

Grundlagen der Statistik

Übung 2

Übersicht über die mit den Übungsaufgaben geprüften Lehrzielgruppen

Lehrzielgruppe 1: Zweidimensionale Häufigkeitsverteilungen

Aufgaben: 2 - 4

Lehrzielgruppe 2: Korrelationsmaße

Aufgaben: 5 und 6, 41 - 43

Lehrzielgruppe 3: Interpretation der Korrelationsmaße

Aufgaben: 1, 7 - 9

Aufgabe 1 (5 RP)

Aus 50 Wertepaaren der Merkmale X und Y wurde ein Korrelationskoeffizient (nach Pearson) von $r = -0,9$ berechnet.

Welche der folgenden Aussagen treffen zu?

- (x aus 5)
- A Der Zusammenhang zwischen den Merkmalen ist annähernd linear.
 - B $r = -0,9$ muß falsch berechnet worden sein, da der Korrelationskoeffizient nur Werte annehmen kann, die vom Betrag kleiner oder gleich 1 sind.
 - C Eine Verdopplung der einen Variablen führt annähernd zu einer Verdopplung der anderen Variablen.
 - D 50 Wertepaare reichen zur Berechnung eines Korrelationskoeffizienten nicht aus.
 - E Es liegt überhaupt kein Zusammenhang vor, da $-0,9$ nah bei 0 liegt.

Aufgabe 2 (10 RP)

Gegeben seien zwei diskrete Merkmale X und Y mit Merkmalsausprägungen $x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 4, y_1 = 1, y_2 = 6, y_3 = 9$.

Die gemeinsame Häufigkeitsverteilung ist durch die folgende Tabelle gegeben:

X	x_1	x_2	x_3	x_4
Y				
y_1	4	3	0	0
y_2	0	4	2	0
y_3	0	1	0	4

Welche der folgenden Aussagen sind richtig?

- (x aus 5)
- A Das arithmetische Mittel der Verteilung Y unter der Bedingung $X = x_2$ ist 5.
 - B $f(y_1|x_1) = 1$.
 - C Die Varianz der Verteilung von Y unter der Bedingung $X = x_2$ ist 8,25.
 - D $f(y_2|x_2) = 0,50$.
 - E $f(x_3) = \frac{1}{18}$.

Aufgabe 3 (10 RP)

Gegeben sei die folgende zweidimensionale Häufigkeitstabelle mit den absoluten Häufigkeiten $h(x_i, y_i)$ der beiden Merkmale X und Y für insgesamt 50 Beobachtungspaare.

	X	x_1	x_2
Y			
y_1			15
y_2			

Weiterhin sei bekannt, daß $f(x_2|y_1) = 0,5$ und $f(x_1|y_2) = 0,5$ gilt.

Welche der folgenden Aussagen sind richtig?

- (x aus 5)
- A Es gilt $h(x_1, y_1) = 15$.
 - B Es gilt $h(x_1, y_2) = 15$.
 - C Es gilt $f(x_1, y_2) = 0,2$.
 - D Die Merkmale X und Y sind empirisch unabhängig.
 - E Mit den gemachten Angaben können die Aussagen A - D nicht auf deren Richtigkeit überprüft werden.

Aufgabe 4 (5 RP)

Welche der folgenden Aussagen sind richtig?

- (x aus 5)
- A Unter einer Korrelationstabelle versteht man die Häufigkeitstabelle zweier metrischer oder ordinalskalierbarer Merkmale.
 - B Die absoluten Häufigkeiten der bedingten Verteilungen können direkt der Häufigkeitstabelle entnommen werden.
 - C Die relative Häufigkeit für das gemeinsame Auftreten der Ausprägungen x_j und y_k der Merkmale X und Y ergibt sich als Produkt der entsprechenden relativen Häufigkeiten der Randverteilungen.
 - D Für gegebene Merkmalsausprägungen werden unterschiedliche Klasseneinteilungen vorgenommen. Eine festgestellte Unabhängigkeit für eine ausgewählte Klasseneinteilung impliziert die Unabhängigkeit für alle vorgenommenen Klasseneinteilungen.
 - E Keine der unter A - D angegebenen Alternativen trifft zu.

Aufgabe 5 (10 RP)

40 Schüler wurden nach ihrer Religionszugehörigkeit und nach ihrer Mathematiknote gefragt. Man erhielt folgendes Umfrageergebnis:

Religionszugehörigkeit (Y)	Mathematiknote (X)					
	1	2	3	4	5	
katholisch (=1)	4	5	8	3	0	
protestantisch (=2)	0	3	8	5	4	

Welche der folgenden Maßzahlen bietet eine korrekte Darstellung des Zusammenhangs zwischen Mathematiknote und Religionszugehörigkeit:

1. Pearsonscher Korrelationskoeffizient;
2. Spearmanscher Rangkorrelationskoeffizient;
3. Kontingenzkoeffizient?

Welchen Wert erhalten Sie für das richtige Maß?

- (1 aus 5)
- A 0,40000
 - B 0,42857
 - C 0,45643
 - D 0,96000
 - E 0,98799

Aufgabe 6 (10 RP)

Sechs Schüler wurden befragt, wieviele Stunden sie pro Woche Sport treiben und wieviele Punkte sie bei den Bundesjugendspielen erreicht haben, wobei die Punkte als ordinal messbar angesehen werden. (Sportstunden =: X, Punkte bei den Bundesjugendspielen = Y)

Y	X					
	2	3	4	5	6	8
20	1					
35				1		
38			1			
53					1	
64		1				
94						1

Berechnen Sie die richtige Maßzahl für den Zusammenhang zwischen den Merkmalen X und Y.

Welchen Wert erhalten Sie?

- (1 aus 5)
- A 0,70
 - B 0,62
 - C 0,64
 - D 0,66
 - E 0,60

Aufgabe 7 (5 RP)

Zwei Lehrer beurteilen die Leistungen von Schülern durch Vergabe von Punkten. Für die sich ergebenden Punktzahlen wird ein Spearmanscher Rangkorrelationskoeffizient von $-0,8$ berechnet.

Welche Aussagen sind richtig?

- (x aus 5)
- A Die von beiden Lehrern vergebenen Punktzahlen stimmen bei 80% der Schüler nicht überein.
 - B Die meisten Schüler, die bei einem Lehrer eine hohe Punktzahl haben, haben bei dem anderen Lehrer eine niedrigere Punktbewertung.
 - C Hinsichtlich der Leistungsreihenfolge haben die beiden Lehrer annähernd entgegengesetzte Vorstellungen von der Schülergruppe.
 - D Keine der Aussagen ist richtig, da der Spearmansche Rangkorrelationskoeffizient keine negativen Werte annehmen kann.
 - E Keine der unter A - D angegebenen Aussagen ist richtig.

Aufgabe 8 (5 RP)

Welche der folgenden Aussagen sind richtig?

- (x aus 5)
- A Die Kovarianz zweier Merkmale X und Y liegt immer zwischen -1 und +1.
 - B Die Größe der Kovarianz zweier Merkmale X und Y hängt nicht von deren Streuungen ab.
 - C Die Kovarianz ist ein Parameter für die gemeinsame Streuung zweier quantitativer Merkmale.
 - D Die Kovarianz und der Pearsonsche Korrelationskoeffizient sind geeignete Maße für lineare und alle nichtlinearen Zusammenhänge zwischen zwei metrisch skalierten Merkmalen.
 - E Die Kovarianz ist ein Maß für die Symmetrie einer zweidimensionalen Häufigkeitsverteilung.

Aufgabe 9 (5 RP)

Welcher der folgenden Aussagen sind richtig?

- (x aus 5)
- A Der Kontingenzkoeffizient läßt sich nur verwenden, wenn die Anzahl der Beobachtungen größer als 30 ist.
 - B Bei nominal skalierten Merkmalen wird der Kontingenzkoeffizient verwendet.
 - C Der Kontingenzkoeffizient und der Spearmansche Rangkorrelationskoeffizient lassen sich bei metrisch skalierten Merkmalen anwenden.
 - D Ein negativer Wert des Kontingenzkoeffizienten besagt, dass an den meisten Untersuchungsobjekten große Werte des einen Merkmals zusammen mit kleinen Werten des anderen Merkmals beobachtet wurden.
 - E Wenn für die Beobachtungswerte zweier ordinal skalierten Merkmale jeweils die Ränge gebildet werden und für diese der Pearsonsche Korrelationskoeffizient berechnet wird, so ist dieses identisch mit der Berechnung des Spearmanschen Rangkorrelationskoeffizienten.
-

Aufgabe 41 (10 RP)

Gegeben seien folgende Beobachtungspaare für die Merkmale X und Y:

i	1	2	3	4	5
x_i	1	2	3	4	5
y_i	3	6	5	7	4

Berechnen Sie den Korrelationskoeffizienten r für die Merkmale X und Y.

Welchen Wert erhalten Sie?

(numerisch)

$r =$

Aufgabe 42 (15 RP)

Die Bevölkerung einer Kleinstadt läßt sich nach den Merkmalen Beteiligung am Erwerbsleben und Geschlecht wie folgt gliedern:

Beteiligung am Erwerbsleben	Geschlecht	
	männlich	weiblich
Erwerbstätige	5285	3715
Erwerbslose	335	365
Nichterwerbspersonen	3180	7120

Zur Untersuchung, ob zwischen den Merkmalen ein Zusammenhang besteht, soll der korrigierte Kontingenzkoeffizient berechnet werden (3 Nachkommastellen).

(numerisch)

$C_{korr} =$

Aufgabe 43 (10 RP)

Bei einer Biologie- und einer Chemieklausur wurden von 9 Studenten die folgenden Punktzahlen erreicht:

Student	1	2	3	4	5	6	7	8	9
Biologie	47	44	51	35	29	22	38	19	12
Chemie	34	31	48	36	23	17	39	28	16

Berechnen Sie den Rangkorrelationskoeffizienten von Spearman. Welchen Wert für r_s erhalten Sie (2 Nachkommastellen)?

(numerisch)

$r_s =$

Aufgabe 43 war die letzte Aufgabe