Grundbegriffe

- **Aufgabe:** Vervollständigen Sie folgende Tabelle:

Statistische Einheit	Merkmal (häufbar?)	Merkmalsausprägungen
		BWL / Informatik / Elektrotechnik
		m / w / d
	Alter	
Arbeitnehmer		$\mathbb{R}^+$ (oder $\mathbb{N}$)
	Bildungsniveau	
	Arbeitszeit	
IT-Unternehmen		$\mathbb{N}$
		$\mathbb{R}^+$
		$\mathbb{R}$
Regionen	Arbeitslosenquote	
	Wirtschaftskraft	
Großstadt	Bevölkerungsdichte	
	Politische Funktion	
		$\mathbb{R}^+$ (oder $\mathbb{N}$)
	Verschuldung (in % des BIP)	

Grundbegriffe

- Lösung: Vervollständigen Sie folgende Tabelle:

Statistische Einheit	Merkmal (häufbar?)	Merkmalsausprägungen
Student	Studienfach (h) Geschlecht Alter	BWL / Informatik / Elektrotechnik m / w / d $\mathbb{R}^+$ (oder $\mathbb{N}$)
Arbeitnehmer	Einkommen Bildungsniveau Arbeitszeit	$\mathbb{R}^+$ (oder $\mathbb{N}$) Abitur / Bachelor / Master $\mathbb{R}^+$
IT-Unternehmen	Mitarbeiterzahl Umsatz Gewinn/Verlust	$\mathbb{N}$ $\mathbb{R}^+$ $\mathbb{R}$
Regionen	Arbeitslosenquote Wirtschaftskraft	$[0; 1]$ $\mathbb{R}^+$
Großstadt	Bevölkerungsdichte Politische Funktion (h)	$\mathbb{Q}$ Mittelzentrum / (Landes)hauptstadt
Staaten	BIP Verschuldung (in % des BIP)	$\mathbb{R}^+$ (oder $\mathbb{N}$) $[0; 100]$

Grundbegriffe

Aufgabe: Geben Sie für folgende Fragestellungen die statistische Einheit, Merkmale, Merkmalsausprägungen sowie die Identifikationskriterien an:

a) Es soll die Qualität von aus der Bodenseeäpfeln in einem Karlsruher Supermarkt untersucht werden.

b) Es soll die Todesursache von Corona-Patienten in Deutschland ermittelt werden.

- Lösung:

	Statistische Einheit	Merkmal	Merkmalsausprägungen	Identifikationskriterien
a)	Bodenseeapfel	Qualität	sehr gut, gut, mittel, schlecht	sachlich: Äpfel vom Bodensee räumlich: Karlsruher Supermarkt zeitlich: Prüftag
b)	Gestorbener Coronapatient	Todesursache	Corona, Lungenkrebs, Unfall, …	sachlich: gestorbener Coronapatient räumlich: Deutschland zeitlich: 2021

Grundbegriffe

Aufgabe: Geben Sie zu folgenden Merkmalen die richtige Skala an und entscheiden Sie über die Häufbarkeit.

Lösung:

	Merkmal	Skala	häufbar?
a)	Wertungsnoten beim Turmspringen		
b)	Matr.-Nr.		
c)	Kinderzahl		
d)	Konto-Nr.		
e)	Restaurantsterne		

Aufgabe: Entscheiden Sie, ob folgende Skalentransformationen für das Merkmal „Familienstand" zulässig sind.

Lösung:

	Merkmal	zulässig
a)	Ledig $\rightarrow A$; verheiratet $\rightarrow B$; Zusammenlebend $\rightarrow C$; Verwitwet $\rightarrow D$ Geschieden $\rightarrow E$	
d)	Ledig, verwitwet, geschieden $\rightarrow$ *alleine lebend*; verheiratet, zusammenlebend $\rightarrow$ *zusammenlebend*	
e)	Ledig $\rightarrow 5$; verheiratet $\rightarrow 3$; Zusammenlebend $\rightarrow 2$; Verwitwet $\rightarrow 4$ Geschieden $\rightarrow 1$	

Grundbegriffe

Aufgabe: Geben Sie zu folgenden Merkmalen die richtige Skala an und entscheiden Sie über die Häufbarkeit.

Lösung:

	Merkmal	Skala	häufbar?
a)	Wertungsnoten beim Turmspringen	Ordinalskala	✓
b)	Matr.-Nr.	Nominalskala	X
c)	Kinderzahl	Absolutskala	X
d)	Konto-Nr.	Nominalskala	✓
e)	Restaurantsterne	Ordinalskala	X

Aufgabe: Entscheiden Sie, ob folgende Skalentransformationen für das Merkmal „Familienstand“ zulässig sind.

Lösung:

	Merkmal	zulässig
a)	Ledig $\rightarrow A$; verheiratet $\rightarrow B$; Zusammenlebend $\rightarrow C$; Verwitwet $\rightarrow D$ Geschieden $\rightarrow E$	✓
d)	Ledig, verwitwet, geschieden $\rightarrow$ *alleine lebend*; verheiratet, zusammenlebend $\rightarrow$ *zusammenlebend*	X
e)	Ledig $\rightarrow 5$; verheiratet $\rightarrow 3$; Zusammenlebend $\rightarrow 2$; Verwitwet $\rightarrow 4$ Geschieden $\rightarrow 1$	✓

Histogramm & Summenhäufigkeiten

- Aufgabe:

20 Teilnehmer einer Klausur haben folgende Punktzahlen erreicht:

$$8, 74, 87, 43, 60, 72, 56, 36, 75, 49, 83, 52, 71, 67, 78, 50, 76, 64, 77, 69$$

a) Stellen Sie die Häufigkeitsverteilung der Klausurpunktzahlen als Histogramm dar.

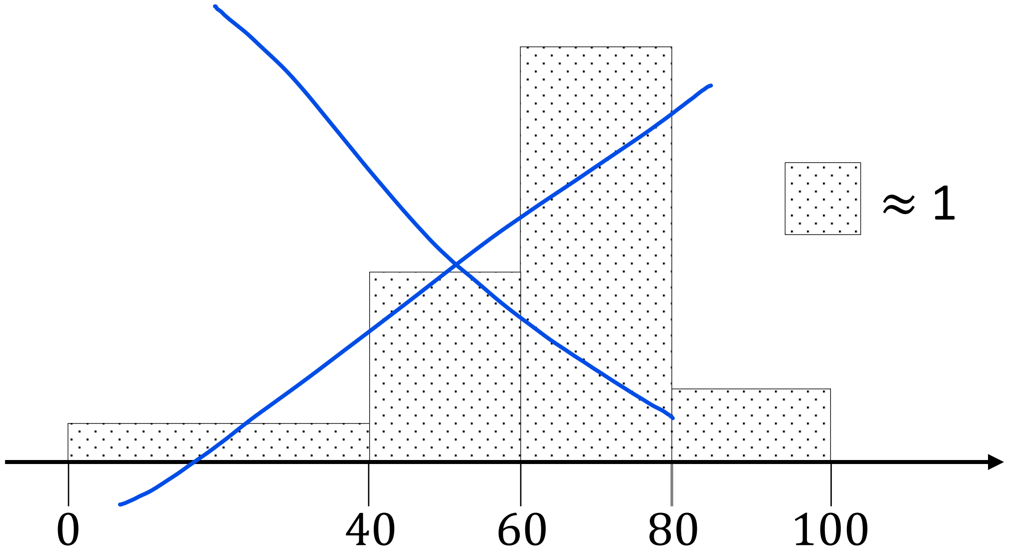
b) Stellen Sie die Summen- & Resthäufigkeiten der Klassen grafisch dar und interpretieren Sie diese.

Histogramm & Summenhäufigkeiten

- <u>Lösung:</u>

a) Histogramm: $\cancel{Häufigkeit = Höhe * Breite}$ $\Rightarrow Höhe = \frac{Häufigkeit}{Breite}$

Klasse	Häufigkeit	*Breite*	*Höhe*
$0-39$	2	4	$0{,}5$
$40-59$	5	2	$2{,}5$
$60-79$	11	2	$5{,}5$
$80-100$	2	$2{,}1$	$\approx 0{,}95$

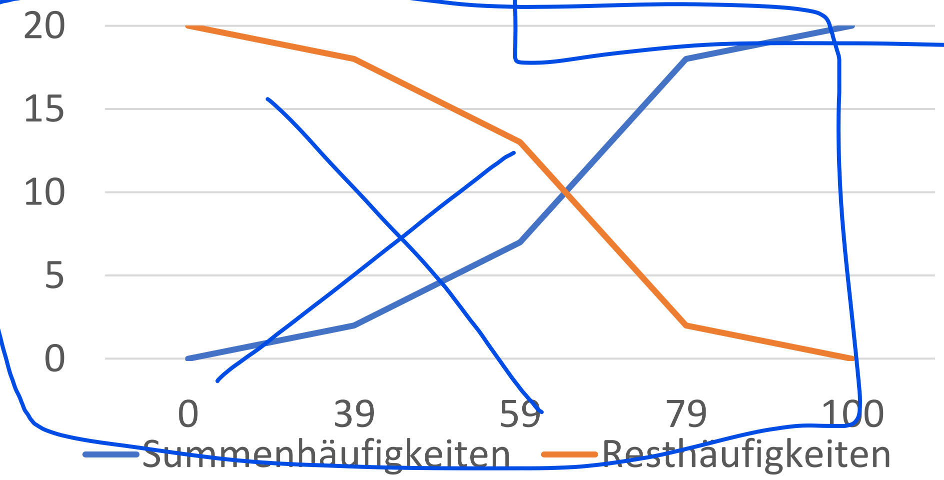


Histogramm & Summenhäufigkeiten

- <u>Lösung:</u>

b) Summen- & Resthäufigkeiten:

Klasse	Summenhäufigkeit	*Resthäufigkeit*	*Interpretation*
	0	20	
$0 - 39$	2	18	2 Studenten haben unter 40 Pkte.; 18 haben 40 Pkte. oder mehr
$40 - 59$	7	13	7 Studenten haben unter 60 Pkte.; 13 haben 60 Pkte. oder mehr
$60 - 79$	18	2	18 Studenten haben unter 80 Pkte.; 2 haben 80 Pkte. oder mehr
$80 - 100$	20	0	20 Studenten haben bis zu 100 Pkte.; keiner hat mehr



Mittelwerte

- Aufgabe:

  Bestimmen Sie zu folgenden Fragestellungen jeweils einen geeigneten Durchschnittswert:

a) Von 12 getesteten Hotels erhielten 6 Hotels 3 Sterne, 2 Hotels 2 Sterne, 3 Hotels 1 Stern und ein Hotel erhielt keinen Stern. Wie viele Sterne wurden im Durchschnitt vergeben?

b) Ein Auto fährt eine $1{,}5$ Stunden lang mit $100\ km/h$. Danach fährt es eine Stunde lang mit $80\ km/h$. Wie hoch ist die Durchschnittsgeschwindigkeit?

c) Ein Auto fährt $30\ km$ mit $80\ km/h$, $50\ km$ mit $100\ km/h$ und $20\ km$ mit $120\ km/h$. Wie hoch ist die Durchschnittsgeschwindigkeit (für die gleiche Strecke in der gleichen Zeit)?

d) Die Lohnsteigerungen in einer Firma betrugen in den letzten drei Jahren $6{,}1\%$, $4{,}2\%$, und $3{,}8\%$. Wie hoch war die durchschnittliche Lohnsteigerungsrate?

Mittelwerte

- <u>Lösung:</u>

a) Median:

	1	2	3	4	5	6	7	8	9	10	11	12
Vergebene Sterne:	0	1	1	1	2	2	3	3	3	3	3	3

$$\tilde{x} = \frac{2+3}{2} = 2{,}5$$

b) Arithmetischer Mittelwert: $\overline{v} = \left(1{,}5h * 100\,\frac{km}{h} + 1h * 80\,\frac{km}{h}\right) : 2{,}5h = 92\,\frac{km}{h}$

c) Harmonischer Mittelwert: $\overline{v}_H = \frac{s_{ges.}}{t_{ges.}} = \frac{s_1+s_2+s_3}{t_1+t_2+t_3} = \frac{s_1+s_2+s_3}{\frac{s_1}{v_1}+\frac{s_2}{v_2}+\frac{s_2}{v_2}} = 96\ km/h$

d) Geometrischer Mittelwert: $\overline{L}_H = \sqrt[3]{1{,}061 * 1{,}042 * 1{,}038} = 1{,}04695$. Das entspricht einer durchschnittlichen Lohnsteigerung von $4{,}695\%$.

Varianz

- <u>Aufgabe:</u>

Wahr oder falsch?

Was ist die Bedeutung der Varianz?

a) Ein Maß für die Unterschiedlichkeit von Kennzahlen in Bezug auf den arithm. Mittelwert. $\boxed{f}$

b) Ein Maß für die Unterschiedlichkeit von Kennzahlen in Bezug auf die Standardabweichung. $\boxed{f}$

c) Ein Maß für die Abweichung der Werte einer Verteilung vom arithmetischen Mittelwert. $\boxed{\checkmark}$

d) Ein Maß für die Differenz zwischen Maximum und Minimum einer Verteilung. $\boxed{f}$

e) Eine durch den Wendepunkt einer Verteilung bestimmte statistische Kennzahl. $\boxed{f}$

f) Ein Maß für die Abweichung der Werte einer Verteilung vom geometrischen Mittelwert. $\boxed{\checkmark}$

Varianz

- Aufgabe:

Zeigen Sie, dass gilt:

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \overline{x^2} - \overline{x}^2$$

Varianz

- <u>Aufgabe:</u>

Zeigen Sie, dass gilt:

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2$$

- <u>Lösung:</u>

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i{}^2 - 2x_i\overline{x} + \overline{x}^2\right) = \frac{1}{n}\sum_{i=1}^{n} x_i{}^2 - 2\overline{x}\frac{1}{n}\sum_{i=1}^{n} x_i + \frac{1}{n}\sum_{i=1}^{n}\overline{x}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i{}^2 - 2\overline{x} * \overline{x} + \frac{1}{n} n\overline{x}^2 =$$

$$= \frac{1}{n}\sum_{i=1}^{n} x_i{}^2 - 2\overline{x}^2 + \overline{x}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2 = \overline{x^2} - \overline{x}^2$$

Box-Plot

- Aufgabe:

Von den Professoren einer Universität wurde das jeweilige Alter erfasst:

$40; 54; 42; 60; 44; 52; 62; 52; 24; 58; 45; 54; 46; 48; 50; 66; 39; 45; 61; 51; 55; 53;$

a) Stellen Sie die Altersverteilung grafisch in einem Box-Plot dar.

b) Ermitteln Sie außerdem $x_{mod}, \overline{x}, R$ sowie die Varianz und die Standardabweichung?

Box-Plot

- Lösung: Geordnete Altersverteilung:

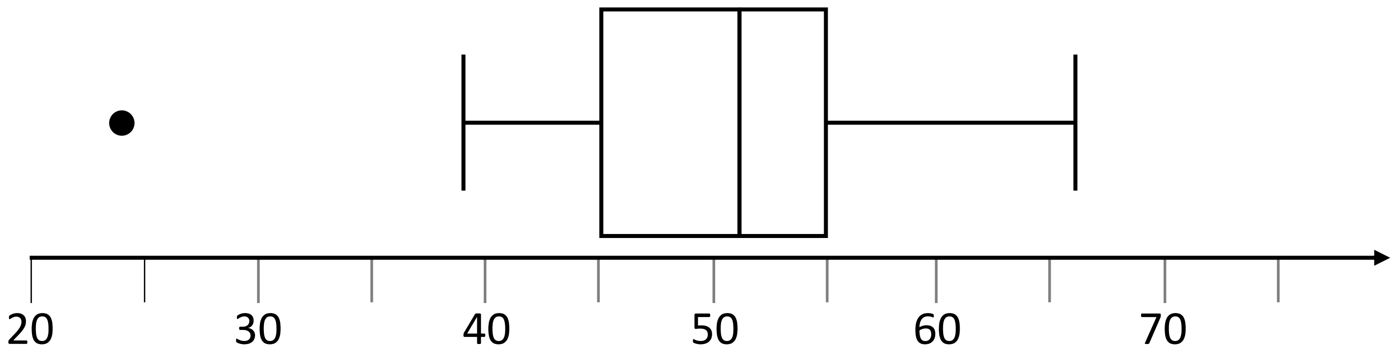
24;	39;	40;	42;	44;	45;	45;	46;	48;	50;	51;	52;	52;	53;	54;	54;	55;	58;	60;	61;	62;	66;
1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.

$\overline{x} = \frac{1101}{22} = 50{,}05$
$x_{mod} = 45; 52; 54;$
Q_1: $22 * \frac{1}{4} = 5{,}5 \; (nicht\ ganzzzahlig) \Rightarrow Q_1 = 6.Dw = 45$
Q_2: $22 * \frac{1}{2} = 11{,}0 \; (ganzzzahlig) \Rightarrow Q_2 = \frac{11.Dw. + 12.Dw}{2} = \frac{51+52}{2} = 51{,}5$
Q_3: $22 * \frac{3}{4} = 16{,}5 \; (nicht\ ganzzzahlig) \Rightarrow Q_3 = 17.Dw = 55$
$IQA = Q_3 - Q_1 = 55 - 45 = 10$
$max.untere\ Whiskergr. = Q_1 - 1{,}5 * \mathrm{IQA} = 45 - 15 = 30;$ $max.obere\ Whiskergr.. = Q_3 + 1{,}5 * \mathrm{IQA} = 55 + 15 = 70$
$\min = 24; \quad max = 66; \quad R = max - min = 66 - 24 = 42$
$tatsächliche\ obere\ Whiskergr.: 66$
$tatsächliche\ untere\ Whiskergr.: 39$

Box-Plot

- Lösung:

$\overline{x} = \frac{1101}{22} = 50{,}05$
$x_{mod} = 45; 52; 54;$
$Q_1 = 45$
$Q_2 = 51{,}5$
$Q_3 = 55$
$IQA = 10$
$\min = 24; \quad max = 66;$
$tatsächliche\ obere\ Whiskergr.: 66$
$tatsächliche\ untere\ Whiskergr.: 39$



Varianz:

$$s^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2 = \frac{1}{22} * 56.927 - 2504{,}55 = 83{,}04; \quad \Rightarrow s = \sqrt{83{,}04} = 9{,}11$$

Unabhängigkeit von Merkmalen

- Aufgabe:

In folgender Häufigkeitstabelle wurden die häufigsten Ausfallgründe von CPU‘s von drei verschiedenen Herstellern erfasst.

	A	B	C
$defektes\ Mainboard$	30	5	15
$Hitzeschaden$	15	10	5
$zu\ schwaches\ bzw.kaputtes\ Netzteil$	15	5	0

Ist der Ausfallgrund unabhängig vom Hersteller?

Unabhängigkeit von Merkmalen

- Lösung:

In folgender Häufigkeitstabelle wurden die häufigsten Ausfallgründe von CPU‘s von drei verschiedenen Herstellern erfasst.

Bedingte Verteilung des Herstellers			
	A	B	C
$defektes\ Mainboard$	0,6	0,1	0,3
$Hitzeschaden$	0,5	0,33	0,16
$zu\ schwaches\ bzw.kaputtes\ Netzteil$	0,75	0,25	0

$$\Rightarrow abhängig$$

Bedingte Häufigkeiten

- Aufgabe:

Für die zweidimensionale Verteilung der relativen Häufigkeiten von 200 Beobachtungswerten ist Folgendes bekannt: $f(x_1|y_3) = 0{,}25$ & $h(y_1) = 60$. Vervollständigen Sie die zugehörige Häufigkeitstabelle:

	y_1	y_2	y_3	
x_1	$0{,}2$			
x_2		$0{,}1$		
			$0{,}4$	

Bedingte Häufigkeiten

- Lösung:

$$f(x_1|y_3) = 0{,}25 \quad \Rightarrow f(x_1; y_3) = 0{,}4 * 0{,}25 = 0{,}1$$

$$h(y_1) = 60 \quad \Rightarrow f(y_1) = \frac{60}{200} = \mathbf{0{,}3}$$

Das markieren, was gegeben war

	y_1	y_2	y_3	
x_1	0,2	0,2	0,1	0,5
x_2	0,1	0,1	0,3	0,5
	0,3	0,3	0,4	1

Lineare Regression

- Aufgabe:

Gegeben sind folgende Beobachtungswertepaare $(x_i; y_i)$: $(-1; 1{,}5), (0; 1), (0; 3), (1; 2), (1; 3), (1{,}5; 2{,}75)$.

(Hinweis: Geben Sie exakte Ergebniswerte an!)

a) Bestimmen Sie die lineare yx – KQ-Regressionsfunktion.

b) Bestimmen Sie den Schnittpunkt der linearen yx-Regressionsfunktion (aus Teilaufgabe a)) mit der linearen xy-Regressionsgeraden.

c) Bestimmen Sie für beide Regressionsfunktionen die Summe der Residuen.

d) Berechnen Sie den linearen KQ-Regressionskoeffizienten a ohne Verwendung der Formel.

e) Zeigen Sie: $\sum_{i=1}^{n} y_i = na + b\sum_{i=1}^{n} x_i \quad \Leftrightarrow \quad \bar{y} = a + b\bar{x}$

Lineare Regression

- Lösung:

x_i	y_i	${x_i}^2$	$x_i{*}y_i$	$(x_i-\overline{x})^2$
-1	1,5	1	$-1{,}5$	289/144
0	1	0	0	25/144
0	3	0	0	25/144
1	2	1	2	49/144
1	3	1	3	49/144
1,5	2,75	2,25	4,125	169/144
2,5	13,25	5,25	7,625	101/24

$$\overline{x} = \frac{5}{12}; \qquad \overline{y} = \frac{53}{24}$$

$$s_x^2 = \frac{1}{6}\left(\frac{289}{144} + \frac{25}{144} + \frac{25}{144} + \frac{49}{144} + \frac{49}{144} + \frac{169}{144}\right) = \frac{1}{6} * \frac{606}{144} = \frac{101}{144}$$

$$COV(X,Y) = \frac{-1{,}5 + 2 + 3 + 4{,}125}{6} - \frac{5 * 53}{12 * 24} = \frac{101}{288}$$

$$a = \frac{\frac{53}{24} * 5{,}25 - \frac{5}{12} * 7{,}625}{6 * \frac{101}{144}} = \frac{\frac{101}{12}}{\frac{101}{24}} = 2 \qquad b = \frac{\frac{101}{288}}{\frac{101}{144}} = \frac{1}{2}$$

$a)\ \hat{y} = \frac{1}{2}x + 2 \quad b)\ S(\overline{x}|\overline{y}) = S\left(\frac{5}{12}\Big|\frac{53}{24}\right) \quad c)\ \sum u_i = 0 \quad d)\ \overline{y} = a + b\overline{x};\ \Leftrightarrow \frac{53}{24} = a + \frac{1}{2}\frac{5}{12};\ \Rightarrow a = 2$

Lineare Regression

- Lösung:

$$e)\ \sum_{i=1}^{n} y_i = na + b\sum_{i=1}^{n} x_i \quad \Leftrightarrow \quad \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}na + \frac{1}{n}b\sum_{i=1}^{n} x_i \quad \Leftrightarrow \quad \bar{y} = a + b\bar{x}$$

Lineare Regression

- Aufgabe:

Von Beobachtungswertpaaren $(x_i; y_i)$ sind die jeweiligen arithmetischen Mittelwerte der Merkmalsausprägungen bekannt: $\bar{x} = 5;\quad \& \quad \bar{y} = 10$. Des weiteren weiß man, dass sich das Merkmal Y um durchschnittlich 18 Einheiten verringert, wenn sich das Merkmal X um 6 Einheiten erhöht.

Wie lautet die lineare yx-Regressionsfunktion?

- Lösung:

$$\left.\begin{aligned} \bar{y} &= a + b\bar{x} \\ b &= -\frac{18}{6} = -3 \end{aligned}\right\} \Rightarrow \quad 10 = a - 3 * 5 \quad \Rightarrow \quad \mathrm{a} = 25 \qquad \Rightarrow \quad \hat{y} = 25 - 3x$$

Lineare Regression

- Aufgabe:

Welche Aussagen über eine KQ-Regressionsfunktion ($\hat{y} = a + bx$) sind wahr bzw. falsch?

a) Sie gibt an, welchen Wert Y durchschnittlich für ein best. Wert von X annimmt. ✓

b) Die Regressionsfunktion gibt den eindeutigen Zusammenhang zwischen X und Y an. f

c) Wenn $a \approx 0$, dann ist der lineare Zusammenhang nur sehr schwach ausgeprägt. f

d) Eine lineare KQ-Regressionsfunktion kann unterschiedl. lineare Zusammenhänge darstellen. ✓

Zusammenhangsanalyse

- Aufgabe:

Tageszeit t [in Std.]	8	10	12	14	16	18
Intensität I [Watt/m²]	101	345	560	578	388	128

Die Journalistin Karla Kolumna misst die Strahlungsintensität der Sonne in Abhängigkeit von der Tageszeit. Aus den Messdaten (siehe Tabelle) berechnet sie den Pearson'schen Korrelationskoeffizienten $r = 0{,}06$. Daraufhin schreibt sie in einem Artikel:

„SENSATION: Strahlungsintensität der Sonne hängt praktisch nicht von der Tageszeit ab! Konservative Physiker völlig ratlos!"

a) Nehmen Sie zu Karla Kolumnas' Aussage Stellung.

b) Wie gut lässt sich der Zusammenhang zwischen Strahlungsintensität und Tageszeit durch die Regressionsfunktion: $\hat{y} = -18{,}482t^2 + 484{,}56\mathrm{t} - 2610{,}2 \quad (\mathrm{t\ in\ Std.})$ beschreiben?

(Hinweis: $\bar{y} = 350;\ s_y^2 = 34.806,\overline{3}$)

Zusammenhangsanalyse

- Lösung:

a) Der Pearson‘sche Korrelationskoeffizient eignet sich nur zur Untersuchung der Ausgeprägtheit linearer Zusammenhänge; für nichtlineare (ist hier der Fall) Zusammenhänge ist er nicht geeignet.

b) $\hat{y} = -18{,}482t^2 + 484{,}56t - 2610{,}2$ (t in Std.)

Tageszeit t [in Std.]	8	10	12	14	16	18		
Intensität I bzw. y_i	101	345	560	578	388	128		$\bar{y} = 350$
$(y_i - \bar{y})^2$	62.001	25	44.100	51.984	1.444	49.284		$s_y^2 = 34.806,\bar{3}$
$\hat{y}_i = \hat{y}(t)$	83,432	387,2	543,112	551,168	411,368	123,712		
$(\hat{y}_i - \bar{y})^2$	71.058,499	1.383,84	37.292,245	40.468,564	3.766,031	51.206,259		$s_{\hat{y}}^2 = 34.195,9$

$\hat{y}(8) = -18{,}482 * 8^2 + 484{,}56 * 8 - 2610{,}2 = 83{,}432$

$$B^2 = \frac{s_{\hat{y}}^2}{s_y^2} = \frac{34.195,9}{34.806,\bar{3}} = 0{,}982$$

Zusammenhangsanalyse

- Aufgabe:

Eine Beta-Version wurde von 9 Testkunden hinsichtlich Benutzerfreundlichkeit und Funktionalität mit einer Punkteskala von 1 (sehr schlecht) bis 10 (sehr gut) bewertet:

Kunde	1	2	3	4	5	6	7	8	9
Benutzerfreundlichkeit	1	2	3	2	5	6	2	7	8
Funktionalität	9	8	7	4	3	7	3	2	1

Bestimmen Sie den Wert eines geeigneten Zusammenhangsmaßes.

Zusammenhangsanalyse

- Lösung:

a)

Kunde	1		2		3		4		5		6		7		8		9	
Benutzerfreundl. (Bewertung | Rang)	1	9	2	7	3	5	2	7	5	4	6	3	2	7	7	2	8	1
Funktionalität (Bewertung | Rang)	9	1	8	2	7	3,5	4	5	3	6,5	7	3,5	3	6,5	2	8	1	9
Rangdifferenz d_i		8		5		1,5		2		$-2{,}5$		$-0{,}5$		0,5		-6		-8
d_i^2		64		25		2,25		4		6,25		0,25		0,25		36		64

$$\Rightarrow r_{sp} = 1 - \frac{6\sum d_i^2}{9(81-1)} = 1 - \frac{1212}{720} = -0{,}68\overline{3};$$

Zusammenhangsanalyse

- Aufgabe:

Eine Software wurde von 6 Kunden hinsichtlich Benutzerfreundlichkeit und Funktionalität mit einer Punkteskala von 1 (sehr gut) bis 6 (sehr schlecht) bewertet:

Kunde	1	2	3	4	5	6
Benutzerfreundlichkeit	1,5	2,5	1,2	2,0	3,0	2,3
Funktionalität	1,4	2,7	1,3	1,8	3,5	2,1

a) Bestimmen Sie den Wert eines geeigneten Zusammenhangsmaßes ohne Berechnung.

Zusammenhangsanalyse

- <u>Lösung:</u>

a)

Kunde	1		2		3		4		5		6	
Benutzerfreundlichkeit (Bewertung | Rang)	1,5	2	2,5	5	1,2	1	2,0	3	3,0	6	2,3	4
Funktionalität (Bewertung | Rang)	1,4	2	2,7	5	1,3	1	1,8	3	3,5	6	2,1	4

$\Rightarrow$ Vollkommen gleichsinniger monotoner Zusammenhang; $\Rightarrow r_{sp} = 1;$

Zeitreihenanalyse

- Aufgabe:

Die Tabelle zeigt den jeweiligen Monatsumsatz eines Unternehmens (in Tsd. €).

Monat	Jan	Feb	Mär	Apr	Mai	Jun	Jul	Aug	Sep	Okt	Nov	Dez
Umsatz [Tsd. €]	77	81	76	77	78	73	74	75	70	71	72	67

a) Ermitteln Sie den Umsatztrend mittels gleitender Durchschnitte geeigneter Ordnung.

b) Ermitteln Sie die monatliche additive Saisonkomponente.

Zeitreihenanalyse

- Lösung:

Monat	Umsatz	Trend (gleitender Durchschnitt 3. Ordnung)	Additive Saisonkomponente
Jan	77		
Feb	81	$78 = (77 + 81 + 76) : 3$	$3 = 81 - 78$
$Mär$	76	$78 = (81 + 76 + 77) : 3$	$-2 = 76 - 78$
Apr	77	$77 = (76 + 77 + 78) : 3$	$0 = 77 - 77$
Mai	78	76	2
Jun	73	75	-2
Jul	74	74	0
Aug	75	73	2
Sep	70	72	-2
Okt	71	71	0
Nov	72	70	2
Dez	67		

Zeitreihenanalyse

- Aufgabe:

Ein Unternehmen hat in 6 Monaten folgende Stückzahlen verkauft:

Monat	Jun	Jul	Aug	Sep	Okt	Nov
Umsatz [Stck.]	53	52	48	58	55	61

Prognostiziert wurden für den vergangenen Juni ein Umsatz von 50 Stck. sowie ein $\hat{\hat{x}}$ in Höhe von 51 Stck.

Erstellen Sie eine Umsatzprognose mittels exponentieller Glättung für den April des Folgejahres (Gewichten Sie dabei Prognosefehler mit 75%.).

Zeitreihenanalyse

- Lösung:

t	x_t	$\hat{x}_t = x^*_{t+1}$	$\hat{\hat{x}}_t$
Mai		50	51
Jun	53	$0{,}75 * 53 + 0{,}25 * 50 = 52{,}25$	$0{,}75 * 52{,}25 + 0{,}25 * 51 = 51{,}9375$
Jul	52	$0{,}75 * 52 + 0{,}25 * 52{,}25 = 52{,}0625$	$0{,}75 * 52{,}0625 + 0{,}25 * 51{,}9375 = 52{,}0313$
Aug	48	$0{,}75 * 48 + 0{,}25 * 52{,}0625 = 49{,}0156$	$0{,}75 * 49{,}0156 + 0{,}25 * 52{,}0313 = 49{,}7695$
Sep	58	$0{,}75 * 58 + 0{,}25 * 49{,}0156 = 55{,}7539$	$0{,}75 * 55{,}7539 + 0{,}25 * 49{,}7695 = 54{,}2578$
Okt	55	$0{,}75 * 55 + 0{,}25 * 55{,}7539 = 55{,}1885$	$0{,}75 * 55{,}1885 + 0{,}25 * 54{,}2578 = 54{,}9558$
Nov	61	$0{,}75 * 61 + 0{,}25 * 55{,}1885 = 59{,}5471$	$0{,}75 * 59{,}5471 + 0{,}25 * 54{,}9558 = 58{,}3993$
Dez			
Jan			
Feb			
$Mär$			

$$\Rightarrow x^{**}_{Nov+4} = 2 * 59{,}5471 - 58{,}3993 + 4\big(0{,}75 * (59{,}5471 - 58{,}3993)\big) = 64{,}14$$

A: Es ist mit einem voraussichtlichem Umsatz von ca. 64 Stck. zu rechnen.

Kombinatorik

- Lösung:

Gegeben seien die Ziffern $1 - 9$

a) Wie viele dreistellige Zahlen lassen sich daraus bilden, wenn jede Ziffer höchstens einmal vorkommen darf?

b) Wie viele dieser Zahlen sind ungerade? (wie viele gerade?)

c) Wieviel dieser Zahlen sind durch 5 teilbar?

d) Wie viele dieser Zahlen sind kleiner als 300, wie viele größer als 400?

Kombinatorik

- <u>Lösung:</u>

Die Lösungen einfach direkt dahinter

a) $|Var_{o.Z.}| = \cancel{\frac{n!}{(n-k)!} = \frac{9!}{(9-3)!}} = 9 * 8 * 7 = 504$

b) $\text{gerade:} \, 8 * 7 * 4 = 224$ ungerade: $8 * 7 * 5 = 280$

c) $\text{durch 5 teilbar:} \, 8 * 7 * 1 = 56$

d) $\text{kleiner als } 300\text{:} \, 2 * 8 * 7 = 112$ $größer\ als\ 400\text{:} \ 6 * 8 * 7 = 336$

Kombinatorik

- Aufgabe:

Wie viele Zeichen können in einem Byte dargestellt werden?

- Lösung:

Die Reihenfolge ist zu berücksichtigen; „0“ bzw. „1“ wird zurückgelegt.

$\Downarrow$ Variation $\qquad$ $\Downarrow$ mit Zurücklegen

$$\Rightarrow |Var_{m.Z.}| = n^k = 2^8 = 256$$

Stochastik

- Aufgabe:

Ein Passwort besteht aus zwei Buchstaben (ausgewählt aus den 26 Buchstaben des Alphabets) und vier Ziffern $(0-9)$, wobei Ziffern, nicht aber Buchstaben mehrfach auftreten dürfen.

Wie groß ist die Wahrscheinlichkeit ein solches Passwort beim ersten Versuch zu knacken?

- Lösung:

$$P(Panzerknacker) = \frac{1}{10^4 * \binom{26}{2} * \frac{6!}{(6-2)!}} = \frac{1}{97.500.000} = 1{,}025641 * 10^{-8}$$

Stochastik

- Aufgabe:

Drei Jäger schießen unabhängig voneinander auf ein Wildschwein. Jäger Andreas hat eine Trefferwahrscheinlichkeit von 70%, Jäger Bernd hat eine Trefferwahrscheinlichkeit von 75% und Jäger Clemens trifft zu 80%.

Wie groß ist die Wahrscheinlichkeit dafür, dass

a) Alle drei treffen

b) die Ente getroffen wird

c) genau einer der drei Jäger trifft

d) nur Jäger Bernd trifft?

Stochastik

- Lösung:

a) $P(3\,T) = 0{,}7 * 0{,}75 * 0{,}8 = 0{,}42$

b) $P(E\ tot) = 1 - P(Ente\ \text{überlebt})$

$= 1 - 0{,}3 * 0{,}25 * 0{,}2 = 0{,}985 = 98{,}5\%$

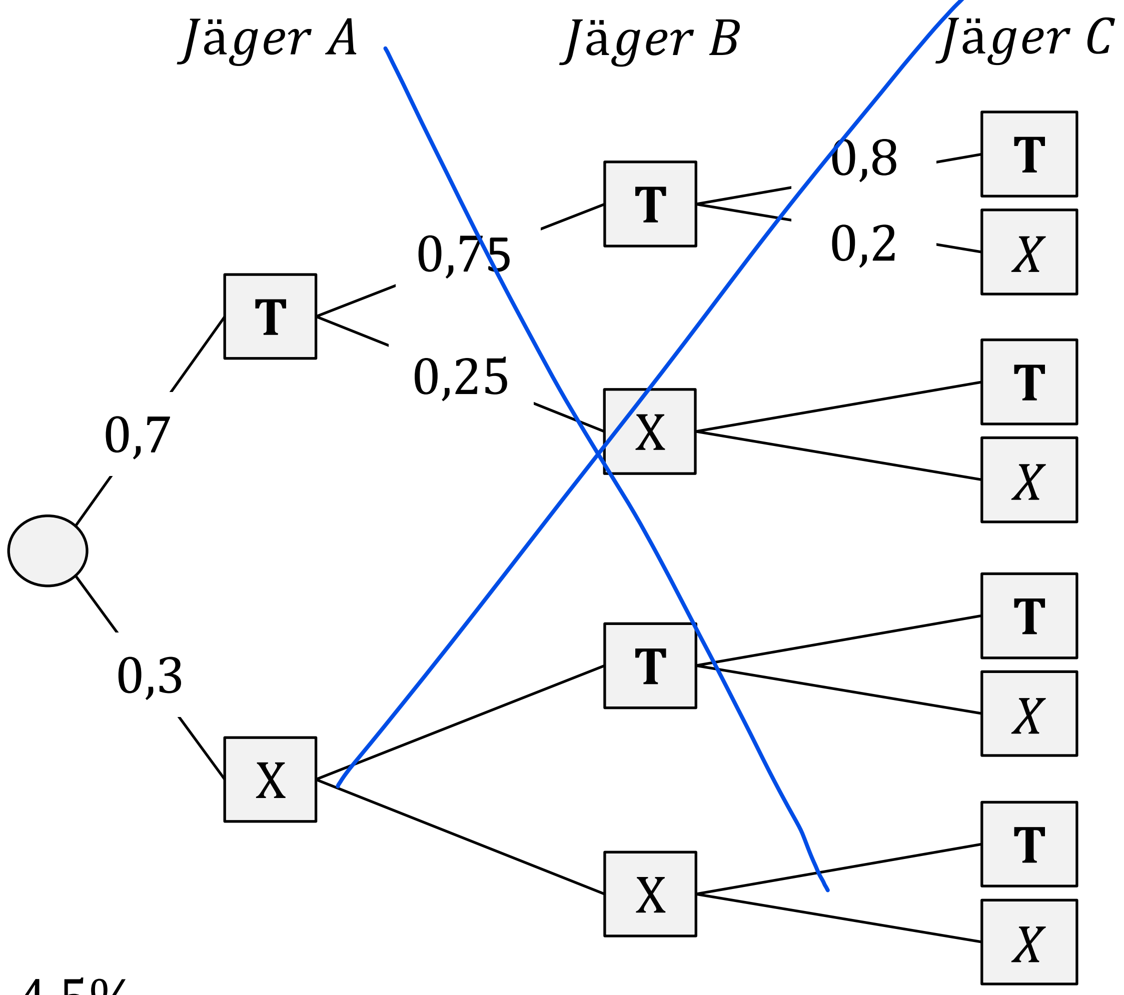
c) $P(genau\ ein\ Jäger\ trifft)$

$= 0{,}7 * 0{,}25 * 0{,}2$

$+ 0{,}3 * 0{,}75 * 0{,}2$

$+ 0{,}3 * 0{,}25 * 0{,}8 = 0{,}14 = 14\%$

d) $P(nur\ Jäger\ B\ trifft) = 0{,}3 * 0{,}75 * 0{,}2 = 0{,}045 = 4{,}5\%$



Stochastik

- Aufgabe:

Bei einem Auto sind innerhalb eines Monats unabhängig voneinander mit einer Wahrscheinlichkeit von 25% die Batterie, mit einer Wahrscheinlichkeit von 15% die Bremsen und mit einer Wahrscheinlichkeit von 10% der Antrieb defekt.

Wie groß ist die Wahrscheinlichkeit einen Monat lang ein wartungsfreies Auto fahren zu können?

- Lösung:

$$P(30\,Tage\ ALLES\ OK) = (1-0{,}25)(1-0{,}15)(1-0{,}1) = 0{,}75 * 0{,}85 * 0{,}9 = 0{,}574$$

Stochastik

- Aufgabe:

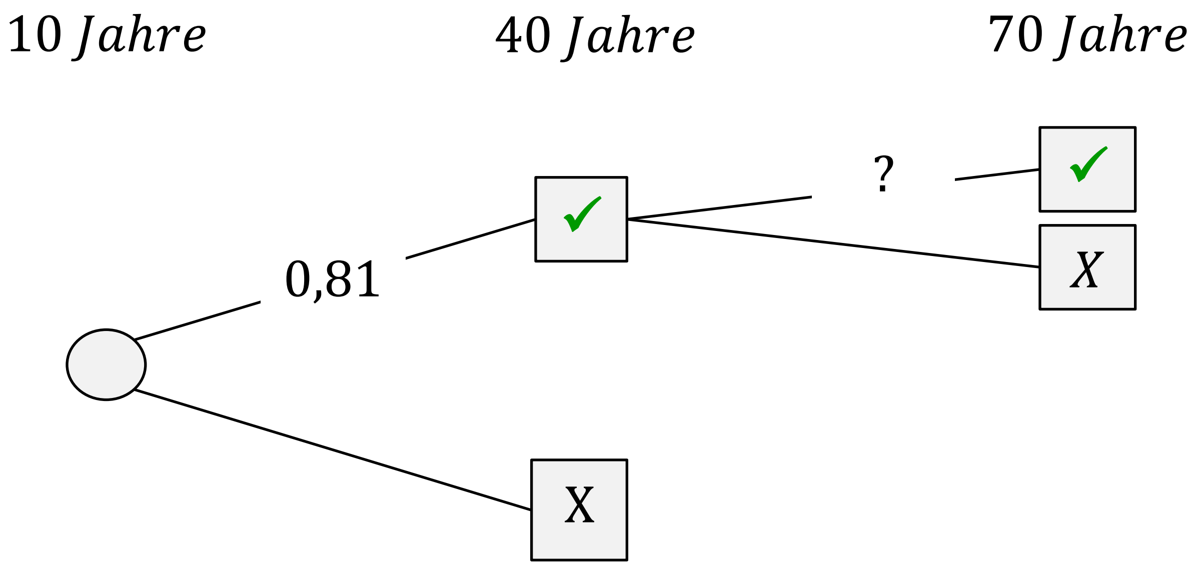
Die Wahrscheinlichkeit, dass ein 10-Jähriger mindestens 40 Jahre alt wird, sei 0,81.

Die Wahrscheinlichkeit, dass ein 10-Jähriger mindestens 70 Jahre alt wird, sei 0,38.

Wie groß ist dann die Wahrscheinlichkeit dafür, dass ein 40-Jähriger mindestens 70 Jahre alt wird?

- Lösung:

$$P(70|40) = \frac{P(40 \cap 70)}{P(40)} = \frac{0{,}38}{0{,}81} = 0{,}4691 = 46{,}91\%$$



Stochastik

- Aufgabe:

Eine Tankstelle verkauft Benzin, Diesel und Gas. Es ist bekannt, dass an dieser Tankstelle 25% der Kunden Benzin tanken und 12% Gas. 10% aller Kunden tanken Benzin und bezahlen elektronisch.

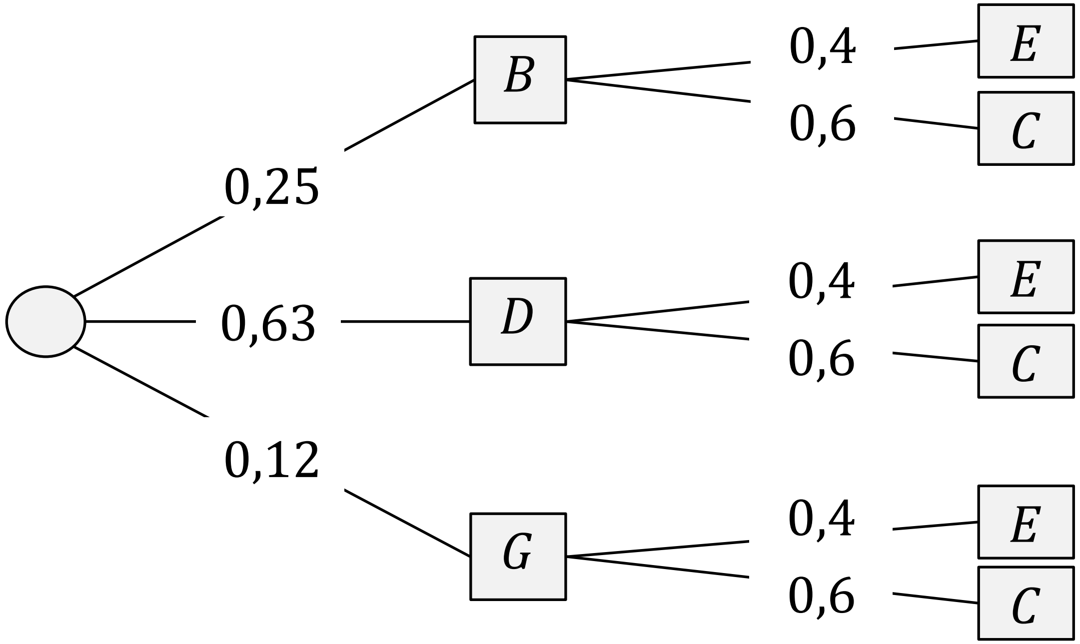
a) Wie groß ist die Wahrscheinlichkeit, dass ein Kunde der Benzin tankt, elektronisch bezahlt?

b) Wie groß ist die Wahrscheinlichkeit einer elektronischen Bezahlungsweise insgesamt? (Gehen Sie dabei davon aus, dass Bezahlweise unabhängig von der getankten Spritsorte ist)

c) Wie groß ist die Wahrscheinlichkeit, dass jemand Diesel getankt hat, wenn er bar bezahlt?

Stochastik

- Lösung:



a) $P(E|B) = \dfrac{P(B \cap E)}{P(B)} = \dfrac{0,1}{0,25} = 0,4$

b) $P(E) = 0,25 * 0,4 + 0,63 * 0,4 + 0,12 * 0,4 = 0,4 * (0,25 + 0,63 + 0,12) = 0,4 * 1 = 0,4 = 40\%$

c) $P(D|C) = \dfrac{P(D) * P(C|D)}{P(C)} = \dfrac{0,63 * 0,6}{0,6} = 0,63 = 63\%$

* Anmerkung: Die gleichen Wahrscheinlichkeiten resultieren aus der angenommenen Unabhängigkeit der Bezahlweise.

Stochastik

- Aufgabe:

Die Analyse von 200 Mails (100 Spam-Mails und 100 normale Mails) hinsichtlich der Wörter „***exklusiv***“ und „***kostenlos***“ ergab folgendes Ergebnis:

	$kein\ Spam$	$Spam$
$exklusiv$	8	32
$kostenlos$	6	12

Berechnen Sie die Wahrscheinlichkeit dafür, dass eine Mail Spam ist, wenn in ihr die Worte „exklusiv“ und „kostenlos“ enthalten sind.

Stochastik

- Lösung:

	$kein\ Spam$	$Spam$
$\boldsymbol{exklusiv}$	6	32
$\boldsymbol{kostenlos}$	8	12

$$P(e|S) = \frac{32}{100} = 0{,}32$$

$$P(k|S) = \frac{12}{100} = 0{,}12$$

$$P(S|e \cap k) = \frac{0{,}32 * 0{,}12}{0{,}32 * 0{,}12 + 0{,}06 * 0{,}08} = 0{,}8888 = 88{,}8\%$$

$$P(e|\overline{S}) = \frac{6}{100} = 0{,}06$$

$$P(k|\overline{S}) = \frac{8}{100} = 0{,}08$$

Wahrscheinlichkeitsverteilungen

- Aufgabe:

Gegeben sei folgende Wertetabelle:

x_i	1	2	3	5
$f(x_i)$	$\frac{c}{6}$	$\frac{c}{4}$	$\frac{c}{12}$	c

a) Welchen Wert muss c annehmen, damit es sich um eine Wahrscheinlichkeitsverteilung handelt?

b) Bestimmen Sie (mit der Lösung für c aus Teilaufgabe a)) $P(0 \leq X < 3)$; $P(1 \leq X < 3)$; $P(1 < X < 3)$;

c) Bestimmen Sie die Verteilungsfunktion.

Wahrscheinlichkeitsverteilungen

- Lösung:

a) $\frac{c}{6}+\frac{c}{4}+\frac{c}{12}+c=1; \quad \Leftrightarrow \quad \frac{2+3+1+12}{12}c=1 \quad \Leftrightarrow \quad c=\frac{2}{3}$

b) $P(0 \le X < 3)=\frac{1}{9}+\frac{1}{6}=\frac{5}{18}; \quad P(1 \le X < 3)=\frac{1}{9}+\frac{1}{6}=\frac{5}{18}; \quad P(1 < X < 3)=\frac{1}{6}$

c)

$$F(x)=\begin{cases} 0 & für\ x<1 \\ \frac{1}{9} & für\ 1 \le x < 2 \\ \frac{5}{18} & für\ 2 \le x < 3 \\ \frac{1}{3} & für\ 3 \le x < 5 \\ 1 & für\ 5 \le x \end{cases}$$

x_i	1	2	3	5
$f(x_i)$	$\frac{1}{9}$	$\frac{1}{6}$	$\frac{1}{18}$	$\frac{2}{3}$

Integralrechnung

- Aufgabe:

Bestimmen Sie folgende Integrale:

$$\int_0^{\frac{1}{3}} 3e^{3x} dx =$$

$$\int_1^e x^{-1} dx =$$

$$\int_0^2 \int_0^1 2xy \, dxdy =$$

Integralrechnung

- Lösung:

Bestimmen Sie folgende Integrale:

$$\int_0^{\frac{1}{3}} 3e^{3x} dx = [e^{3x}]_0^{\frac{1}{3}} = e^1 - e^0 = e - 1$$

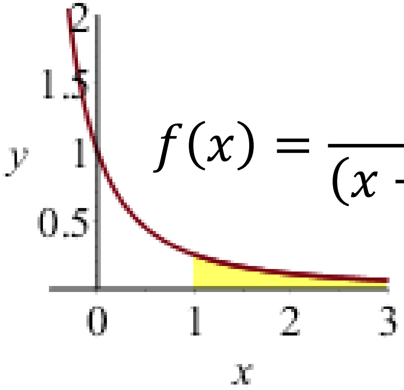
$$\int_1^e x^{-1} dx = \int_1^e \frac{1}{x} dx = [\ln(x)]_1^e = \ln(e) - \ln(1) = 1 - 0 = 1$$

$$\int_0^2 \int_0^1 2xy\, dx dy = \int_0^2 2y \left[\frac{1}{2}x^2\right]_0^1 dy = \int_0^2 2y \left[\frac{1}{2} - 0\right]_0^1 dy = \int_0^2 y\, dy = \left[\frac{1}{2}y^2\right]_0^2 = \frac{1}{2} * 2^2 - 0 = \frac{4}{2} = 2;$$

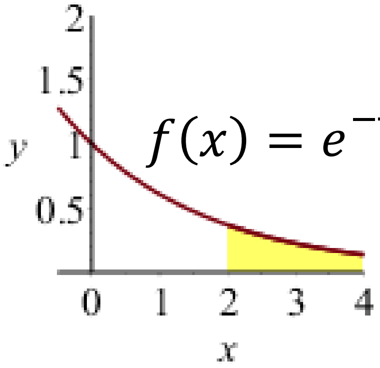
Integralrechnung

- <u>Aufgabe:</u>

Bestimmen Sie für folgende Funktionen jeweils das markierte uneigentliche Integral:



$$f(x) = \frac{1}{(x+1)^2}$$

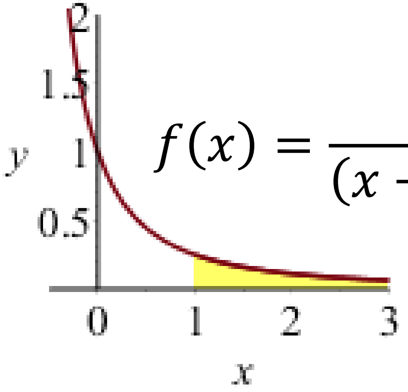


$$f(x) = e^{-\frac{1}{2}x}$$

Integralrechnung

- Lösung:

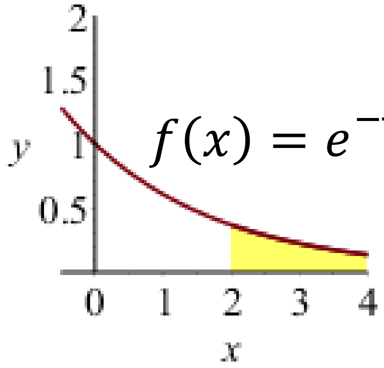
Bestimmen Sie für folgende Funktionen jeweils das markierte uneigentliche Integral:



$$f(x) = \frac{1}{(x+1)^2}$$

$$\int_1^{\infty} \frac{1}{(x+1)^2}\,dx = \int_1^{\infty} (x+1)^{-2}\,dx = \left[-(x+1)^{-1}\right]_1^{\infty} = -\frac{1}{\infty+1} + \frac{1}{1+1} = $$

$$= 0 + \frac{1}{2} = \frac{1}{2}$$



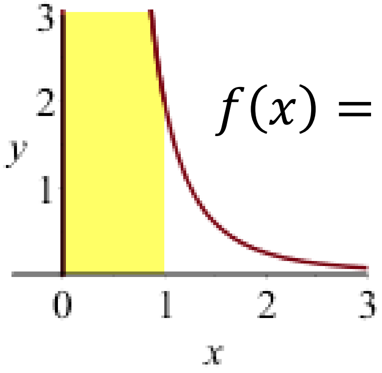
$$f(x) = e^{-\frac{1}{2}x}$$

$$\int_2^{\infty} e^{-\frac{1}{2}x}\,dx = \left[-2e^{-\frac{1}{2}x}\right]_2^{\infty} = -2e^{-\frac{1}{2}*\infty} + 2e^{-\frac{1}{2}*2} = -\frac{2}{e^{\frac{1}{2}*\infty}} + \frac{2}{e^1} = 0 + \frac{2}{e} = \frac{2}{e}$$

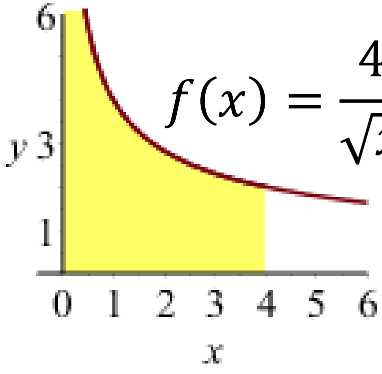
Integralrechnung

- Aufgabe:

Bestimmen Sie für folgende Funktionen jeweils das markierte uneigentliche Integral:



$$f(x) = \frac{2}{x^3}$$

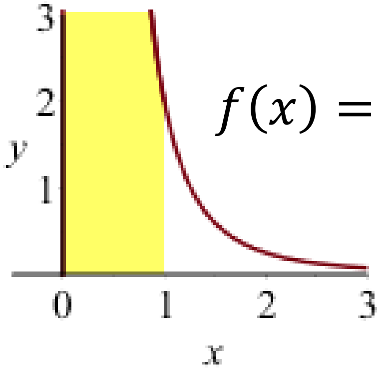


$$f(x) = \frac{4}{\sqrt{x}}$$

Integralrechnung

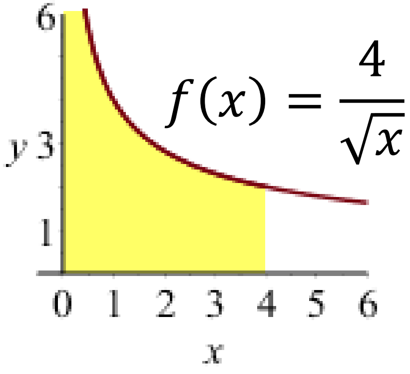
- Lösung:

Bestimmen Sie für folgende Funktionen jeweils das markierte uneigentliche Integral:



$$f(x) = \frac{2}{x^3}$$

$$\int_0^1 \frac{2}{x^3}dx = \int_0^1 2x^{-3}dx = [-x^{-2}]_0^1 = \left[-\frac{1}{x^2}\right]_0^1 = -\frac{1}{1^2} + \lim_{x\to 0}\frac{1}{x^2} = -1 + \infty = \infty$$



$$f(x) = \frac{4}{\sqrt{x}}$$

$$\int_0^4 \frac{4}{\sqrt{x}}dx = \int_0^4 4x^{-\frac{1}{2}}dx = 4\left[2x^{\frac{1}{2}}\right]_0^4 = 4\left(2\sqrt{4} - 2\sqrt{0}\right) = 4(4-0) = 16$$

Wahrscheinlichkeitsverteilungen

- Aufgabe:

Gegeben sei folgende Funktion: $f(x) = \begin{cases} \frac{1}{2}x^2 & für\ 0 < x < a \\ 0 & sonst \end{cases}$

Welchen Wert muss a annehmen, damit $f(x)$ eine Dichtefunktion einer Zufallsvariablen ist?

- Lösung: $\int_0^a \frac{1}{2}x^2 dx = 1; \quad \Leftrightarrow \quad \frac{1}{2}\left[\frac{1}{3}x^3\right]_0^a = 1 \Leftrightarrow \quad \frac{1}{3}a^3 = 2 \quad \Rightarrow a = \sqrt[3]{6}$

Wahrscheinlichkeitsverteilungen

- Aufgabe:/ Lösung:

Begründen Sie, warum die folgenden Funktionen nicht als Dichtefunktion einer stetigen Zufallsvariable aufgefasst werden können?

a) $f(x) = \begin{cases} 0 & für\ x \leq 0 \\ \frac{3}{4}x^2 & für\ 0 < x < 5 \\ 0 & für\ 5 \leq x \end{cases}$

$$\frac{3}{4}\int_0^5 x^2 dx = \frac{125}{4} > 1; \Rightarrow f(x) \text{ ist keine Dichtefunktion}$$

b) $f(x) = \begin{cases} 0 & für\ x \leq 0 \\ x & für\ 0 < x \leq 1 \\ 1 & für\ 1 < x \end{cases}$

$$\int_0^1 x\, dx = \frac{1}{2} < 1; \Rightarrow f(x) \text{ ist keine Dichtefunktion}$$

c) $f(x) = \begin{cases} 0 & für\ x \leq 0 \\ 2x & für\ 0 < x \leq 1 \\ 1 & für\ 1 < x \end{cases}$

$f(x) = 1\ für\ 1 < x; \Rightarrow f(x)$ ist keine Dichtefunktion;
(das gilt analog auch zusätzlich bei der vorherigen Funktion)

Wahrscheinlichkeitsverteilungen

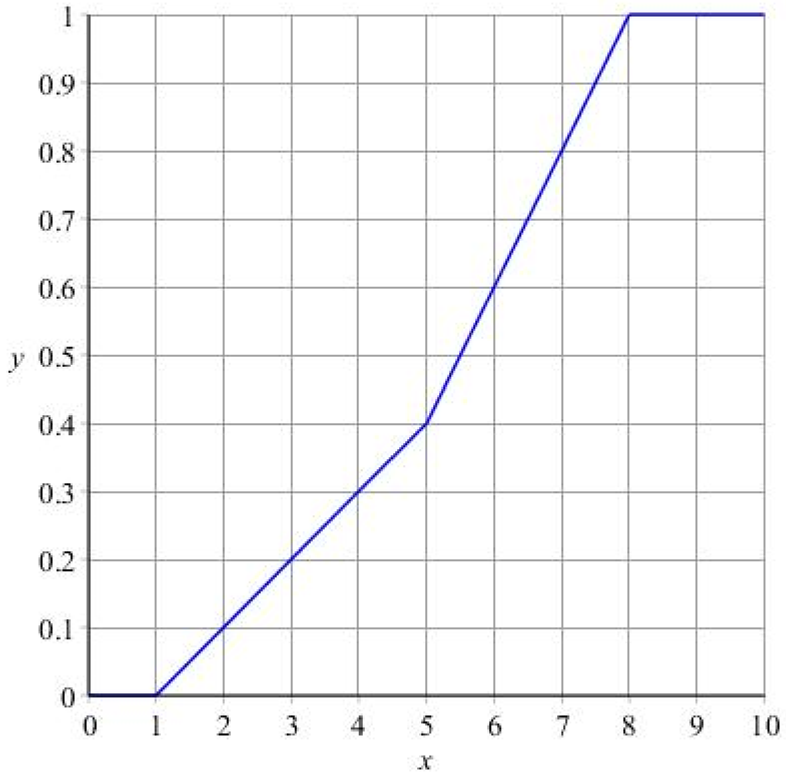
- Aufgabe:

Gegeben sei die dargestellte Verteilungsfunktion:

a) Bestimmen Sie $P(X \leq 7)$.

b) Bestimmen Sie $P(3 < X \leq 4 \; oder \; X \geq 6)$

c) Bestimmen Sie die Dichtefunktion.



- Lösung:

a) $P(X \leq 7) = 0{,}8$

b) $P(3 < X \leq 4 \; oder \; X \geq 6) = (0{,}3 - 0{,}2) + (1 - 0{,}6) = 0{,}5$

c) $$f(x) = \begin{cases} 0{,}1 & für \; 1 \leq x < 5 \\ 0{,}2 & für \; 5 \leq x < 8 \\ 0 & sonst \end{cases}$$

Wahrscheinlichkeitsverteilungen

- Aufgabe:

Zeigen Sie: $\sum_i (x_i - \mu)^2 f(x_i) = \sum_i {x_i}^2 f(x_i) - \mu^2$

- Lösung:

$$\sum_i (x_i - \mu)^2 f(x_i) = \sum (x_i{}^2 - 2x_i\mu + \mu^2) * f(x_i) = \sum_i {x_i}^2 f(x_i) \; -2\mu \underbrace{\sum x_i * f(x_i)}_{=\mu} \; +\mu^2 \underbrace{\sum f(x_i)}_{=1}$$

$$= \sum_i {x_i}^2 f(x_i) - 2\mu^2 + \mu^2 = \sum_i {x_i}^2 f(x_i) - \mu^2$$

Wahrscheinlichkeitsverteilungen

- Aufgabe:

Gegeben sei die Dichtefunktion einer stetigen Zufallsvariablen X: $f(x) = \begin{cases} \frac{1}{4}x^3 & für\ 0 < x < 2 \\ 0 & sonst \end{cases}$

a) Bestimmen Sie die Verteilungsfunktion.

b) Bestimmen Sie $P(1 < X \leq 2)$.

c) Bestimmen Sie den Erwartungswert und die Varianz.

- Lösung:

a) $F(x) = \begin{cases} 0 & für \quad x < 0 \\ \frac{1}{16}x^4 & für\ 0 \leq x < 2 \\ 1 & für\ \ 2 \leq x \end{cases}$

b) $P(1 < X \leq 2) = F(2) - F(1) = 1 - \frac{1}{16} = \frac{15}{16}$

c) $\mu = \int_0^2 \frac{1}{4}x^4\,\mathrm{dx} = \frac{1}{4}\left[\frac{1}{5}x^5\right]_0^2 = \frac{8}{5};$

$\sigma^2 = \int_0^2 \frac{1}{4}x^5\,\mathrm{dx} - \frac{64}{25} = \frac{64}{24} - \frac{64}{25} = \frac{8}{75};$

Wahrscheinlichkeitsverteilungen

- Aufgabe:

Ein Unternehmen produziert rechteckige Betonbauteile deren Länge und Breite produktionsbedingt zufällig (und unabhängig voneinander) schwanken. Die Bauteile sind im Durchschnitt $1000\ cm$ lang (mit einer Varianz von $0{,}025$) und $500\ cm$ breit (mit einer Varianz von $0{,}012$).

Quelle: Youtube: Die zehn größten Baukatastrophen

Bestimmen Sie den Erwartungswert sowie die Standardabweichung für den Umfang der Bauteile.

- Lösung:

$$U = 2L + 2B \Rightarrow \quad \mu_U = 2\mu_L + 2\mu_B = 2 * 1000\ cm + 2 * 500\ cm = 3000\ cm$$

$$\sigma_U^2 = 2^2\sigma_L^2 + 2^2\sigma_B^2 = 4 * 0{,}025 + 4 * 0{,}012 = 0{,}148 \quad \Rightarrow \quad \sigma_U = \sqrt{0{,}148} = 0{,}3847\ cm$$

Wahrscheinlichkeitsverteilungen

- Aufgabe:

Gegeben sei folgende Zufallsvariable X mit der Dichtefunktion: $f(x) = \begin{cases} \frac{3}{8}x^2 & für\ 0 < x < 2 \\ 0 & sonst \end{cases}$

Berechnen Sie Erwartungswert und Standardabweichung für die Zufallsvariable $Y = -2X + 4{,}25$

- Lösung:

$$\mu_X = \int_0^2 \frac{3}{8}x^3 dx = \frac{3}{8}\left[\frac{1}{4}x^4\right]_0^2 = \frac{3*16}{32} = \frac{3}{2} \qquad \Rightarrow \qquad \mu_Y = a\mu_X + b = -2*\frac{3}{2} + 4{,}25 = 1{,}25$$

$$\sigma_X^2 = \int_0^2 \frac{3}{8}x^4 dx - \mu_X^2 = \frac{3}{8}\left[\frac{1}{5}x^5\right]_0^2 - \frac{9}{4} = \frac{3*32}{40} - 1{,}5^2 = 0{,}15$$

$$\Rightarrow \sigma_Y^2 = a^2\sigma_X^2 = (-2)^2 * 0{,}15 = 0{,}6 \qquad \Rightarrow \sigma_Y = \sqrt{0{,}6} \approx 0{,}77$$

Wahrscheinlichkeitsverteilungen

- Aufgabe:

Die paarweise stochastisch unabhängigen Zufallsvariablen

$X_1,\ X_2,\ X_3$ haben folgende Erwartungswerte und Varianzen:

Zufallsvariable	X_1	X_2	X_3
μ_{X_i}	-1	1	2
$\sigma^2_{X_i}$	$0,25$	$0,5$	$0,75$

Berechnen Sie Erwartungswert und Standardabweichung der Zufallsvariable

$Y = 2X_1 - 3X_2 + X_3 + 0,25$

- Lösung:

$$\mu_Y = 2 * (-1) - 3 * 1 + 1 * 2 + 0,25 = -2,75;$$

$$\sigma_Y = \sqrt{\sigma_Y^2} = \sqrt{2^2 * 0,25 + (-3)^2 * 0,5 + 1^2 * 0,75} = \sqrt{6,25} \approx 2,5$$

Wahrscheinlichkeitsverteilungen

- Aufgabe:/ Lösung (Binomialverteilung):

  Studenten bestehen mit einer Wahrscheinlichkeit von $0{,}75$ die Statistikklausur. Wie hoch ist die Wahrscheinlichkeit, dass von 10 Studenten

a) keiner die Klausur besteht (ohne Tab.!)? $P(B=0)=(1-0{,}75)^{10}=9{,}5*10^{-7}$

b) höchstens 6 die Klausur bestehen? $P(B\leq 6)\overset{*}{=}P(\overline{B}\geq 4)=1-P(\overline{B}\leq 3)=1-0{,}7759=0{,}2241$

c) mindestens 6 die Klausur bestehen? $P(B\geq 6)=1-P(B\leq 5)=P(\overline{B}\leq 4)=0{,}9219$

d) genau einer durchfällt? $P(\overline{B}=1)=\left(P(B=9)\right)=0{,}1877$

* Umrechnung nötig, da Tabelle eigentlich nur p –Werte $\leq 0{,}5$ berücksichtigt:

$P(B)=0{,}75;\ \ bzw.\, P(\overline{B})=1-0{,}75=0{,}25$

Binomialverteilung

- Aufgabe:/ Lösung:

Über einen Nachrichtenkanal werden Zeichen übertragen. Durch Störeinflüsse wird jedes Zeichen mit der unbekannten Wahrscheinlichkeit p falsch übertragen. (Die Störung der einzelnen Zeichen erfolgt dabei unabhängig voneinander)

Wie groß darf p höchstens sein, wenn die Wahrscheinlichkeit, dass bei 100 übertragenen Zeichen mehr als eines falsch übertragen wird, höchstens 10% betragen darf?

- Lösung:

a) $\boldsymbol{W'keit}, dass\ bei\ 100\ Zeichen\ mehr\ als\ eines\ falsch\ \text{übertragen}\ wird, \boldsymbol{\text{beträgt}\ \text{höchstens}}\ 10\%$

$$n = 100$$

$$= \boldsymbol{P(X} > 1\boldsymbol{)} \leq 0{,}1 \quad \Leftrightarrow \quad 1 - P(X \leq 1) \leq 0{,}1 \quad \Leftrightarrow \quad 0{,}9 \leq P(X \leq 1) \quad \Rightarrow \quad p \approx 0{,}54\%$$

Binomialverteilung

- Aufgabe:/ Lösung:

Bei einer Binomialverteilung gibt es mit einer Wahrscheinlichkeit von mindestens 87% höchstens zehn Treffer

a) Bestimmen Sie (näherungsweise) den extremsten Wert für p, wenn $n = 20$ ist.

b) Bestimmen Sie (näherungsweise) den extremsten Wert für n, wenn $\mathrm{p} = 0{,}45$ ist.

c) Handelt es sich bei den ermittelten Werten in a) und b) jeweils um einen Minimal- oder einen Maximalwert?

- Lösung:

$a) \quad P(X \leq 10) \geq 0{,}87; \quad n = 20; \qquad \Rightarrow p \leq 0{,}4$

$b) \quad P(X \leq 10) \geq 0{,}87; \quad p = 0{,}45; \qquad \Rightarrow n \leq 18$

Wahrscheinlichkeitsverteilungen

- Aufgabe:

  Wie hoch ist die Wahrscheinlichkeit beim Lotto „6 aus 49“

$a)$ 6 Richtige zu tippen?

b) den Jackpot zu knacken? (Hinweis: Um den Jackpot zu knacken muss zusätzlich zu den 6 Richtigen auch noch die Superzahl (eine Zahl von $0-9$) richtig sein.)

- Lösung (Hypergeometrische Verteilung):

$$a)\quad P(6\ Richtige) = H(6|49;6;6) = \frac{\binom{6}{6}\binom{49-6}{6-6}}{\binom{49}{6}} = \frac{1}{13.983.816} \approx 7{,}15 * 10^{-8}$$

$$b)\quad P(Jackpot) = H(6|49;6;6) * \frac{1}{10} = \frac{1}{139.838.160} \approx 7{,}15 * 10^{-9}$$

Wahrscheinlichkeitsverteilungen

- Aufgabe:

  Die Elektronik-Fachhandelskette Jupiter hat in einer Großstadt 3 Filialen. Die Anzahl der Kunden, die die Filialen pro Stunde betreten sei mit den folgenden Parametern verteilt:

Parameter	F_1	F_2	F_3
μ_{F_i}	1	2	$2{,}5$
$\sigma^2_{F_i}$	1	2	$2{,}5$

  Die Ankünfte der Kunden seien voneinander unabhängig.

Wie groß ist die Wahrscheinlichkeit, dass

a) innerhalb einer Stunde insgesamt 8 Kunden die Filialen betreten?

b) innerhalb einer Stunde nicht mehr als 2 Kunden die Filialen betreten?

c) insgesamt 12 Kunden die 3 Filialen innerhalb von 2 Stunden betreten?

Poissonverteilung

- Lösung:

$$a)\quad \left.\begin{aligned} \mu_Y &= 1*1+1*2+1*2{,}5 = 5{,}5 \\ \sigma_Y^2 &= 1^2*1+1^2*2+1^2*2{,}5 = 5{,}5 \end{aligned}\right\} \mu_Y = \sigma_Y^2 \quad \Rightarrow Y \sim Ps(5{,}5) \quad Ps(8|5{,}5) = \frac{5{,}5^8}{8!}e^{-5{,}5} \approx 0{,}085$$

$$b)\quad P(Y \leq 2) = Ps(0|5{,}5) + Ps(1|5{,}5) + Ps(2|5{,}5) = \frac{5{,}5^0}{0!}e^{-5{,}5} + \frac{5{,}5^1}{1!}e^{-5{,}5} + \frac{5{,}5^2}{2!}e^{-5{,}5} \approx 0{,}088$$

$$c)\quad \mu_{2Y} = 2(1*1+1*2+1*2{,}5) = 11 \quad \Rightarrow Y \sim Ps(11) \quad Ps(12|11) = \frac{11^{12}}{12!}e^{-11} \approx 0{,}1094$$

Wahrscheinlichkeitsverteilungen

- Aufgabe:

Die Lebensdauer von Akkus des Typs Powerix sei normalverteilt mit einer durchschnittlichen Lebenserwartung $3 \pm 0{,}5\ Jahren$.

Wie groß ist die Wahrscheinlichkeit, dass

a) ein Akku weniger als 2 Jahre hält?

b) ein Akku länger als 3 Jahre hält?

c) ein Akku länger als 4 Jahre hält?

d) ein Akku länger als $4{,}5$ Jahre hält?

e) ein Akku zwischen $2{,}5 - 3{,}5$ Jahre hält?

f) zwei Akkus länger als $15\ Monate$ halten?

Normalverteilung

- Lösung:

$$P(X<2):\ \Rightarrow\quad z=\frac{x-\mu}{\sigma}=\frac{2-3}{0{,}5}=-2\qquad \Rightarrow P(z<-2)=P(z>2)=1-P(z<2)=0{,}0228$$

$$P(X>3):\ \Rightarrow\quad z=\frac{x-\mu}{\sigma}=\frac{3-3}{0{,}5}=0\qquad \Rightarrow P(z>0)=1-P(z<0)=0{,}5$$

$$P(X>4):\ \Rightarrow\quad z=\frac{x-\mu}{\sigma}=\frac{4-3}{0{,}5}=2\qquad \Rightarrow P(z>2)=0{,}0228$$

$$P(X>4{,}5):\ \Rightarrow\quad z=\frac{x-\mu}{\sigma}=\frac{4{,}5-3}{0{,}5}=3\qquad \Rightarrow P(z>3)=1-P(z<3)=1-0{,}9987=0{,}0013$$

Normalverteilung

- Lösung:

$e)\quad P(2{,}5 < X < 3{,}5):\ \Rightarrow\quad z_1 = \frac{2{,}5-3}{0{,}5} = -1;\ bzw.\ z_2 = \frac{3{,}5-3}{0{,}5} = 1;$

$\Rightarrow P(-1 < z < 1) = P(z < 1) - P(z < -1) = P(z < 1) - \big(1 - P(z < 1)\big)$

$= P(z < 1) - 1 + P(z < 1) = 2P(z < 1) - 1 = 2 * 0{,}8413 - 1 = 0{,}6826$

$f)\quad P(X > 1{,}25):\ \Rightarrow\quad z = \frac{1{,}25-3}{0{,}5} = -3{,}5;$

$\Rightarrow P\big(z > (-3{,}5)\big) = 1 - P(z < -3{,}5) = 1 - \big(1 - P(z < 3{,}5)\big) = P(z < 3{,}5) = 0{,}9998$

$P(X_1 > 1{,}25)\ \cap P(X_2 > 1{,}25) = 0{,}9998 * 0{,}9998 = 0{,}9996$

Wahrscheinlichkeitsverteilungen

- Aufgabe:

  Der Durchmesser von Feinbauteilen ist $N(\mu; 0{,}1mm)$-verteilt, bei einem durchschnittlichen Durchmesser von $10\ mm$.

a) Es seien nur Bauteile mit einem Durchmesser von $9{,}92\ mm$ bis $10{,}18\ mm$ brauchbar. Wie groß ist der Ausschussanteil?

b) Es seien nur Bauteile mit einem Maximaldurchmesser von $10{,}02\ mm$ brauchbar. Wie groß darf die Standardabweichung des Durchmessers höchstens sein, damit der Ausschussanteil nicht mehr als 15% beträgt

Normalverteilung

- Lösung:

a) $\mu = 10mm \Rightarrow X \sim N(10mm; 0{,}1mm) \Rightarrow P(9{,}92 \leq X \leq 10{,}18)$: $\Rightarrow$ $z_1 = \frac{9{,}92-10}{0{,}1} = -0{,}8$; $bzw.$ $z_2 = \frac{10{,}18-10}{0{,}1} = 1{,}8$;

$$\Rightarrow P(-0{,}8 < z < 1{,}8) = P(z < 1{,}8) - P(z < -0{,}8) = P(z < 1{,}8) - \big(1 - P(z < 0{,}8)\big)$$

Ausschussanteil: $1 - 0{,}7522 = 0{,}2478$

b) $P\left(Z \leq z = \frac{X-\mu}{\sigma}\right) > 0{,}85$: $\Rightarrow$ $z = 1{,}04$;

$$\Rightarrow 1{,}04 = \frac{10{,}02 - 10}{\sigma} \qquad \Rightarrow \sigma = 0{,}019 \text{ mm}$$

χ^2 -Verteilung

- Aufgabe:

Die Zufallsvariable sei χ^2-verteilt mit 15 Freiheitsgraden.

a) Bestimmen Sie $P(6{,}262 \leq X \leq 32{,}801)$.

b) Für welches x_0 gilt $F(x_0) = 0{,}75$?

c) Wie viele unabhängige Zufallsvariablen hat eine χ^2-verteilte Zufallsvariable für die gilt $P(X \leq 18{,}101) = 0{,}3$?

- Lösung:

$a)$ $P(6{,}262 \leq X \leq 32{,}801) = P(X \leq 32{,}801) - P(X \leq 6{,}262) = 0{,}995 - 0{,}025 = 0{,}97.$

$b)$ $x_0 = 18{,}245$

$c)$ $v = 22$

Approximation von Verteilungen

- Aufgabe:

  Bei der Überprüfung von 100.000 LED-Lampen wurde festgestellt, dass 3.500 von mangelhafter Qualität waren.

  Bestimmen Sie die Wahrscheinlichkeit, mit der sich in einer Stichprobe von 100 Lampen mindestens 3 Mangelhafte befinden?

- Lösung:

$$\frac{M}{N} = \frac{3.500}{100.000} = 0{,}035 \; < 0{,}1 \qquad n = 100 \; > 30 \qquad \frac{n}{N} = \frac{100}{100.000} \; < 0{,}05$$

$$\Rightarrow H(100.000; 3.500; 100) \approx Ps\left(n * \frac{M}{N} = \frac{100 * 3.500}{100.000} = 3{,}5\right)$$

$$\Rightarrow P\left(L_{mangelh.} \geq 3\right) = 1 - F(2) = 1 - \sum_{i=0}^{2} Ps(i|3{,}5) \approx 1 - 0{,}32 = 0{,}68$$

Approximation von Verteilungen

- Aufgabe:

Gegeben sei die Zufallsvariable X mit $\mathrm{X} \sim B(50; 0{,}38)$.

Bestimmen Sie $P(X \leq 25)$.

- Lösung:

$$n * p * (1-p) = 11{,}78 > 9 \quad \Rightarrow B(50; 0{,}38) \approx N\left(50 * 0{,}38; \sqrt{50 * 0{,}38 * (1-0{,}38)}\right) = N(19; 3{,}4322)$$

Stetigkeitskorrektur

$$P(X \leq 25) = P\left(Z \leq \frac{25{,}5 - 19}{3{,}4322}\right) \approx P(Z \leq 1{,}894) = 0{,}97089$$

Stichprobenauswahl

- Aufgabe:

Aus einer Kundenkartei von 8.000 Kunden sollen mindestens 350 zufällig systematisch ausgewählt werden; begonnen wird mit dem 12. Kunden der Kartei.

a) Bestimmen Sie die nächsten 3 zu ziehenden Kunden.

b) Aus welchem Zahlenbereich darf der erste zu ziehende Kunde ausgewählt werden?

- Lösung:

$a)\quad \lfloor 8.000 : 350 \rfloor = \lfloor 22{,}85 \rfloor = 22; \qquad \Rightarrow 12.,\ 34.,56.,78.,usw.$

$b)\quad K_1 \in \{1; \dots; x\}:\ 22 * 350 + x = 8000; \Rightarrow x = 300$

Schätzverfahren

- Aufgabe:

Aus einer großen Kartoffellieferung werden 10 Säcke entnommen und deren jeweiliges Gewicht notiert: $9{,}5; 10{,}5; 10{,}0; 10{,}0; 10{,}2; 10{,}0; 10{,}4; 9{,}6; 9{,}8; 10{,}0 \quad [kg]$

(Das Gewicht der Säcke sei näherungsweise normalverteilt.)

Bestimmen Sie ein 95% –Konfidenzintervall für das durchschnittliche Kartoffelsackgewicht der Lieferung.

- Lösung:

$\overline{x} = 10; \quad \sigma_S = 0{,}3 \Rightarrow \hat{\sigma}_{\mu_S}^2 \approx \frac{\sigma_S^2}{n-1}$ bzw. $\hat{\sigma}_{\mu_S} = \frac{\sigma_S}{\sqrt{n-1}} = \frac{0{,}3}{\sqrt{9}} = 0{,}1$

$n - 1 < 30; \Rightarrow \overline{X} \sim t(v = 9) \Rightarrow t = 2{,}262$

$\Rightarrow \mu_u = 10 - 2{,}262 * 0{,}1 = 9{,}7738; \;\; \& \quad \mu_o = 10 + 2{,}262 * 0{,}1 = 10{,}2262$

Schätzverfahren

- Aufgabe:

  Eine Stichprobe von 65 Batterien ergibt eine mittlere Lebensdauer von 75 ± 8 Stunden. (Die Lebensdauer ist näherungsweise normalverteilt.)

a) Wie groß ist die mittlere Lebensdauer der 1.500 Batterien einer Lieferung bei einem Konfidenzintervall von 90%?

b) Wie groß ist die mittlere Lebensdauer der 1.500 Batterien einer Lieferung mindestens bei einem Konfidenzintervall von 90%?

c) Wie groß ist die mittlere Lebensdauer der 1.500 Batterien einer Lieferung höchstens bei einem Konfidenzintervall von 92%?

Schätzverfahren

- Lösung:

$$a)\quad \frac{n}{N} < 0{,}05;\ \sigma\ unbekannt\ \Rightarrow \hat{\sigma}_{\mu_S} = \frac{\sigma_S}{\sqrt{n-1}} = \frac{8}{\sqrt{64}} = 1;\quad 1-\alpha = 90\%$$

$$n-1 > 30\ \ \&\ zweiseitiges\ Konfidenzintervall \qquad \Rightarrow z = 1{,}645$$

$$\Rightarrow \mu_{u/o} = 75 \pm 1{,}645$$

$$a)\quad \frac{n}{N} < 0{,}05;\ \sigma\ unbekannt\ \Rightarrow \hat{\sigma}_{\mu_S} = \frac{\sigma_S}{\sqrt{n-1}} = \frac{8}{\sqrt{64}} = 1; 1-\alpha = 90\%$$

$$n-1 > 30\ \ \&\ einseitiges\ Konfidenzintervall \qquad \Rightarrow z = 1{,}28$$

$$\Rightarrow \mu_u = 75 - 1{,}28 * 1 = 73{,}72$$

$$a)\quad \frac{n}{N} < 0{,}05;\ \sigma\ unbekannt\ \Rightarrow \hat{\sigma}_{\mu_S} = \frac{\sigma_S}{\sqrt{n-1}} = \frac{8}{\sqrt{64}} = 1;\ \ 1-\alpha = 92\%$$

$$n-1 > 30\ \ \&\ einseitiges\ Konfidenzintervall \qquad \Rightarrow z = 1{,}41$$

$$\Rightarrow \mu_o = 75 + 1{,}41 * 1 = 76{,}41$$

Schätzverfahren

- Aufgabe:

Eine unbekannte Wahrscheinlichkeit p soll mit einer Abweichung von $\pm 1\%$ genau geschätzt werden.

a) Wie hängen Konfidenzniveau und Stichprobenumfang zusammen?

b) Berechnen Sie den Stichprobenumfang für ein Konfidenzniveau von $85\ \%$.

(Hinweis: $\Phi(z) = \frac{1+(1-\propto)}{2}$)

c) Welchen Hintergedanken könnten Meinungsforscher haben, wenn sie ihre Umfrageergebnisse publizieren, ohne den Stichprobenumfang anzugeben?

Schätzverfahren

- Lösung:

a) Je höher das Konfidenzniveau desto höher der Stichprobenumfang.

$b)$ $\Phi(z) = \frac{1+(1-\propto)}{2} = \frac{1+0{,}85)}{2} = 0{,}925; \Rightarrow z = 1{,}44; \quad \Rightarrow n \geq= \frac{z^2}{4\varepsilon^2} = \frac{1{,}44^2}{4*0{,}01^2} = 5184$

c) Die Genauigkeit von Umfrageergebnissen lässt sich ohne Angabe des Stichprobenumfangs nicht einschätzen. Leser von Umfragen gehen bei Statistischen Instituten aber unterbewusst von großen Stichprobenumfängen aus, die aber (aus Kostengründen) nicht unbedingt vorliegen müssen.

Testverfahren

- <u>Aufgabe:</u>

Ein Jäger behauptet, dass er höchstens 20% aller Wildenten, auf die er schießt, nicht trifft. Bei der letzten Jagd gab er 15 Schüsse ab und brachte nur 9 Wildenten mit nach Hause.

a) Kann seine Behauptung deshalb als Jägerlatein bezeichnet werden (Signifikanzniveau 5%)?

b) Formulieren und berechnen Sie obige Aufgabe nicht über die „Nicht-Trefferwahrscheinlichkeit", sondern über die Trefferwahrscheinlichkeit.

c) Angenommen der Jäger würde behaupten er hätte eine Trefferwahrscheinlichkeit von 45%. Wie viele Wildenten dürfte der Jäger höchstens bzw. mindestens treffen, damit er bei seiner Behauptung bleiben kann (Signifikanzniveau 5%)?

Testverfahren

- Aufgabe:

  Ein Medikament A heilt eine Krankheit bei 85% der Patienten. Ein konkurrierender Arzneihersteller behauptet, dass sein Medikament B noch besser wirkt, und führt eine Testreihe an 108 Patienten durch.

  Bei mindestens wie vielen Patienten muss das Medikament B die Krankheit heilen, damit man auf einem Signifikanzniveau von 5% bei Medikament B von einer besseren Heilungswirkung als bei Medikament A ausgehen kann.

Testverfahren

- Aufgabe:

  Ein Hersteller liefert Pflastersteine, von denen höchstens 20% zweite Wahl sein sollen. Der Empfänger akzeptiert die Lieferung, falls höchstens 25 Steine zweite Wahl sind. Dazu überprüft er 100 Steine.

  Bei welchem Signifikanzniveau trifft der Empfänger seine Entscheidung?