

Biometrie und Methodik (Statistik) - WiSem08/09

2. Übungsblatt: Darstellende Statistik

Aufgabe 1. Geben Sie Mittelwert und Standardabweichung für die stat. Gesamtheit der folgenden Urlisten an: a) 7,7,7,7,7,7; b) 7,5,5,7,5,7,7,5; c) 2.5,-1.5,-1.5,-1.5,2.5,-1.5,2.5,2.5; d) 3,2,1,2,1,3,2,3,2,1,2,2. (Möglichst ohne Taschenrechner!)

Aufgabe 2. Prüfen Sie durch Bestimmung der bedingten empirischen Verteilung, bei welcher der Tabellen die Merkmale x und y abhängig sind:

a)

Merkmal	a_1	a_2	a_3	$n(b_j)$
b_1	4	6	10	20
b_2	5	7	8	20
b_3	11	7	2	20
$n(a_i)$	20	20	20	60

b)

Merkmal	a_1	a_2	a_3	$n(b_j)$
b_1	2	5	3	10
b_2	6	15	9	30
b_3	4	10	6	20
$n(a_i)$	12	30	18	60

Aufgabe 3. In einer Untersuchung über die Dauer von Arbeitslosigkeit von Jugendlichen werden zufällig 100 Jugendliche befragt. Das Ergebnis ist die folgende Kontingenztabelle (abs. Häufigkeiten):

Ausbildung/Dauer d. Arbeitslosigk.	Mittelfristig	Langfristig	Summe
Keine Ausbildung	19	18	37
Lehre	43	20	63
Summe	62	38	100

- a) Beurteilen Sie zunächst ohne Rechnung, ob das Merkmal "Ausbildung" mit dem Merkmal "Dauer der Arbeitslosigkeit" zusammenhängt bzw.unabhängig davon ist.
- b) Bestätigen Sie dann Ihren Eindruck, indem Sie das Cramérsche Kontingenzmaß berechnen. Beurteilen Sie die Stärke des Zusammenhangs zwischen den Merkmalen.

Aufgabe 4. Bei der folgenden Kontingenztabelle ist das Merkmal X maximal statistisch abhängig vom Merkmal Y . Trotzdem sind X und Y unkorreliert. Überprüfen Sie diese Aussage, indem Sie die empirische Kovarianz berechnen:

X/Y	$b_1 = 1$	$b_2 = 2$	$b_3 = 3$	Summe
$a_1 = 10$	10	0	10	20
$a_2 = 20$	0	10	0	10
Summe	10	10	10	30

Aufgabe 5. Bei fünf zufällig ausgewählten Kleinkindern wird der Wortschatz X und die Körpergröße Y in cm gemessen. Das Maß für X sei die Anzahl der verwendeten Wörter in einem Aufsatz über die Sommerferien. Zudem wird das Alter der Kinder Z (in Jahren) bestimmt. Die Urliste, die Sie erhalten sei:

i	1	2	3	4	5
x_i	37	30	20	28	35
y_i	130	112	108	114	136
z_i	12	7	6	7	13

- a) Zeichnen Sie zunächst das Streudiagramm von X gegen Y und berechnen Sie den Korrelationskoeffizienten $\rho_{X,Y}$.
- b) Soll man nun glauben, dass Körpergröße und Wortschatz in einem kausalen Zusammenhang stehen? Eher nicht. Um das zu belegen, berechnen Sie die Korrelationskoeffizienten $\rho_{X,Z}$ sowie $\rho_{Y,Z}$ und begründen Sie, wie die starke Korrelation zwischen X und Y vermutlich zustande kommt.

KLAUSURAUFGABEN

Klausur-Aufgabe 1. Besteht das folgende Objekt aus **statistischen Einheiten**? Antworten Sie durch ankreuzen von Ja oder Nein!

Nr.	Objekt	Ja	Nein
1.a)	Die Grundgesamtheit	<input type="radio"/>	<input type="radio"/>
1.b)	Merkmalsausprägung	<input type="radio"/>	<input type="radio"/>
1.c)	Stichprobe	<input type="radio"/>	<input type="radio"/>
1.d)	Werteklasse eines Merkmals	<input type="radio"/>	<input type="radio"/>

Klausur-Aufgabe 2. Betrachten Sie die folgenden Tabellen **A** und **B**, die Stichproben-Informationen über das quantitative Merkmal X enthalten:

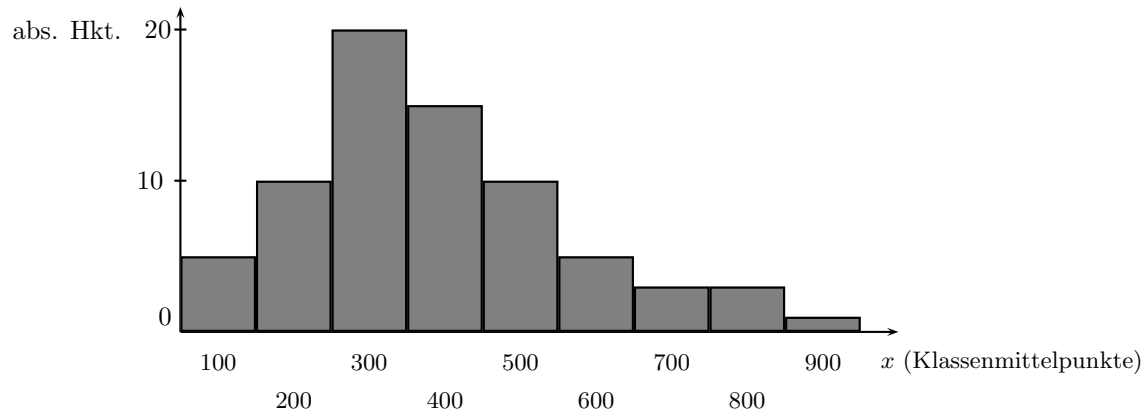
A:	Stat. Einheit Nr.	1	2	3	4	5	6
	X -Wert	7	9	10	13	14	15

B:	X -Klasse	7-9	10-12	13-15	Σ
	Anzahl	2	1	3	6

Welche der folgenden Aussagen ist korrekt? Antworten Sie durch ankreuzen von Ja und Nein!

Nr.	Aussage	Ja	Nein
2.a)	Tabelle A ist eine Urliste	<input type="radio"/>	<input type="radio"/>
2.b)	In Tabelle B werden absolute Summenhäufigkeiten dargestellt	<input type="radio"/>	<input type="radio"/>
2.c)	Tabelle B geht aus Tabelle A durch Klassenbildung hervor	<input type="radio"/>	<input type="radio"/>
2.d)	Der Stichprobenumfang in Tabelle A war 15	<input type="radio"/>	<input type="radio"/>

Klausur-Aufgabe 3. Betrachten Sie das folgende Balkendiagramm absoluter Häufigkeiten:



Kreuzen sie jeweils den zu dem Diagramm passenden Wert (ggf. gerundet) für das arithmetische Mittel \bar{x} und die empirische Standardabweichung s an!

- | | |
|-----------------------------|-----------------------------|
| $\bar{x} =$ | $s =$ |
| <input type="radio"/> 204.3 | <input type="radio"/> 10.2 |
| <input type="radio"/> 387.5 | <input type="radio"/> 182.8 |
| <input type="radio"/> 900 | <input type="radio"/> 611.3 |

Klausur-Aufgabe 4. Betrachten Sie die folgende Kontingenztabelle relativer Häufigkeiten für Merkmale X und Y (gewonnen aus einer Stichprobe vom Umfang $n = 30$).

X/Y	b_1	b_2	\sum
a_1	0.1	0.4	0.5
a_2	0.4	0.1	0.5
\sum	0.5	0.5	1

Kreuzen sie jeweils den zu der Tabelle passenden Wert für die **bedingte relative Häufigkeit** $h(X = a_1|Y = b_2)$ und das **Cramérsche Kontingenzmass** C an!

- | | |
|---------------------------|---------------------------|
| $h(X = a_1 Y = b_2) =$ | $C =$ |
| <input type="radio"/> 10% | <input type="radio"/> 0.1 |
| <input type="radio"/> 20% | <input type="radio"/> 0.6 |
| <input type="radio"/> 80% | <input type="radio"/> 1 |

Klausur-Aufgabe 5. Betrachten Sie die folgenden vier **Streudiagramme** zweier Merkmale X und Y :

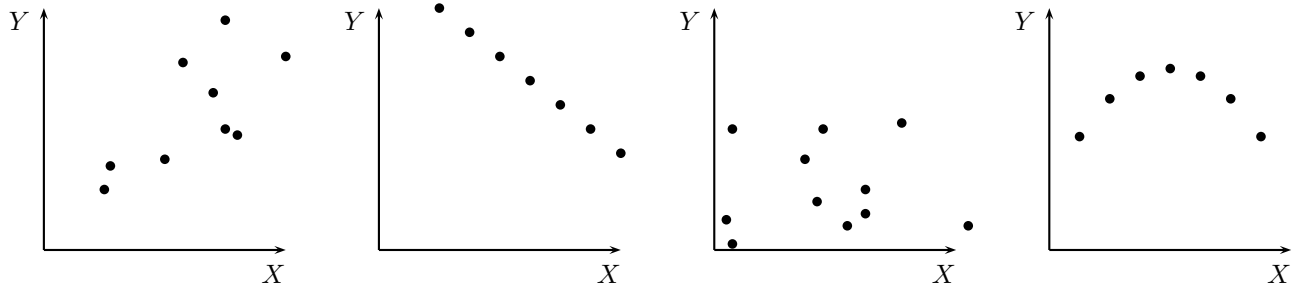


Diagramm 1

Diagramm 2

Diagramm 3

Diagramm 4

Kreuzen Sie den zum jeweiligen Diagramm passenden Wertebereich für den **Korrelationskoeffizienten** $\rho_{X,Y}$ an!

	$\rho_{X,Y} = -1$	$-0.9 \leq \rho_{X,Y} \leq -0.2$	$\rho_{X,Y} \approx 0$ bzw. $= 0$	$0.2 \leq \rho_{X,Y} \leq 0.9$	$\rho_{X,Y} = +1$
Diagr. 1	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Diagr. 2	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Diagr. 3	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Diagr. 4	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Klausur-Aufgabe 6. In einem Versuch wird die Giftproduktion einer Pilzkultur in Abhängigkeit von dem Zuckergehalt einer Nährlösung betrachtet. Die Giftkonzentration Y wird in mg/cm^3 gemessen und der Zuckergehalt X in g/cm^3 . Der Zusammenhang der Merkmale wird durch eine Stichprobe von 5 Zuckerkonzentrationen untersucht. Die Versuchsergebnisse finden Sie in der folgenden Tabelle:

Merkmal/Wert Nr.	1	2	3	4	5
X=Zuckerkonzentration (in g/cm^3)	0.2	0.4	0.6	0.8	1.0
Y=Giftkonzentration (in mg/cm^3)	0.4	0.5	0.8	0.8	0.9

- Zeichnen sie das Streudiagramm.
- Berechnen Sie die Durchschnittswerte \bar{x} und \bar{y} der Merkmale X und Y sowie deren empirische Varianzen s_X^2 und s_Y^2 .
- Berechnen Sie die empirische Kovarianz $Cov_{X,Y}$ sowie den Korrelationskoeffizienten $\rho_{X,Y}$ (gerundet auf 3 Stellen hinterm Komma).
- Bestimmen Sie die Ausgleichsgrade der Form $f(x) = \hat{\alpha} + \hat{\beta}x$ aus dem linearen Regressionsmodell.
- Zeichnen Sie die Ausgleichsgrade in das Streudiagramm und äussern sie sich kurz dazu (etwa 2-3 Sätze), ob das lineare Regressionsmodell die Schadstoffkonzentration Y gut durch die Aenderung der Zuckerkonzentration X erklärt. Stützen Sie Ihre Aeusserung durch eine Interpretation Ihrer Zeichnung und des Anteils der erklärten Streuung R^2 .

Hinweis: Die fünf Versuchsdurchführungen sind als **Stichprobe** vom Umfang $n = 5$ zu sehen.