2 * 4)$$

c)

$$s^2 = \frac{1}{n} \sum_{i=1}^k (x_i - \bar{x})^2 * n_i$$

s_n² = nicht zu berechnen, da x_i – Werte nicht gegeben

$$s_n^2 = \frac{1}{4} ((9,7 - 10)^2 + (10,5 - 10)^2 + (10 - 10)^2 + (9,8 - 10)^2)$$

$$= \frac{1}{4} * (0,09 + 0,25 + 0 + 0,04) = 0,095$$

$$\bar{x}_1 = \frac{1}{6} * 55,2 = 9,2$$

$\bar{x}_B = 10$ (siehe Aufgabenteila)

$\tilde{x} = 9,52$ (siehe Aufgabenteil b)

$$= \frac{1}{10}(s_1^2 * 6 + 0,095 * 4) + \frac{1}{10}((9,2 - 9,52)^2 * 6 + (10 - 9,52)^2 * 4)$$

$$= 0,6 * s_1^2 + 0,038 + 0,06144 + 0,09216$$

$$= 0,6 * s_1^2 + 0,1916g^2$$

Interne Varianz immer über Streuungszzerlegungssatz berechnen, wenn möglich.

$$d) Z_B = x_{\binom{n+1}{2}} \Rightarrow x_{\binom{5+1}{2}} = x_3 = 9,9g$$

$$\frac{1}{5}(9,7 + 9,8 + 10 + 10,5 + x_5) = 10$$

$$9,7 + 9,8 + 10 + 10,5 + x_5 = 50$$

$$x_5 = 10g$$

e) stetiges Merkmal und kardinal skaliert, das Histogramm ist zu benutzen
 Vorarbeit: Gruppenbildung der Gewichte – Klasseneinteilung nötig (da stetig)

Aufgabe 23

Geschlecht (x)	Votum (y)		Summe
	positiv	negativ	
männlich	24	36	60
weiblich	12	28	40
Summe	36	64	100

$$b) h(\text{männlich/negativ}) = \frac{36}{64} = 0,5625 = 56,25\%$$

$h(\text{männlich/positiv}) = h(\text{weiblich/positiv})$ - unabhängig

$$\frac{24}{36} = \frac{2}{3} \neq \frac{1}{3} = \frac{12}{36} \Rightarrow \text{Merkmale sind abhängig}$$

Aufgabe 24

Alter	selten	oft	Regelmäßig	Summe
20	5	10	5	20
20-40	5	20	25	50
40-65	5	5	5	15
ab 65	5	5	5	15
Summe	20	40	40	100

$$h(\text{regelmäßig}/\leq 20) = \frac{5}{20} = \frac{1}{4} = 25\%$$

$$h(>65/\text{selten}) = \frac{5}{20} = \frac{1}{4} = 25\%$$

b) K^* - da $0 \leq K^* \leq 1$ (da Kauf nominal ist)

Aufgabe 25

Geschlecht	60 PS	75 PS	90 PS	Summe
weiblich	24 <small>(60*120/200=36)</small>	24	72	120
männlich	36	16	28	80
Summe	60	40	100	200

$$n_{ij} = 24 \quad \bar{n}_{ij} = \frac{n_{.j} * n_{.i}}{n} = \frac{120 * 60}{200} = 36$$

$n_{ij} = 24 \neq 36 = \bar{n}_{ij} \Rightarrow$ Merkmale sind abhängig

n_{ij}	\bar{n}_{ij}	$\frac{(n_{ij} - \bar{n}_{ij})^2}{\bar{n}_{ij}}$
24	36	$(24-36)^2/36=4$
24	24	0
72	60	2,4
36	24	6
16	16	0
28	40	3,6
Summe:		16

$$\chi^2 = 16$$

$$K = \sqrt{\frac{\chi^2}{n + \chi^2}} = \sqrt{\frac{16}{200 + 16}} = 0,2722$$

$$M = \min(\text{Spalten, Zeilen}) = \min(2,3) = 2$$

$$K_{\max} = \sqrt{\frac{M-1}{M}} = \sqrt{\frac{2-1}{2}} = \sqrt{0,5}$$

$$K^* = \frac{K}{K_{\max}} = \frac{0,2722}{\sqrt{0,5}} = 0,3849$$

Aufgabe 26

	ledig	verheiratet	verwitwet	geschieden	Summe
Raucher	50	40	20	80	190
Nichtraucher	330	340	400	240	1310
Summe	380	380	420	320	1500

a)

$$n_{11} = 50 \quad \bar{n}_{11} = \frac{190 * 380}{1900} = 48,13$$

$n_{11} \neq \bar{n}_{11}$ Merkmale sind abhängig

n_{ij}	\bar{n}_{ij}	$\frac{(n_{ij} - \bar{n}_{ij})^2}{\bar{n}_{ij}}$
50	48,130	$(50-48,13)^2/48,13=0,0724$
40	48,130	1,3743
20	53,200	20,7188
80	40,530	38,4282
330	331,867	0,0105
340	331,867	0,1993
400	336,800	3,0050
240	279,467	5,5735
Summe:		69,382

$$\chi^2 = 69,382$$

$$K = \sqrt{\frac{\chi^2}{n + \chi^2}} = \sqrt{\frac{69,382}{1500 + 69,382}} = 0,2103$$

$$M = \min(k,l) = \min(2,4) = 2$$

$$K_{\max} = \sqrt{\frac{M-1}{M}} = \sqrt{\frac{2-1}{2}} = \sqrt{0,5}$$

$$K^* = \frac{K}{K_{\max}} = \frac{0,2103}{\sqrt{0,5}} = 0,2974$$

mäßiger Zusammenhang, nicht starker

c) $K_{\max} = \sqrt{0,5} = 0,7071$

Aufgabe 27

$$0 \leq \chi^2 < \infty$$

$$0 \leq K \leq K_{\max} < 1$$

$$0 \leq K^* \leq 1$$

Aufgabe 28

Regressionsanalyse: es müssen beide Merkmale kardinalskaliert sein!!!

Preis (p)	3	4	5	6	12
Absatz (y)	14	13	10	7	6

$$y' = a + bp$$

Koordinatensystem: x-Achse – Preis, y-Achse – Absatz

Zuerst das arithmetische Mittel bestimmen:

$$\bar{p} = \frac{3+4+5+6+12}{5} = 6 \quad \bar{y} = \frac{14+13+10+7+6}{5} = 10$$

p_i	y_i	$(p_i - \bar{p})$	$(y_i - \bar{y})$	$(p_i - \bar{p})(y_i - \bar{y})$	$(p_i - \bar{p})^2$	$(y_i - \bar{y})^2$
3	14	-3	4	-12	9	16
4	13	-2	3	-6	4	9
5	10	-1	0	0	1	0
6	7	0	-3	0	0	9
12	6	6	-4	-24	36	16
Summe:				-42	50	50

$$b = \frac{\sum (p_i - \bar{p})(y_i - \bar{y})}{\sum (p_i - \bar{p})^2} = \frac{-42}{50} = -0,84$$

$$a = \bar{y} - b \cdot \bar{p} = 10 - (-0,84 \cdot 6) = 10 + 0,84 \cdot 6 = 15,04$$

$$y' = 15,04 - 0,84 \cdot p$$

b) $p=2$

$$y' = 15,04 - 0,84 \cdot 2 = 13,36 \text{ (in 100 St.)}$$

Der Absatz beträgt 1336 Stück.

c)

$$r = \frac{-42}{\sqrt{50 \cdot 50}} = -0,84 \quad B=r^2 = (-0,84)^2 = 0,7056$$

Aufgabe 29

x: Seitenzahl des Originals y: Seitenzahl der Übersetzung (abhängig vom Original)

$$y' = a + bx$$

$$b = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{S_{xy}}{S_{x^2}} \quad \bar{x} = 150 \quad \bar{y} = 210 \quad S_{x^2} = 30 \quad S_{y^2} = 60 \quad S_{xy} = 42$$

$$a = \bar{y} - b\bar{x} = 210 - 1,4 \cdot 150 = 0 \quad y' = 1,4x$$

Erhöht sich die Seitenzahl des Originals um eine Seite, so steigt die Seitenzahl der Übersetzung um 1,4 Seiten.

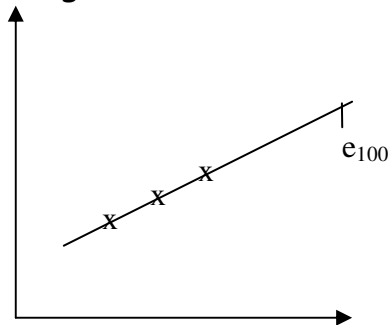
$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} = \frac{S_{xy}}{\sqrt{S_{x^2} \cdot S_{y^2}}} = \frac{42}{\sqrt{30 \cdot 60}} = 0,99 \quad B=r^2 = 0,99^2 = 0,98$$

98% der Variation von y ist durch die Veränderung von x erklärbar.

Aufgabe 30

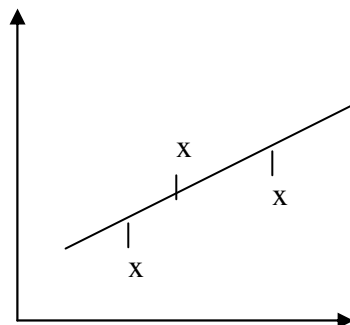
Geringe Steigung der Regressionsgeraden, wenn x um eine Einheit steigt, verändert sich y tendenziell nur wenig; keine Aussage über Stärke des Zusammenhangs zwischen x und y möglich.

Aufgabe 31



$$\sum e_i = e_{100} > 0$$

Rechenfehler da hier $\sum e_i = e_{100}$



$$\sum e_i^2 \rightarrow \min \quad \sum e_i = 0$$

Aufgabe 32

- Richtig, da bei $r=1$ vollständige Linearität
- Falsch, $-0,9$ ist vom Betrag her kleiner als eins, liegt also in dem angegebenen Bereich
- Falsch, $-0,9$ bedeutet einen gegenläufigen Zusammenhang
- Falsch, da r auf $[-1;1]$ normiert, somit liegt $-0,9$ eher nahe -1 als 0

Aufgabe 33

Sportstunden	2	5	4	3	6	8
Punkte	20	35	38	53	44	94

Sportstunden	2	3	4	5	6	8
Rangziffer	1	2	3	4	5	6

Punkte	20	35	38	44	53	94
Rangziffer	1	2	3	4	5	6

Sportstunden	1	4	3	2	5	6	Summe
Punkte	1	2	3	5	4	6	
d_i	0	2	0	-3	1	0	
d_i^2	0	4	0	9	1	0	14

$$r_{sp} = 1 - \frac{6 \sum d_i^2}{n * (n^2 - 1)} = 1 - \frac{6 \sum (R_i - R_i')^2}{(n-1) * n(n+1)} = 1 - \frac{6 * 14}{6 * (6^2 - 1)} = 1 - 0,4 = 0,6$$

Aufgabe 34

- Falsch, man kann keine Aussage über die Prozentzahl der Übereinstimmungen treffen
- Falsch, da die Berechnung auf einer ordinalen Skala beruht.
- Richtig, da der Wert in der Nähe von -1 liegt, was einer Entgegengesetzten Vorstellung entspräche.
- Falsch, da der Wertebereich auf $[-1;1]$ normiert ist.

Aufgabe 35

Varianz-Kovarianz-Matrix

	x_1	x_2	x_3
x_1	VAR_1	COV_{12}	...
x_2	COV_{12}	VAR_2	
x_3	...		VAR_3

Korrelations-Matrix

	x_1	x_2	x_3
x_1	1		
x_2		1	
x_3			1

- weder noch, da nicht quadratisch
- Varianz-Kovarianz-Matrix
keine Korrelations-Matrix da $| -3,3 | > 1$
- Varianz-Kovarianz-Matrix
keine Korrelations-Matrix, da keine 1 auf der Hauptdiagonale
- weder noch, da nicht symmetrisch
- weder noch, da negative Werte auf der Hauptdiagonale

Aufgabe 36

a)

$$\bar{x}_j = \frac{1}{n} \sum_{r=1}^n x_{rj} \quad (\text{mit } r \text{ Zeilen und } j \text{ Spalten}) \quad \bar{x} = \begin{cases} 25 \\ 55 \\ 175 \end{cases}$$

Mittelwertvektor gibt „Durchschnittsperson“ an.

b)

$$S_j^2 = \frac{1}{n} \sum (x_{rj} - \bar{x}_j)^2 \quad n=4 \quad 586 \quad 1300 \quad 500$$

$$s_j^2 \quad 146,5 \quad 325 \quad 125 \quad S_{ij} = \frac{1}{n} \sum (x_{ri} - \bar{x}_i)(x_{rj} - \bar{x}_j)$$

$x_{11} - \bar{x}_1$	$x_{12} - \bar{x}_2$	$x_{13} - \bar{x}_3$
-8	-5	15
-1	5	5
20	25	-5
-11	-25	-15

$x_{11} - \bar{x}_1$ ⁻²	$x_{12} - \bar{x}_2$ ⁻²	$x_{13} - \bar{x}_3$ ⁻²
64	25	225
1	25	25
400	625	25
121	625	225

$(x_{11} - \bar{x}_1) * (x_{12} - \bar{x}_2)$	$(x_{11} - \bar{x}_1) * (x_{13} - \bar{x}_3)$	$(x_{12} - \bar{x}_2) * (x_{13} - \bar{x}_3)$
40	-120	-75
-5	-5	25
500	-100	-125
275	165	375
810	-60	200
202,5	-15	50

$$VAR - COV = \begin{pmatrix} 146,5 & 202,5 & -15 \\ 202,5 & 325 & 50 \\ -15 & 50 & 125 \end{pmatrix}$$

$$Z_{ij} = \frac{x_{ij} - \bar{x}_j}{S_j} \quad Z_{11} = \frac{-8}{\sqrt{146,5}} = -0,6610 \quad Z_{12} = \frac{-5}{\sqrt{325}} = -0,2773$$

$$Z_{ij} = \begin{pmatrix} -0,6610 & -0,2773 & 1,3416 \\ -0,0826 & 0,2773 & 0,4472 \\ 1,6523 & 1,3867 & -0,4472 \\ -0,9088 & -1,3867 & -1,3416 \end{pmatrix}$$

$$r_{ij} = \frac{S_{ij}}{S_i * S_j} \quad r_{12} = \frac{202,5}{\sqrt{146,5 * 325}} = 0,9280$$

$$r_{13} = \frac{-15}{\sqrt{146,5 * 125}} = -0,1108$$

$$R = \begin{pmatrix} 1 & 0,9280 & -0,1108 \\ 0,9280 & 1 & 0,2481 \\ -0,1108 & 0,2481 & 1 \end{pmatrix}$$

$$r_{23} = \frac{50}{\sqrt{325 * 125}} = 0,2461$$

- Alter und Gewicht sind stark positiv korreliert
- Alter und Größe sind schwach negativ korreliert
- Gewicht und Größe sind schwach positiv korreliert

Aufgabe 37

$$r = \frac{S_{xy}}{\sqrt{S_x^2 * S_y^2}} = \frac{3}{\sqrt{4 * 9}} = 0,5 \quad x \rightarrow y \quad y \rightarrow x$$

$$b = \frac{S_{xy}}{S_x^2} \quad b' = \frac{S_{xy}}{S_y^2} \quad \hat{y} = a + bx$$

$$a = \bar{y} - b\bar{x} \quad a' = \bar{x} - b'\bar{y} \quad \hat{x} = a' + b'y$$

$$b) \quad b = \frac{3}{4} = 0,75 \quad a = -2 - 0,75 * 4 = -5$$

$$c) \quad b = \frac{3}{9} = 0,3 \quad a' = 4 - 0,3 * (-2) = 4,6$$

$$B = r^2 = 0,5^2 = 0,25 \quad B = b * b' = 0,75 * 0,3 = 0,25$$

Aufgabe 38

Jahr	98/99	99/00	00/01	01/02	02/03	03/04	04/05	Summe	n=5
Jahresüberschuss	400	500	540	570	660			2670	
t_i^*	1	2	3	4	5	6	7		
$t_i^* * y_i$	400	1000	1620	2280	3300			8600	
$(t_i^*)^2$	1	4	9	16	25			55	

$$\hat{y} = a + bt^* \quad b = \frac{n * \sum t_i^* y_i - \sum t_i^* * \sum y_i}{n * \sum t_i^2 - (\sum t_i^*)^2} \quad b = \frac{5 * 8600 - 15 * 2670}{5 * 55 - 15^2} = \frac{2950}{50} = 59$$

$$a = \bar{y} - b * \bar{t}^* \quad a = \frac{1}{n} \sum y_i - b * \frac{1}{n} \sum t_i^* = \frac{2670}{5} - 59 * \frac{15}{5} = 357$$

$$\hat{y} = 357 + 59 * t_i^* \quad \hat{y} = 357 + 59 * 7 = 770$$

Es ist mit einem Jahresüberschuß von 770 Mio € zu rechnen.

Aufgabe 39

	2000		2001		2002		2003	
y_{ij}	I	II	I	II	I	II	I	
	160	240	220	320	260	320	268	
		200	230	270	290	290	294	
G_{ij}	-	215	250	280	290	292	-	
$y_{ij} - G_{ij}$	-	25	-30	40	-30	28	-	

	I	II
2000	-	25
2001	-30	40
2002	-30	28
Summe	-60	93
\tilde{S}_j	-30	31

$$\sum \tilde{S}_j = -30 + 31 = 1$$

$$\frac{1}{2} \sum \tilde{S}_j = \frac{1}{2} * 1 = 0,5$$

$$S_1 = -30 - 0,5 = -30,5 \quad S_2 = 31 - 0,5 = 30,5$$

Aufgabe 40

Jahr	2000			2001			2002			2003		Summe
Tertial	1	2	3	1	2	3	1	2	3	1	2	
Absatz y_i	86	122	101	89	113	110	86	125	95	92	125	1144
G_{ij}	-	103	104	101	104	103	107	102	104	104	-	
$y_{ij} - G_{ij}$	-	19	-3	-12	9	7	-21	23	-9	-12	-	
t_i^*	1	2	3	4	5	6	7	8	9	10	11	66
$t_i^* * y_i$	86	244	303	356	565	660	602	1000	855	920	1375	6966
$(t_i^*)^2$	1	4	9	16	25	36	49	64	81	100	121	506

$$b) \quad b = \frac{n * \sum t_i^* y_i - \sum t_i^* * \sum y_i}{n * \sum t_i^2 - (\sum t_i^*)^2} \quad n=11 \quad b = \frac{11 * 6966 - 66 * 1144}{11 * 506 - 66^2} = \frac{1122}{1210} = 0,9273$$

$$a = \bar{y} - b \bar{t}^* = \frac{1}{n} \sum y_i - b * \frac{1}{n} \sum t_i = \frac{1}{11} * 1144 - 0,9273 * \frac{1}{11} * 66 = 98,4362$$

$$G = T + Z \quad Z = G - T$$

Zyklische Komponente kann nur für 2003/I berechnet werden, da für das zweite Tertial die glatte Komponente nicht berechnet werden konnte.

$$\hat{y} = 98,4362 + 0,9273 * t^* \quad \hat{y}_{2003/I} = 98,4362 + 0,9273 * 10 = 107,7092 \quad Z_{2003/I} = 104 - 107,7092 = -3,709$$

	I	II	III
2000		19	-3
2001	-12	9	7
2002	-21	23	-9
2003	-12		
Summe	-45	51	-5
\bar{S}_j	-15	17	-5/3

$$\sum \tilde{S}_j = -15 + 17 - \frac{5}{3} = \frac{1}{3}$$

$$\frac{1}{3} \sum \tilde{S}_j = \frac{1}{3} * \frac{1}{3} = \frac{1}{9}$$

$$S_1 = -15 - \frac{1}{9} = 15,1 \quad S_2 = 17 - \frac{1}{9} = 16,8$$

$$S_3 = -\frac{5}{3} - \frac{1}{9} = 1,7 \quad S_1 + S_2 + S_3 = 0$$

	2000		
	1	2	3
y_{ij}	86	122	101
S_j	-15,1	16,8	-1,7
$y_{ij} - S_j$	101,1	105,1	102,7

Aufgabe 41

	2000		2003	
	p_0	q_0	p_t	q_t
V	10	100	20	108
G	100	5	80	3

$$P_{0,t}^L = \frac{\sum_{i=1}^n p_t(i) * q_0(i)}{\sum_{i=1}^n p_0(i) * q_0(i)} = \frac{20 * 100 + 80 * 5}{10 * 100 + 100 * 5} = 1,6$$

Die im Basisjahr verbrauchten Mengen sind im Berichtsjahr um 60% teurer als im Basisjahr.

$$Q_{0,t}^P = \frac{\sum_{i=1}^n p_t(i) * q_t(i)}{\sum_{i=1}^n p_t(i) * q_0(i)} = \frac{20 * 108 + 80 * 3}{20 * 100 + 80 * 5} = 1$$

Die Preise des Berichtsjahres zugrunde gelegt, ist der Verbrauch im Berichtsjahr weder gestiegen noch gesunken.

Aufgabe 42

	2000		2003	
	p_0	q_0	p_t	q_t
A	1.000	120.000	2.000	75.000
B	4.000	15.000	5.000	12.000
C	8.000	5.000	10.000	5.000

a)
Umsatz: 120.000.000€
 $\frac{120.000.000}{1000} = 120.000$

$$Q_{0,t}^L = \frac{\sum_{i=1}^n p_0(i) * q_t(i)}{\sum_{i=1}^n p_0(i) * q_0(i)} = \frac{1000 * 75000 + 4000 * 12000 + 8000 * 5000}{1000 * 120000 + 4000 * 15000 + 8000 * 5000} = 0,7409$$

Der Verbrauch im Berichtsjahr, gewichtet mit den Preisen des Basisjahres ist um ca. 26% gesunken.

b)

$$P_{0,t}^P = \frac{\sum_{i=1}^n p_t(i) * q_t(i)}{\sum_{i=1}^n p_0(i) * q_t(i)} = \frac{2000 * 75000 + 5000 * 120000 + 10000 * 5000}{1000 * 75000 + 4000 * 120000 + 8000 * 5000} = 1,595$$

Waren im Basisjahr jene Mengen des Berichtsjahres verbraucht worden, so wären die Preise im Berichtsjahr um ca. 60% gestiegen.

Aufgabe 43

$$P_{2000,2003}^P = \frac{P_{2003} * q_{2003}}{P_{2000} * q_{2003}} = 1,6 = \frac{250}{P_{2000}} \Leftrightarrow P_{2000} = \frac{250}{1,6} = 156,25$$

2000 hat das Menü 156,25€ gekostet.

Aufgabe 44

- Permutation, da alle 10 Wagen berücksichtigt werden

Anzahl möglicher Kombinationen: $10! = 3.628.800$

Es gibt 3.628.800 mögliche Arten die Wagen anzuordnen. Es ist daher nicht möglich alle Anordnungen auszuprobieren.

Aufgabe 45

- Kombination, da 3 aus 5
- ohne Wiederholung, da jeder Schlüssel nur einmal benutzt wird
- mit Berücksichtigung der Anordnung, da die Schlüssel in der richtigen Reihenfolge benutzt werden müssen

$$\text{Anzahl möglicher Kombination: } \frac{n!}{(n-k)!} = \frac{5!}{(5-3)!} = 60$$

Der Einbrecher hat eine Chance von 1/60 den Safe zu öffnen.

Aufgabe 46

- Kombination, da 2 aus 10
- ohne Wiederholung, da die Vasen nicht zurückgestellt werden
- ohne Berücksichtigung der Anordnung, da die Reihenfolge der Vasenwahl keine Rolle spielt

$$\text{Anzahl möglicher Kombination: } \binom{n}{k} = \binom{10}{2} = 45$$

Es gibt 45 unterschiedliche Möglichkeiten zwei Vasen aus diesen 10 auszuwählen. Die Wahrscheinlichkeit die beiden Vasen II Wahl zu ziehen wäre also 1/45.

Aufgabe 47

- Kombination, da Ziffern (3 aus 10) und Buchstaben (2 aus 26)
- mit Wiederholung, da alle Symbole mehrfach verwendet werden können
- mit Berücksichtigung der Anordnung

$$\text{Anzahl möglicher Kombination: } n^k * n^k = 26^2 * 10^3 = 676.000$$

Es lassen sich 676.000 verschiedene Kennzeichen bilden.

Aufgabe 48

- Kombination, da 6 aus 10
- mit Wiederholung, da auch mehrere Klausuren auf dem gleichen Papier gedruckt werden können
- ohne Berücksichtigung der Anordnung, da die Reihenfolge der Farbwahl egal ist

Anzahl möglicher Kombination: $\binom{n+k-1}{k} = \binom{10+6-1}{6} = 5005$

Es gibt 5005 Möglichkeiten die Klausuren auf farbigem Papier drucken zu lassen.

Aufgabe 49

- Permutation, da alle Fächer belegt werden
- mit Wiederholung, da mehrer Fächer mit gleichen Süßigkeiten belegt werden können

Anzahl möglicher Permutation: $\frac{n!}{n_1! * n_2! * \dots * n_k!} = \frac{10!}{2! * 3! * 2! * 1! * 1! * 1!} = 151.200$

Es gibt 151.200 Möglichkeiten, den Automaten zu bestücken.

Aufgabe 50

	Modus	Median	arithmetisches Mittel	Varianz
a)	Mercedes	–	–	–
b)	Klasse B	Klasse B	–	–
c)	0,75	6,6	6,425	1,3756

€ pro m ²	n _i	h _i	Δx _i	f _i = $\frac{h_i}{\Delta x_i}$	F _i	x _i [*]	x _i [*] * n _i	(x _i [*] - \bar{x}) * n _i
3,50 – 5,50	40	0,2	2	0,1	0,2	4,5	180	148,2250
5,50 – 6,50	40	0,2	1	0,2	0,4	6	240	7,2250
6,50 – 7,00	60	0,3	0,5	0,6	0,7	6,75	405	6,3375
7,00 – 7,50	40	0,2	0,5	0,4	0,9	7,25	290	27,2250
7,50 – 9,50	20	0,1	2	0,05	1	8,5	170	86,1125
	200						1285	275,1250

Median:

$$Z = x_{\frac{n+1}{2}} = x_{\frac{31+1}{2}} = x_{16}$$

Z liegt bei F(x) = 0,5

n _i	F _i
6,5	0,4
Z	0,5
7	0,7

$$\frac{dx}{Dx} = \frac{dF}{DF} \quad \frac{Z-6,5}{0,5} = \frac{0,1}{0,3} \quad Z-6,5=0,16 \quad Z=6,6$$

$$dx = Z - 6,5 \quad Dx = 7 - 6,5 = 0,5 \quad dF = 0,5 - 0,4 = 0,1 \quad DF = 0,7 - 0,4 = 0,3$$

arithmetisches Mittel:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k x_i^* * n_i \Rightarrow \frac{1}{200} * 1285 = 6,425 \quad S^2 = \frac{1}{n} \sum_{i=1}^k (x_i^* - \bar{x}) * n_i = \frac{1}{200} * 275,1250 = 1,3756$$

Graphische Darstellung				
	absolute Häufigkeit	relative Häufigkeit	Verteilungsfunktion	Dichtefunktion
a)	Kreisdiagramm	Kreisdiagramm	–	–
b)	Stab-, Säulendiagramm	Stab-, Säulendiagramm	Treppenfunktion	–
c)	Histogramm	Histogramm	Polygonzug	Häufigkeitspolygon

Aufgabe 51

n_{ij}	unter 15 min	15-30 min	mehr als 30 min	Σ
Männer	30	50	20	100
Frauen	10	40	50	100
Σ	40	90	70	200

n_{ij}	unter 15 min	15-30 min	mehr als 30 min	Σ
Männer	20	45	35	100
Frauen	20	45	35	100
Σ	40	90	70	200

$$\frac{40 \cdot 100}{200} = 20$$

$$\chi^2 = \frac{(30-20)^2}{20} + \frac{(50-45)^2}{45} + \frac{(20-35)^2}{35} + \frac{(10-20)^2}{20} + \frac{(40-45)^2}{45} + \frac{(50-35)^2}{35}$$

$$= 5 + 0,5 + 6,4 + 5 + 0,5 + 6,4 = 23,9684$$

$$K = \sqrt{\frac{23,9684}{200 + 23,9684}} = 0,3271 \quad M = \min(3,2) = 2 \quad K_{\max} = \sqrt{\frac{2-1}{2}} = \sqrt{0,5} = 0,7071 \quad K^* = \frac{0,3271}{0,7071} = 0,4626$$

Aufgabe 52

	n_i	\bar{x}_i	S_i^2	$\bar{x}_i - n_i$	$S_i^2 * n_i$	$(\bar{x}_i - \bar{x})^2 * n_i$
1	180	48,2	36	8676	6480	50,8869
2	270	46,5	22	12555	5940	368,5297
3	215	47,1	48	10126,5	10320	69,4375
4	248	49,1	29	12176,8	7192	508,3417
5	193	47,6	41	9186,8	7913	0,9003
Σ	1106			52721,1	37845	998,0961

$$\bar{x} = \frac{1}{n} \sum \bar{x}_i * n_i \Rightarrow \frac{1}{1106} * 52721,1 = 47,6683$$

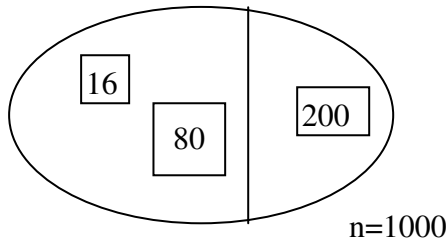
$$S^2 = \frac{1}{n} \sum S_i^2 * n_i + \frac{1}{n} \sum (\bar{x}_i - \bar{x})^2 * n_i = \frac{1}{1106} * 37845 + \frac{1}{1106} * 998,0961 = 35,1203$$

Aufgabe 53

Merkmale	Kontingenzkoeffizient	Rangkorrelationskoeffizient nach Spearman	Korrelationskoeffizient nach Bravais-Pearson
Geschlecht-Alter	X		
Alter-Körpergröße	X	X	X
Steuerklasse-Einkommen	X	X	
Automarke-Geschlecht	X		
IQ-Schulabschluss	X	X	

$$b) W(A_1) = \frac{4}{52} \quad W(A_2/A_1) = \frac{3}{51} \quad W(A_1 \cap A_2) = \frac{4}{52} * \frac{3}{51} = \frac{1}{221} = 0,00452$$

Aufgabe 59



a) Insgesamt haben 1000 Studenten teilgenommen, und von diesen 1000 sind 200 durchgefallen, so dass 800 bei der Prüfung erfolgreich waren
 A = erfolgreicher Student

$$W(A) = \frac{800}{1000} = 0,8$$

b) B = Note: sehr gut

$$W(B/A) = \frac{W(A \cap B)}{W(A)} = \frac{16}{0,8}$$

$$W(B/A) = \frac{16}{800} = 0,02$$

x_i	n_i
sehr gut	16
gut	80
nicht bestanden	200

Insgesamt hatten 800 Studenten die Prüfung erfolgreich abgelegt. Von diesen 800 erzielten 16 die Note sehr gut

Aufgabe 60

A = infiziertes Rind

B = Test ist positiv

\bar{A} = nicht infizierte Rind $W(A/B)$

\bar{B} = Test ist negativ

$$W(A) = 0,001$$

$$W(\bar{A}) = 0,999$$

$$W(B/A) = \frac{999}{1000} = 0,999$$

$$W(B/\bar{A}) = 1\% = 0,01$$

$$W(A/B) = \frac{W(B/A) * W(A)}{W(B/A) * W(A) + W(B/\bar{A}) * W(\bar{A})} = \frac{0,999 * 0,001}{0,999 * 0,001 + 0,01 * 0,999}$$

$$= \frac{0,000999}{0,000999 + 0,010989} = 0,0909$$

Aufgabe 61

A_1 = Hauptschüler

$$W(A_1) = 0,4 \quad W(B/A_1) = 0,1$$

A_2 = Abiturient

$$W(A_2) = 0,5 \quad W(B/A_2) = 0,05$$

A_3 = Hochschulabsolvent

$$W(A_3) = 0,1 \quad W(B/A_3) = 0,02$$

B = arbeitslos

$$W(A_3/B)$$

\bar{B} = nicht arbeitslos

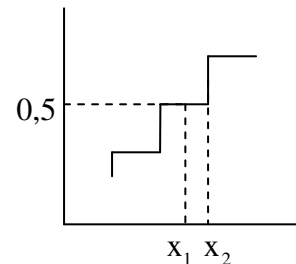
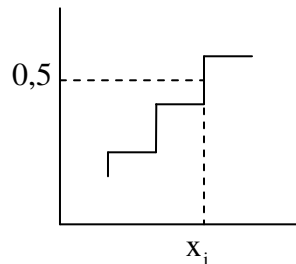
$$W(A_3/B) = \frac{W(B/A_3) * W(A_3)}{W(B/A_3) * W(A_3) + W(B/A_1) * W(A_1) + W(B/A_2) * W(A_2)}$$

$$= \frac{0,02 * 0,1}{0,02 * 0,1 + 0,1 * 0,4 + 0,05 * 0,5} = \frac{0,002}{0,067} = 0,0299$$

Aufgabe 62

$$\Omega = \begin{pmatrix} 12 & 12 & 13 & 14 & 15 & 16 \\ 21 & 22 & 23 & 24 & 25 & 26 \\ 31 & 32 & 33 & 34 & 35 & 36 \\ 41 & 42 & 43 & 44 & 45 & 46 \\ 51 & 52 & 53 & 54 & 55 & 56 \\ 61 & 62 & 63 & 64 & 65 & 66 \end{pmatrix} \quad x = \begin{pmatrix} 0 & -1 & -2 & -3 & 9 & 5 \\ 1 & 0 & -1 & -2 & -3 & -4 \\ 2 & 1 & 0 & -1 & -2 & -3 \\ 3 & 2 & 1 & 0 & -1 & -2 \\ 4 & 3 & 2 & 1 & 0 & -1 \\ 5 & 4 & 3 & 2 & 1 & 0 \end{pmatrix} \quad x = \text{Differenz der Augenzahlen}$$

x_i	$f(x_i)$	$F_i(x)$
-5	1/36	1/36
-4	2/36	3/36
-3	3/36	6/36
-2	4/36	10/36
-1	5/36	15/36
0	6/36	21/36
1	5/36	26/36
2	4/36	30/36
3	3/36	33/36
4	2/36	35/36
5	1/36	36/36



$$\sum(x) = \sum x_i * f(x_i) = 0$$

Aufgabe 63

x_i	$f(x_i)$	$F_i(x)$
1	1/6	1/6
4	1/6	2/6
9	1/6	3/6
16	1/6	4/6
25	1/6	5/6
36	1/6	6/6

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

$X = \text{Quadrat der Augenzahl}$

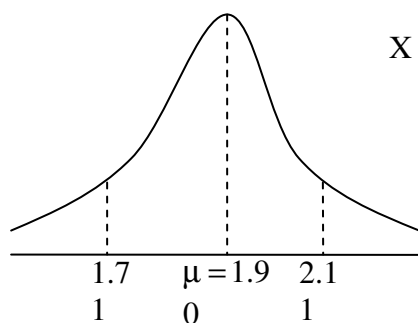
$$X = \{1, 4, 9, 16, 25, 36\}$$

$$D = 1, 4, 9, 16, 25, 36$$

$$Z = 12,5$$

$$\sum(x) = \sum x_i * f(x_i) = 1 * \frac{1}{6} + \frac{4}{6} + \frac{9}{6} + \frac{16}{6} + \frac{25}{6} + \frac{36}{6} = \frac{91}{6} = 15,1667$$

Aufgabe 64



$$X \sim NV(\mu, \sigma) \quad T \sim (0,1)$$

$$X \sim NV(1.9, 0.2)$$

$$W(1.7 \leq X \leq 2.1)$$

$$W = \left(\frac{1.7 - \mu}{\sigma} \leq T \leq \frac{2.1 - \mu}{\sigma} \right) = W = \left(\frac{1.7 - 1.9}{0.2} \leq T \leq \frac{2.1 - 1.9}{0.2} \right) =$$

$$W = (-1 \leq T \leq 1) = \Phi(1) = 68,27\%$$

b)

$$W(X > 2.2) = W\left(T > \frac{2.2 - 1.9}{0.2}\right) = W(T > 1.5) = 0.5 - 0.5\Phi(1.5)$$

$$= 0.5 - 0.5 * 0.86639 = 0.066805$$

⇒ 6,68% der Enten wiegt mehr als 2,2 kg

c) $W(X=1.9)=0$

Aufgabe 65

$X \sim NV(50; 2.4)$

$$W(50.96 \leq X \leq 51.92) = W\left(\frac{50.96 - 50}{2.4} \leq T \leq \frac{51.92 - 50}{2.4}\right) = W(0.4 \leq T \leq 0.8)$$

$$0.5 * \phi(0.8) - 0.5 * \phi(0.4) = 0.5 * 0.57629 - 0.5 * 0.31084 = 0.132725$$

\Rightarrow 13,27% der Brötchen wiegen zwischen 50,96 und 51,92g.

Aufgabe 66

- richtig
- richtig
- richtig
- richtig
- falsch, Wahrscheinlichkeit einer Variablen ist gleich Null
- richtig, Dichte generell definiert $f(x) \geq 0$ im Fall der Normalverteilung gilt $f(x_i) > 0$