Variante **A**

Statistik

Technische Universität Dortmund — Fakultät Wirtschaftswissenschaften — 8. Februar 2023

Bitte tragen sie ihre Daten sorgfältig und leserlich ein:

1. Prüfungsversuch ☐ Ja Nein ☐

Matrikelnummer ☐ ☐ ☐ ☐ ☐ ☐ Nachname ______________

Studiengang ______________ Vorname ______________

Bearbeitungshinweise:

Diese Klausur besteht aus 10 Aufgaben. Alle Aufgaben sind unabhängig voneinander lösbar.

Jede Aufgabe hat vier Antwortmöglichkeiten, von denen jeweils genau eine zutreffend ist.

Markieren sie die jeweils zutreffende Antwortmöglichkeit jeder Aufgabe auf diesem Deckblatt. Es werden ausschließlich ihre Markierungen der jeweiligen Antwortmöglichkeiten a) bis d) auf diesem Deckblatt gewertet. Skizzen und Anmerkungen werden nicht bewertet.

Sie dürfen entweder eine oder zwei Antwortmöglichkeiten für jede Aufgabe markieren.

Markieren sie genau eine Antwortmöglichkeit, so erhalten sie bei Markierung der zutreffenden Antwortmöglichkeit drei Punkte für die entsprechende Aufgabe.

Markieren sie genau zwei Antwortmöglichkeiten, so erhalten sie bei Markierung der zutreffenden Antwortmöglichkeit einen Punkt für die entsprechende Aufgabe.

In allen anderen Fällen erhalten sie null Punkte für diese Aufgabe.

Bitte verwenden sie einen Kugelschreiber oder nicht zu starken Filzstift.

Bei 18 von maximal 30 erreichbaren Punkten ist die Klausur in jedem Fall bestanden.

Die Bearbeitungszeit beträgt 60 Minuten.

Viel Erfolg!

Markierung: ☒ Korrektur: ■

Korrekturhinweis: Wenn sie irrtümlich ein falsches Kästchen angekreuzt haben, malen sie dieses bitte vollständig aus und kreuzen sie eindeutig erkennbar die zutreffende Antwort an.

	a)	b)	c)	d)
Aufgabe 1	☐	☐	☐	☐
Aufgabe 2	☐	☐	☐	☐
Aufgabe 3	☐	☐	☐	☐
Aufgabe 4	☐	☐	☐	☐
Aufgabe 5	☐	☐	☐	☐
Aufgabe 6	☐	☐	☐	☐
Aufgabe 7	☐	☐	☐	☐
Aufgabe 8	☐	☐	☐	☐
Aufgabe 9	☐	☐	☐	☐
Aufgabe 10	☐	☐	☐	☐

Aufgabe 1

Zu Silvester wurden Feuerwerkskörper in verschiedenen Verpackungen verkauft. Die folgende Tabelle listet die Anzahl der verkauften Verpackungen zu den jeweiligen Preisen auf:

Verpackungspreis in €	1,00	2,00	2,50	3,00	4,00	5,00
Anzahl der Verkäufe	7	18	29	19	21	6

Wie groß ist der Anteil der verkauften Verpackungen, die 3,00€ oder weniger kosten?

a) **Dieser Anteil beträgt 73%**. (452 Kreuze, 94%)

b) Dieser Anteil beträgt 19%. (13 Kreuze, 3%)

c) Dieser Anteil beträgt 100%. (1 Kreuz, 0%)

d) Dieser Anteil beträgt 27%. (20 Kreuze, 4%)

Im Durchschnitt wurden 1,01 Kreuze gesetzt und 2,81 Punkte erzielt.

Lösungsweg

Es wurden 7+18+29+19=73 Verpackungen verkauft, die 3€ oder weniger kosten.

Insgesamt wurden 7+18+29+19+21+6=100 Verpackungen verkauft.

Daher beträgt der gefragte Anteil 73/100 = 73%.

Aufgabe 2

Eine Stichprobe von elf Personen haben in der Statistikklausur folgende Punktzahlen erreicht:

$$9, 13, 25, 30, 18, 30, 27, 21, 4, 29, 25$$

Wie lauten das arithmetische Mittel $\bar{x}_a$ und der Median $\bar{x}_m$ dieser Stichprobe?

a) $\bar{x}_a = 19,\bar{5}$ und $\bar{x}_m = 25$ (4 Kreuze, 1%)

b) $\mathbf{\bar{x}_a = 21}$ **und** $\mathbf{\bar{x}_m = 25}$ (463 Kreuze, 97%)

c) $\bar{x}_a = 15$ und $\bar{x}_m = 15$ (0 Kreuze, 0%)

d) $\bar{x}_a = 21$ und $\bar{x}_m = 15$ (12 Kreuze, 3%)

Im Durchschnitt wurden 1,00 Kreuze gesetzt und 2,89 Punkte erzielt.

Lösungsweg

Das arithmetische Mittel lautet:

$$\bar{x}_a = \frac{1}{11}\left(9 + 13 + 25 + 30 + 18 + 30 + 27 + 21 + 4 + 29 + 25\right) = \frac{231}{11} = 21$$

Um den Median zu bestimmen, ordnen wir die Stichprobe:

$$4, 9, 13, 18, 21, 25, 25, 27, 29, 30, 30$$

und bestimmen den Wert der Beobachtung, welche an 6. Stelle in der Mitte liegt: $\bar{x}_m = 25$.

Aufgabe 3

Für eine Stichprobe mit n Beobachtungspaaren $(x_i, y_i), i = 1, \ldots, n$ sei die empirische Kovarianz gegeben durch $s_{xy} = 5$.

Betrachten Sie nun die transformierten Daten $z_i = -2 \cdot y_i + 1,\ i = 1, \ldots, n$.

Wie lautet die empirische Kovarianz s_{xz} der Paare (x_i, z_i)?

a) $s_{xz} = -9$ (216 Kreuze, 45%)

b) $s_{xz} = 10$ (71 Kreuze, 15%)

c) $s_{xz} = 9$ (69 Kreuze, 14%)

d) $\mathbf{s_{xz} = -10}$ (199 Kreuze, 42%)

Im Durchschnitt wurden 1,56 Kreuze Gesetzt und 1,08 Punkte erzielt.

Lösungsweg

Die Formel für die empirische Kovarianz s_{xz} der Paare (x_i, z_i) lautet:

$$s_{xz} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_a)(z_i - \bar{z}_a)$$

Wir berechnen zunächst $\bar{z}_a$:

$$\bar{z}_a = \frac{1}{n}\sum_{i=1}^{n} z_i = \frac{1}{n}\sum_{i=1}^{n}(-2 \cdot y_i + 1) = -2\frac{1}{n}\sum_{i=1}^{n} y_i + \frac{1}{n}\sum_{i=1}^{n} 1 = -2 \cdot \bar{y}_a + 1,$$

wobei wir nach dem zweiten Gleichheitszeichen z_i durch $-2 \cdot y_i + 1$ ersetzt haben.

Nun ersetzen wir z_i und $\bar{z}_a$ in der Formel für s_{xz}:

$$\begin{aligned}
s_{xz} &= \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_a)(-2 \cdot y_i + 1 - (-2 \cdot \bar{y}_a + 1)) \\
&= \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_a)(-2 \cdot y_i + 1 + 2 \cdot \bar{y}_a - 1) \\
&= \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_a)(-2 \cdot y_i + 2 \cdot \bar{y}_a) \\
&= -2 \cdot \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_a)(y_i - \cdot \bar{y}_a) \\
&= -2 \cdot s_{xy} = -2 \cdot 5 = -10
\end{aligned}$$

(Siehe auch Formelsammlung unter „Emp. Kovarianz“: $s_{xz} = a * s_{xy}$, hier $a = -2$)

Aufgabe 4

Gegeben Sei ein Warenkorb mit den drei Gütern $i = 1, 2, 3$. Die Preise und Mengen dieser Güter seien für die beiden Perioden $t = 0, 1$ gegeben durch

i	$p_0(i)$	$p_1(i)$	$q_0(i)$	$q_1(i)$
1	15	9	24	20
2	18	12	20	15
3	24	20	15	9

Der Preisindex nach Paasche hat auf Grundlage dieser Daten den Wert $P_{01}^P = 0,687$, wobei auf die dritte Nachkommastelle gerundet wurde.

Welche der folgenden Aussagen gilt für die Preisindizes nach Laspeyeres (P_{01}^L), Paasche (P_{01}^P) und Fischer (P_{01}^F), welche auf Basis dieser Daten berechnet werden?

a) $P_{01}^P < P_{01}^L < P_{01}^F$ (29 Kreuze, 6%)

b) $P_{01}^L < P_{01}^F < P_{01}^P$ (71 Kreuze, 15%)

c) $P_{01}^F < P_{01}^L < P_{01}^P$ (40 Kreuze, 8%)

d) $\mathbf{P_{01}^P < P_{01}^F < P_{01}^L}$ (373 Kreuze, 78%)

Im Durchschnitt wurden 1,07 Kreuze gesetzt und 2,29 Punkte erzielt.

Lösungsweg

In der Vorlesung wurde argumentiert, dass der Preisindex nach Fischer immer zwischen den beiden anderen Preisindizes liegen muss. Daher entfallen Antwortmöglichkeiten a) und c) und wir müssen nur noch den Preisindex nach Laspeyres berechnen (siehe Formelsammlung):

$$P_{01}^L = \frac{\sum_{i=1}^3 p_1(i) \cdot q_0(i)}{\sum_{i=1}^3 p_0(i) \cdot q_0(i)} = \frac{9 \cdot 24 + 12 \cdot 20 + 20 \cdot 15}{15 \cdot 24 + 18 \cdot 20 + 24 \cdot 15} = \frac{216 + 240 + 300}{360 + 360 + 360} = \frac{746}{1080} = 0,691$$

(Es gilt nun $P_{01}^F = \sqrt{P_{01}^P \cdot P_{01}^L} = 0,689$.)

Demnach gilt $P_{01}^P = 0,687 < P_{01}^F = 0,689 < P_{01}^L = 0,691$.

Aufgabe 5

In der Vorbereitung auf die Klausur konnten Studierende Vorschläge für die Formelsammlung einreichen.

Die folgende Tabelle listet einige Angaben zur Abgabe eines Vorschlags und zum Geschlecht der Person auf:[1]

	weiblich	männlich
Vorschlag eingereicht	10	4
ohne Vorschlag	18	X

Betrachten Sie die Hypothese, dass das Geschlecht einer Person keinen Einfluss darauf hat, ob diese Person einen Vorschlag einreicht.

Welchen Wert würden Sie für X erwarten, falls diese Hypothese stimmt?
(Für welchen Wert X wäre der Kontingenzkoeffizient nahe null?)

Runden Sie dabei auf eine ganze Zahl.

a) $\mathbf{X = 7}$ (311 Kreuze, 65%)

b) $X = 24$ (79 Kreuze, 16%)

c) $X = 9$ (64 Kreuze, 13%)

d) $X = 13$ (72 Kreuze, 15%)

Im Durchschnitt wurden 1,10 Kreuze gesetzt und 1,82 Punkte erzielt.

Lösungsweg

Wenn das Geschlecht einer Person keinen Einfluss darauf hat, ob diese Person einen Vorschlag einreicht, dann sollte das Verhältnis von weiblich zu männlich bei den Personen, die einen Vorschlag eingereicht haben ungefähr gleich sein zu dem Verhältnis von weiblich zu männlich bei den Personen, die keinen Vorschlag eingereicht haben. Dies bedeutet, dass

$$\frac{10}{4} \approx \frac{18}{X} \Leftrightarrow X \approx 18 \cdot \frac{4}{10} = 7,2$$

Alternativ kann unter der gestellten Hypothese argumentiert werden, dass das Verhältnis der Personen mit Vorschlag zu Personen ohne Vorschlag bei beiden Geschlechtern ähnlich sein sollte. Dies bedeutet, dass

$$\frac{10}{18} \approx \frac{4}{X} \ ,$$

was zum gleichen Ergebnis führt.

Die Berechnung des Kontingenzkoeffizienten mittels $\tilde{h}_{ij}$ und χ^2 ist zu aufwendig und würde zu viel Zeit beanspruchen.

Für $X = 7$ ergibt sich der Wert $K = 0,01$.

[1] Daten zum Geschlecht laut Lehrevaluation, die Option „divers“ wurde dort von null Personen gewählt.

Aufgabe 6

Betrachten Sie die unabhängigen Würfe eines roten und eines blauen Würfels mit je vier Seiten mit den Zahlen eins bis vier. Alle Seiten der Würfel fallen mit gleicher Wahrscheinlichkeit. Der rote Würfel bestimmt die zehner-Stelle und der blaue Würfel die einer-Stelle einer Zufallszahl X mit den Werten zwischen 11 und 44.

Wie groß ist die Wahrscheinlichkeit, dass eine Zahl $X \leq 33$ gewürfelt wird?

a) $\mathbb{P}(X \leq 33) = \frac{33}{16}$ (43 Kreuze, 9%)

b) $\mathbb{P}(X \leq 33) = \frac{1}{2}$ (64 Kreuze, 13%)

c) $\mathbb{P}(\mathbf{X} \leq \mathbf{33}) = \mathbf{\frac{11}{16}}$ (293 Kreuze, 61%)

d) $\mathbb{P}(X \leq 33) = \frac{3}{8}$ (124 Kreuze, 26%)

Im Durchschnitt wurden 1,09 Kreuze gesetzt und 1,69 Punkte erzielt.

Lösungsweg

Die folgende Tabelle listet die Würfelergebnisse und den Wert der Zufallsvariable X auf, wobei alle Zahlen kleiner oder gleich 33 fett gedruckt sind:

	1	2	3	4
1	**11**	**12**	**13**	**14**
2	**22**	**22**	**23**	**24**
3	**31**	**32**	**33**	34
4	41	42	43	44

Es gibt also 11 Werte von X, die kleiner oder gleich 33 sind. Insgesamt gibt es $4 \cdot 4 = 16$ mögliche Werte von X. Alle Werte haben die gleiche Wahrscheinlichkeit. Also beträgt die Wahrscheinlichkeit, dass $X \leq 33$:

$$\mathbb{P}(X \leq 33) = \frac{11}{16}$$

Aufgabe 7

Rainer Z. kreuzt jeweils mit gleicher Wahrscheinlichkeit und unabhängig bei jeder der zehn Aufgaben dieser Klausur eine der vier Antwortmöglichkeiten an.

Für jedes zutreffende Kreuz erhält Rainer drei Punkte.

Welche der folgenden Aussagen in Bezug auf die Summe der Punkte Y, die Rainer in dieser Klausur erzielt ist richtig?

a) $\mathbf{Y = 3X}$**, wobei** $\mathbf{X \sim \mathcal{B}\left(10, \frac{1}{4}\right)}$ (275 Kreuze, 57%)

b) $Y \sim \mathcal{N}\left(\frac{30}{4}, \frac{3}{16}\right)$ (64 Kreuze, 13%)

c) $Y = 30X$, wobei X Bernoulli-verteilt mit $p = \frac{1}{4}$ (191 Kreuze, 40%)

d) $Y \sim \mathcal{P}\left(\frac{30}{4}\right)$ (54 Kreuze, 11%)

Im Durchschnitt wurden 1,22 Kreuze gesetzt und 1,40 Punkte erzielt.

Lösungsweg

Seien X_i, $i = 1, \ldots, 10$, unabhängige Bernoulli-Variablen, welche je mit Wahrscheinlichkeit $\frac{1}{4}$ den Wert 1 haben und mit Wahrscheinlichkeit $\frac{3}{4}$ den Wert 0. Interpretation: Kreuzt Rainer in Aufgabe $i = 1, \ldots, 10$ die zutreffende Antwort an, gilt $X_i = 1$ und andernfalls gilt $X_i = 0$.

Die Summe aller zehn Zufallsvariablen X_i, also $X = \sum_{i=1}^{10} X_i$, ist dann binomialverteilt mit den Parametern $n = 10$ und $p = \frac{1}{4}$, also $X \sim \mathcal{B}(10, \frac{1}{4})$. Interpretation: X zeigt die Anzahl der zutreffend geratenen Aufgaben an.

Pro zutreffendem Kreuz gibt es drei Punkte, also ist die Gesamtpunktzahl Y gegeben durch:

$$Y = 3X \sim \mathcal{B}(10, \frac{1}{4})$$

Aufgabe 8

Es sei die Zufallsvariable X gleichverteilt auf dem Intervall $\left[\frac{1}{3}, \frac{2}{3}\right]$.

Für welches $x_{4/5}$ gilt $\mathbb{P}(X \leq x_{4/5}) = \frac{4}{5}$? Berechnen Sie für $\alpha = \frac{4}{5}$ das α-Quantil $x_{4/5}$!

a) $x_{4/5} = \frac{2}{5}$ (104 Kreuze, 22%)

b) $x_{4/5} = \frac{1}{5}$ (109 Kreuze, 23%)

c) $x_{4/5} = \frac{4}{5}$ (129 Kreuze, 27%)

d) $\mathbf{x_{4/5} = \frac{3}{5}}$ (231 Kreuze, 48%)

Im Durchschnitt wurden 1,20 Kreuze gesetzt und 1,21 Punkte erzielt.

Lösungsweg

Um die Aufgabe zu beantworten müssen wir die Verteilungsfunktion $F(x) = \mathbb{P}(X \leq x)$ bestimmen.

Bei gleichverteilten stetigen Zufallsvariablen hat die Verteilungsfunktion eine konstante Steigung auf dem Wertebereich der Zufallsvariablen, sie hat also die Form $F(x) = ax + b$.

Die Wahrscheinlichkeit, dass X einen Wert kleiner oder gleich $\frac{1}{3}$ annimmt, ist 0 und die Wahrscheinlichkeit, dass X einen Wert kleiner oder gleich $\frac{2}{3}$ annimmt, ist 1.

Demnach gilt $F(\frac{1}{3}) = a \cdot \frac{1}{3} + b = 0$ und $F(\frac{2}{3}) = a \cdot \frac{2}{3} + b = 1$.

Die erste Gleichung nach b umgeformt ergibt $b = -a\frac{1}{3}$.
Eingesetzt in die zweite Gleichung ergibt $a \cdot \frac{2}{3} - a \cdot \frac{1}{3} = 1 \Leftrightarrow a = 3$. Damit ist dann $b = -3\frac{1}{3} = -1$.

Es gilt also

$$F(x) = \begin{cases} 0 & \text{, falls } x < \frac{1}{3} \\ 3x - 1 & \text{, falls } \frac{1}{3} \leq x \leq \frac{2}{3} \\ 1 & \text{, falls } x > \frac{2}{3} \end{cases}$$

Gesucht ist nun der Wert $x_{4/5}$ mit $F(x_{4/5}) = \frac{4}{5}$. Wir lösen also die Gleichung

$$3 \cdot x_{4/5} - 1 = \frac{4}{5} \Leftrightarrow 3 \cdot x_{4/5} = 1 + \frac{4}{5} = \frac{9}{5} \Leftrightarrow x_{4/5} = \frac{3}{5}$$

Der kurze Lösungsweg für ein beliebiges Quantil x_α bei gleichverteilter Zufallsvariable auf dem Intervall $[\underline{x}, \overline{x}]$ lautet:

$$x_\alpha = (1 - \alpha) \cdot \underline{x} + \alpha \cdot \overline{x}$$

Hier mit $\alpha = \frac{4}{5}$, $\underline{x} = \frac{1}{3}$ und $\overline{x} = \frac{2}{3}$:

$$x_{4/5} = \frac{1}{5} \cdot \frac{1}{3} + \frac{4}{5} \cdot \frac{2}{3} = \frac{1}{15} + \frac{8}{15} = \frac{9}{15} = \frac{3}{5}$$

Aufgabe 9

Die Zufallsvariable X sei normalverteilt mit Erwartungswert $\mu = -3$ und Varianz $\sigma^2 = 9$.

Bestimmen Sie $\mathbb{P}(X \leq 0)$!

Sie finden eine Tabelle der Wahrscheinlichkeiten der Standardnormalverteilung auf der letzen Seite dieser Klausur.

a) $\mathbb{P}(\mathbf{X \leq 0}) \approx \mathbf{84}\%$ (194 Kreuze, 41%)

b) $\mathbb{P}(X \leq 0) \approx 0,8413\%$ (175 Kreuze, 37%)

c) $\mathbb{P}(X \leq 0) \approx 50\%$ (102 Kreuze, 21%)

d) $\mathbb{P}(X \leq 0) \approx 16\%$ (91 Kreuze, 19%)

Im Durchschnitt wurden 1,17 Kreuze gesetzt und 1,04 Punkte erzielt.

Lösungsweg

Um die Tabelle im Anhang nutzen zu können, müssen wir die Zufallsvariable standardisieren, wir müssen also ihren Erwartungswert subtrahieren und danach durch ihre Standardabweichung dividieren.

Es gilt:

$$X \leq 0 \Leftrightarrow X - (-3) \leq 0 - (-3) \Leftrightarrow X + 3 \leq 3 \Leftrightarrow \frac{X+3}{\sqrt{9}} \leq \frac{3}{\sqrt{9}} \Leftrightarrow \frac{X+3}{3} \leq 1$$

Die Zufallsvariable $\frac{X+3}{3}$ ist nun standardnormalverteilt und es gilt:

$$\mathbb{P}\left(\frac{X+3}{3} \leq 1\right) = \Phi(1)$$

Wir können nun in der Tabelle im Anhang ablesen, welchen Wert die Verteilungsfunktion der Standardnormalverteilung Φ an der Stelle 1 hat. Hierzu suchen wir die Zeile, welche in der Spalte z den Wert 1 hat und lesen in dieser Zeile in der Spalte mit der Bezeichnung „0" den Wert 0,8413 ab.

Es gilt also $\Phi(1) = 0,8413$. Mit Wahrscheinlichkeit 0,8413, also mit ungefähr 84%, nimmt die Zufallsvariable X einen Wert kleiner oder gleich null an.

Anmerkung: der Wert $0,8413\%$ entspricht der Dezimalzahl 0,008413.

Aufgabe 10

Gegeben seien zwei diskrete Zufallsvariablen X und Y, welche gemeinsam der in der Tabelle angegebenen Wahrscheinlichkeitsverteilung folgen:

		y_1	y_2	y_3
		-1	0	1
x_1	0	0	1/3	0
x_2	1	1/3	0	1/3

Die Zelle der Zeile x_i und Spalte y_j gibt dabei die Wahrscheinlichkeit an, dass X den Wert x_i und Y den Wert x_j annimmt: $f_{X,Y}(x_i, y_j)$.

Welche der folgenden Aussagen ist wahr?

a) Die Zufallsvariablen X und Y sind unabhängig und $cov(X,Y) = 0$. (170 Kreuze, 35%)

b) Die Zufallsvariablen X und Y sind nicht unabhängig und $cov(X,Y) \neq 0$. (160 Kreuze, 33%)

c) **Die Zufallsvariablen X und Y sind nicht unabhängig und $\mathbf{cov(X,Y) = 0}$.** (207 Kreuze, 43%)

d) Die Zufallsvariablen X und Y sind unabhängig und $cov(X,Y) \neq 0$. (100 Kreuze, 21%)

Im Durchschnitt wurden 1,33 Kreuze gesetzt und 0,90 Punkte erzielt.

Lösungsweg

Die Kovarianz von X und Y ist definiert durch

$$cov(X,Y) = \mathbb{E}[X \cdot Y] - \mathbb{E}[X] \cdot \mathbb{E}[Y]$$

Es gilt

$$\begin{aligned}
\mathbb{E}[X \cdot Y] &= 0 \cdot (-1) \cdot f_{X,Y}(0,-1) + 0 \cdot 0 \cdot f_{X,Y}(0,0) + 0 \cdot 1 \cdot f_{X,Y}(0,1) \\
&+ 1 \cdot (-1) \cdot f_{X,Y}(1,-1) + 1 \cdot 0 \cdot f_{X,Y}(1,0) + 1 \cdot 1 \cdot f_{X,Y}(1,1) \\
&= 1 \cdot (-1) \cdot \underbrace{f_{X,Y}(1,-1)}_{\frac{1}{3}} + 1 \cdot 1 \cdot \underbrace{f_{X,Y}(1,1)}_{\frac{1}{3}} - \frac{1}{3} + \frac{1}{3} = 0 \\
\mathbb{E}[Y] &= -1 \cdot (\underbrace{f_{X,Y}(0,-1)}_{0} + f_{X,Y}(1,-1)) + 0 \cdot (f_{X,Y}(0,0) + f_{X,Y}(1,0)) + 1 \cdot (\underbrace{f_{X,Y}(0,1)}_{0} + f_{X,Y}(1,1)) \\
&= -1 \cdot \underbrace{f_{X,Y}(1,-1)}_{\frac{1}{3}} + 1 \cdot \underbrace{f_{X,Y}(1,1)}_{\frac{1}{3}} = -\frac{1}{3} + \frac{1}{3} = 0
\end{aligned}$$

Also gilt $cov(X,Y) = 0 - \mathbb{E}[X] \cdot 0 = 0$.

Wir prüfen nun die Unabhängigkeit der beiden Zufallsvariablen und berechnen hierfür $\mathbb{P}(X=0, Y=-1)$ und $\mathbb{P}(X=0) \cdot \mathbb{P}(Y=-1)$. Falls X und Y unabhängig sind, müssen diese beiden Werte übereinstimmen.

$$\begin{aligned}
\mathbb{P}(X=0, Y=-1) &= f_{X,Y}(0,-1) = 0 \\
\mathbb{P}(X=0) &= f_{X,Y}(0,-1) + f_{X,Y}(0,0) + f_{X,Y}(0,1) = 0 + \frac{1}{3} + 0 = \frac{1}{3} \\
\mathbb{P}(Y=-1) &= f_{X,Y}(0,-1) + f_{X,Y}(1,-1) = 0 + \frac{1}{3} = \frac{1}{3}
\end{aligned}$$

Demnach gilt $\mathbb{P}(X=0, Y=-1) = 0 \neq \frac{1}{3} \cdot \frac{1}{3} = \mathbb{P}(X=0) \cdot \mathbb{P}(Y=-1)$ und die beiden Zufallsvariablen X und Y sind nicht unabhängig.

Formelsammlung **deskriptive Statistik**; Empirische Verteilungsfunktion $F_n(x) = \sum_{i=1}^{n} h(a_i) \leq$ x Anteil Beobachtungen, die höchstens den Wert x haben

Lageparameter; seien $x_1, x_2, \dots, x_n$ Ausprägungen eines metrischen, ordinalen oder quantitativen Merkmals; Wachstumsfaktor=neu/alt; Wachstumsrate=$(neu - alt)/alt$

Arithmetisches Mittel $\bar{x}_a = \frac{1}{n}\sum_{i=1}^{n} x_i$
Gewichtetes Mittel $\bar{x}_a^g = \sum_{i=1}^{n} g_i x_i$ mit $g_i \geq 0$
Harmonisches Mittel $\bar{x}_h = \frac{n}{\frac{1}{x_1}+\dots+\frac{1}{x_n}}$ mit $x_i > 0$
Geometrisches Mittel $\bar{x}_g = \sqrt[n]{x_1 * x_2 * \dots * x_n}$

Es gilt: $\bar{x}_g \leq \bar{x}_a$ und $\bar{x}_g = \bar{x}_a \Leftrightarrow x_1 = x_2 = \dots = x_n$

Kontingenztabelle h_{ij} abs. Häufigkeit
Koeffizient: $\tilde{h}_{ij} = \frac{h_i * h_j}{n}$
Abhängigk.: $x^2 = \sum_{i=1}^{m(i)} \sum_{j=1}^{m(j)} \frac{(h_{ij} - \tilde{h}_{ij})^2}{\tilde{h}_{ij}}$
Kontingenz.: $K = \sqrt{\frac{x^2}{n + x^2}}$ $0 \leq K \leq \sqrt{\frac{M-1}{M}}$
$x^2 = 0, K = 0 \Leftrightarrow h_{21} * h_{12} = h_{11} * h_{22}$

$x_{0,5} =$ Median $\bar{x}_m = \begin{cases} x_{(\frac{n+1}{2})} \text{ falls } n \text{ ungerade} \\ \frac{1}{2}\left(x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}\right) \text{ falls } n \text{ gerade} \end{cases}$
wobei $x_{(i)}$ der i-te Wert der aufsteigend geordneten Daten ist

Ges. Wachstumsrate
r=$(1 + r_m)^p - 1$
Allg. Wachstumsrate
$K_n = K_{n-1}(1 + r_n)$

Modalwert $\bar{x}_{mod}$ tritt am häufigsten in einer geordneten Stichprobe auf

Lineare Transformation
Arithmetisches Mittel: Seien X,Y,Z metrische Merkmale mit x_i, y_i, z_i und i = 1,…,n und $z_i = cx_i + dy_i + e \Rightarrow \bar{z}_a = c\bar{x}_a + d\bar{y}_a + e$
Median: Seien X,Y metrische Merkmale mit x_i, y_i und i = 1,…,n und $y_i = cx_i + d \Rightarrow \bar{y}_m = c\bar{x}_m + d$

$2^x = 8 \Leftrightarrow x = log_2 8$
absolute Häufigkeit $H(a_j) \Rightarrow Höhe = h(a_i)/Breite$
relative Häufigkeit $h(a_j) = \frac{1}{n} H(a_j)$

Optimalitätseigenschaft von $\bar{x}_a$ und $\bar{x}_m$
Das arithmetische Mittel minimiert die Summe der quadrierten Abweichungen von $x_1, \dots, x_n$ von einer Zahl: $\bar{x}_a = argmin_{z \in \mathbb{R}} \sum_{i=1}^{n} (x_i - z)^2$
Der Median minimiert die Summe der absoluten Abweichungen von $x_1, \dots, x_n$ von einer Zahl: $\bar{x}_m = argmin_{z \in \mathbb{R}} \sum_{i=1}^{n} |x_i - z|$

Jeder Wert x_p mit $0 < p < 1$ und $x_{(\lceil np \rceil - 1)} \leq x_p \leq x_{(\lceil np \rceil)}$ heißt p-Quantil. $x_p = \begin{cases} x_{(\lceil np \rceil)} \text{ falls } np \notin \mathbb{Z} \\ \frac{1}{2}\left(x_{(np)} + x_{(np+1)}\right) \text{ falls } np \in \mathbb{Z} \end{cases}$, es gilt: $F_n(x_p) \leq p$
$x_{0,25} = unteres\ Quartil$ (25%); $x_{0,75} = oberes\ Quartil$ (75%), $\lceil x \rceil = min\{k \in \mathbb{Z} | k \geq x\}$, $[a, b)\ \{x | a \leq x < b\}$

Streuungsmaße; $y_i = ax_i + b, i = 1, \dots, n \wedge a, b \in \mathbb{R} \Leftrightarrow R_y = |a| R_x;\ d_y = |a| d_x;\ \Delta_y = |a| \Delta_x;\ s_y^2 = a^2 s_x^2 \Leftrightarrow s_y = |a| s_x$ analog Kovarianz; Lorenzkurve$(i/n;\ \sum_{j=1}^{i} x_j / \sum_{l=1}^{n} x_l)$

Spannweite $R_x = max_i\, x_i - min_i\, x_i = x_n - x_1$
Mittlere absolute Abweichung $d_x = \frac{1}{n}\sum_{i=1}^{n} |x_i - \bar{x}_m|$
Mittlere absolute Diff $\Delta_x = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} |x_i - x_j|$

Herfindahl $H_x = \sum_{i=1}^{n} \left(\frac{x_i}{\sum_{j=1}^{n} x_j}\right)^2$
$\frac{1}{n} \leq H_x \leq 1 \wedge H_x = \frac{1}{n} \Leftrightarrow$
$alle\ x_i\ gleich$

Emp. Varianz (mittlere quad. Abweichung) $s_x^2 = \frac{1}{n}\sum_{n=1}^{n} (x_i - \bar{x}_a)^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}_a^2$
Standardabweichung $s_x = \sqrt{s_x^2} = \sqrt{\frac{1}{n}\sum_{n=1}^{n} (x_i - \bar{x}_a)^2}$
Gini-Koeffizient $G_x = \Delta_x / 2\bar{x}_a$ und $0 \leq G_x \leq ((n-1)/n) < 1$ bzw.
$G_x = \frac{2}{n * \sum_{i=1}^{n} x_i} * \left(\sum_{i=1}^{n} x_i * i\right) - 1 - \frac{1}{n}$

Preis -&Mengenindizes
Laspeyres $P_{0t}^L = \frac{\sum_{i=1}^{n} p_t(i) * q_0(i)}{\sum_{i=1}^{n} p_0(i) * q_0(i)}$; $Q_{0t}^L = \frac{\sum_{i=1}^{n} p_0(i) * q_t(i)}{\sum_{i=1}^{n} p_0(i) * q_0(i)}$
Paasche $P_{0t}^P = \frac{\sum_{i=1}^{n} p_t(i) * q_t(i)}{\sum_{i=1}^{n} p_0(i) * q_t(i)}$; $Q_{0t}^P = \frac{\sum_{i=1}^{n} p_t(i) * q_t(i)}{\sum_{i=1}^{n} p_t(i) * q_0(i)}$
Fisher $P_{0t}^F = \sqrt{P_{0t}^L * P_{0t}^P}$; $Q_{0t}^F = \sqrt{Q_{0t}^L * Q_{0t}^P}$

$p_o(i)$ Preis i.d. Basisperiode 0
$q_0(i)$ Menge i.d. Basisperiode 0
$p_t(i)$ Preis i.d. Periode t
$q_t(i)$ Menge i.d. Berichtsperiode t

Emp. Kovarianz $s_{xy} = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x}_a)(y_i - \bar{y}_a) = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}_a \bar{y}_a \wedge s_{xz} = a * s_{xy}$
Korrelation $r_{xy} = \frac{s_{xy}}{s_x * s_y}$ linearer Zusammenhang zwischen X und Y
Rangkorrelation $r_{xy}^{SP} = \frac{\sum_{i=1}^{n} (rk(x_i) - \frac{1+n}{2})(rk(y_i) - \frac{1+n}{2})}{\sqrt{\sum_{i=1}^{n} (rk(x_i) - \frac{1+n}{2})^2} * \sqrt{\sum_{i=1}^{n} (rk(y_i) - \frac{1+n}{2})^2}}$ $r_{xy}^{SP} = 1 - \frac{6\sum_{i=1}^{n} (rk(x_i) - rk(y_i))^2}{n(n-1)(n+1)}$

Wahrscheinlichkeitsrechnung; $\Omega = \{\omega_1, \dots, \omega_n\}$ Ergebnismenge, $\{\omega_i\} \in \Omega$ Ereignisse
$\Omega = \{(1,1)\}$ Zahlenpaar, $\Omega = \{1, \dots, n\} x \{1, \dots, n\}$ kartes. Produkt, $|\Omega|$=Anzahl Elemente
Zufallsexperiment mit k Ereignissen wird n-mal wiederholt$\Rightarrow k^n$ Ereignisse

Laplace: alle Ereignisse gleiche Wahrscheinlichkeit $\frac{1}{|\Omega|}$, A ein Ereignis $\Rightarrow P(A) = \frac{|A|}{|\Omega|}$
$P(\Omega) = 1; P(\emptyset) = 0; P(\bar{A}) = 1 - P(A);\ P(A \cup B) = P(A) + P(B) - P(A \cap B)$
$P(B|A) = \frac{P(A \cap B)}{P(A)} \triangleq$ bedingte Wahrscheinlichkeit von B, gegeben A
$P(B|A) = P(B)\ und\ P(A|B) = P(A)$ stoch unabhängig $\Leftrightarrow P(A \cap B) = P(A) * P(B)$
Formel v Bayes Sei A_i Partition von Ω dann $P(A_j|B) = \frac{P(B|A_j) * P(A_j)}{\sum_{i=1}^{m} P(B|A_i) * P(A_i)}$ (a-posteriori Wslkeit)

Schnittmenge $A \cap B \triangleq \omega \in \Omega$ sowohl in A als auch in B
Vereinigungsmenge $A \cup B \triangleq \omega \in \Omega$ in A oder B
Komplementärmenge $\bar{A}$ von $A \triangleq \omega \in \Omega$, die nicht in A liegen
Differenzmenge $A \setminus B \triangleq \omega \in \Omega$, die in A aber nicht in B liegen
A und B disjunkt $\Leftrightarrow A \cap B = \emptyset$
Potenzmenge $\rho(A) \triangleq$ Menge aller Teilmengen von A
Binomialkoeffizient $\binom{n}{k} = \frac{n!}{(n-k)!k!}$ wieviele Teilmengen von Ω mit genau k Elementen

Vollst. Wslkeit: $A_1, \dots, A_n$ eine Partition von $\Omega \Rightarrow A_i \cap A_j = \emptyset$ und $A_1 \cup \dots A_m = \Omega$ dann
$P(B) = \sum_{i=1}^{m} P(B|A_i) * P(A_i)$

Ω Ergebnismenge eines Zufallsexperiment und $n \in \mathbb{N} \Rightarrow X: \Omega \to \mathbb{R}^n, \omega \to X(\omega)$; wobei X(ω) Realisation einer Zufallsvariablen
diskret$\Rightarrow$ nur endlich oder abzählbar unendlich(=|ℕ|) viele Werte annehmen
stetig$\Rightarrow$überabzählbar viele Werte annehmen
$f(x) = \mathbb{P}(X = x)$ heißt Wahrscheinlichkeitsfunktion von X
$F(x) = \mathbb{P}(X \leq x)$ heißt Verteilungsfunktion von X $F(x) = \sum_{i \in \mathcal{J}: x_i \leq x} f(x_i)$; $\mathbb{P}(a < X \leq b) = F(b) - F(a); 0 \leq F(x) \leq 1 \forall x \in \mathbb{R};\ \lim_{x \to -\infty} F(X) = 0\ und\ \lim_{x \to \infty} F(X) = 1$
1. Ableitung der Verteilungsfunktion heißt Dichtefunktion f(x)=F`(X); F(X) = $\int_{-\infty}^{x} f(t)dt$; $\mathbb{P}(a < X \leq b) = F(b) - F(a) = \int_a^b f(t)dt$; $\lim_{n \to \infty} f(x) = \lim_{n \to -\infty} f(x) = 0$; $\int_{-\infty}^{\infty} f(x)dx = 1$

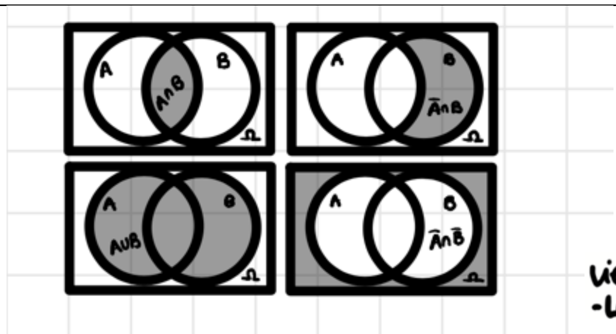
Erwartungswert von Zufallsvariable X heißt $\mathbb{E}[X] = \sum_{i \in \mathcal{J}} x_i f(x_i) \Leftrightarrow$ X diskret und f die Wahrscheinlichkeitsfunktion und $\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) dx \Leftrightarrow$ X stetig und f die Dichtefunktion
Varianz von X heißt $\sigma^2 = \sum_{i \in \mathcal{J}} (x_i - \mathbb{E}(X))^2 f(x_i) \Leftrightarrow$ X diskret und f die Wahrscheinlichkeitsfunktion und $\sigma^2 = \int_{-\infty}^{\infty} (x - \mathbb{E}(x))^2 f(x) dx \Leftrightarrow$ X stetig und f die Dichtefunktion
$\sigma^2 = \mathbb{E}(x^2) - \mathbb{E}(x)^2; \mathbb{E}[XY] = \mathbb{E}[X] * \mathbb{E}[Y]$

Gleichverteilung, wenn Intervalle gleicher Länge gleich wahrscheinlich sind: $\mathbb{E}(x) = \frac{a+b}{2}, \sigma^2 = \frac{(b-a)^2}{12}$; Dichtefunktion $f(x) = \begin{cases} 0 \text{ falls } x < a \text{ oder } x > b \\ \frac{1}{b-a} \text{ falls } a \leq x \leq b \end{cases}$
Verteilungsfunktion $F(x) = \begin{cases} 0 \text{ falls } x < a \\ \frac{x-a}{b-a} \text{ falls } a \leq x \leq b \\ 1 \text{ falls } x > b \end{cases}$ Quantile: $x_\alpha = (1 - \alpha) * a + \alpha * b; F(X_\alpha) = \alpha$
Normalverteilung, Dichte: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right)$; $\mathbb{E}(x) = \mu$; $var(x) = \sigma^2$ Standardnormalverteilung, Dichte: $\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$;
Affine Transformation standardnorm. Zufallsvariablen $\mathbb{E}(x) = a; var(x) = b^2$; Quantile normalverteilter Zufallsvariablen $F_x(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$ und α-Quantil $z_\alpha = \frac{x_\alpha - \mu}{\sigma}$ für $X \sim \mathcal{N}(\mu, \sigma^2)$
Zentraler Grenzwertsatz $\lim_{n \to \infty} \mathbb{P}\left(\frac{\bar{x}_n - \mu}{\sigma/\sqrt{n}} \leq x\right) = \Phi(x)$; $\mathbb{E}[\bar{X}_n] = \mu$; $var(\bar{X}_n) = \sigma^2/n$

Kovarianz $\sigma_{XY} = cov(X, Y) = \mathbb{E}((X - \mathbb{E}(x))(Y - \mathbb{E}(Y)))$
Korrelation $\rho_{XY} = corr(X, Y) = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$; $-1 \leq corr(X, Y) \leq 1$
Bedingter erwartungswert von g(y) gegeben X=x definiert als $\mathbb{E}[g(y)|X = x] = \sum_{i \in \mathcal{J}} g(y_i) * f_{Y|X}(y_i|x)$
Bedingte Varianz $var(Y|X = x) = \sum_{i \in \mathcal{J}} (y_i - \mathbb{E}[Y])^2 * f_{y|X}(y|x)$

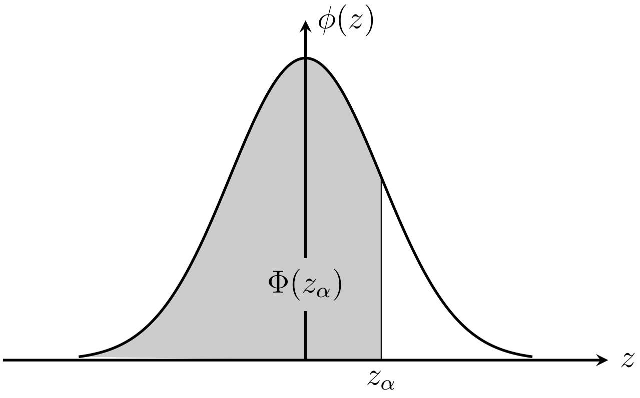
$A \cup B$ ist Komplementärmenge zu $\bar{A} \cap \bar{B} = \overline{A \cup B}$
$A \cap B$ ist Komplementärmenge zu $\bar{A} \cup \bar{B} = \overline{A \cap B}$
$P(A) = P(A \cap B) + P(A \cap \bar{B}) = P(B) * P(A|B) + P(\bar{B}) * P(A|\bar{B})$
$A \cap \bar{B}$ = in A, aber nicht in B
$\bar{A} \cap \bar{B}$ = nicht in A und nicht in B

Binomialverteilung: $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$; $\mathbb{E}[X] = n * p$; $var(X) = np(1-p)$
Poisson: $\mathbb{P}(X = k) = \begin{cases} (\lambda^k / k!) * \exp(-\lambda) \text{ falls k=0,1,2,...} \\ 0 \end{cases}$; $\mathbb{E}[X] = var(X) = \lambda$



viel Erfolg!
-Laura

Wahrscheinlichkeiten der Standardnormalverteilung $\mathcal{N}(0,1)$



z	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	0,5000	0,504	0,508	0,512	0,516	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986