

Prof. Dr. Peter von der Lippe

Klausur in Statistik VWA Essen

neue Regelung (verkürzter Stoff)

Bitte schreiben Sie hier Ihren Namen auf das Deckblatt. **Bitte** neben dieser Aufgabenstellung **keine weitere Blätter benutzen**. Lösungen nach Möglichkeit nur in die hierfür vorgesehen freien Stellen dieses Aufgabenblattes (oder auf der Rückseite) eintragen.

Name:

(nicht ausfüllen)

Punkte	Note

Die Klausur besteht aus zwei Teilen, Multiple Choice Aufgaben und Rechenaufgaben. Erlaubte Hilfsmittel sind:

- Taschenrechner,
- unbeschriebene Formelsammlung mit Tabelle der Standardnormalverteilung.

Viel Erfolg!

11	01	03	
<p>Wie erkenne ich bei einem Merkmal (einer Variable) welche Skalenart vorliegt?</p> <p>a) wenn die Merkmalsausprägungen als Zahlen gegeben sind, dann liegt eine metrische Skala vor;</p> <p>b) das hängt von der Art des Merkmals ab, es gibt Merkmale, die messbar sind (d.h. mit einer metrischen Skala dargestellt werden können) und Merkmale (z.B. qualitative Merkmale) bei denen das nicht der Fall ist;</p> <p>c) indem ich mir überlege, welche Art der Transformation der Skala zulässig ist, in dem Sinne, dass der Informationsgehalt der Zahlen nicht verloren geht;</p> <p>d) alle Antworten a) bis c) sind richtig.</p>			

11	08	02a	
<p>Einem Streuungsdiagramm kann man entnehmen</p> <p>a) Stärke und Richtung (positiv, negativ) der Korrelation;</p> <p>b) ob eine Scheinkorrelation zwischen der Variable x und der Variable y vorliegt,</p> <p>c) ob x die Ursache für y oder y die Ursache für x ist,</p> <p>d) alle Antworten a) bis c) sind richtig.</p>			

11	08	06	
<p>Bei einer Untersuchung lag der Korrelationskoeffizient r zwischen dem Alter x des Ehemannes und dem Alter y der Ehefrau nahe bei $+1$. Das bedeutet, dass:</p> <p>a) bei den meisten Ehepaaren Mann und Frau etwa gleich alt sind;</p> <p>b) die Differenz d_i zwischen dem Alter des Ehemannes und der Ehefrau (also $d_i = x_i - y_i$) bei den meisten der n Ehepaare i ($i = 1, \dots, n$) etwa gleich groß ist;</p> <p>c) je älter der Ehemann ist, desto älter ist in der Regel auch die Ehefrau und umgekehrt (je jünger,... desto jünger...)</p> <p>d) alle Antworten a) bis c) sind richtig.</p>			

11	08	09e	
<p>Bei der Frage nach dem Verhältnis zwischen Korrelation und Kausalität (= Ursache-Wirkungs-Zusammenhang) gilt:</p> <p>a) auch ohne (direkte) Kausalität können zwei Variablen miteinander korrelieren;</p> <p>b) wenn Kausalität, dann auch immer Korrelation, aber die Umkehrung gilt nicht;</p> <p>c) beide Aussagen a) und b) sind richtig;</p> <p>d) beide Aussagen a) und b) sind falsch.</p>			

13	10	02 Teil davon	
----	----	---------------	--

Gegeben seien die Preise von drei Waren A, B, und C (die zusammen die Gesamtheit der Güter der Lebenshaltung privater Haushalte ausmachen mögen) und

Warenpreise zur Zeit 0 und zur Zeit $t = 4$

A	$p_{A0} = 15$	$p_{A4} = 20$
B	$p_{B0} = 25$	$p_{B4} = 40$
C	$p_{C0} = 20$	$p_{C4} = 30$

a) egal, welche Mengen zur Zeit 0 und zur Zeit t konsumiert wurden, ein sinnvoller Index sollte einen Wert annehmen zwischen

mindestens und höchstens

b) Angenommen der Preis von C entwickelt sich wie folgt:

t = 0	t = 1	t = 2	t = 3	t = 4
$p_{C0} = 20$	$p_{C1} = 18$	$p_{C2} = 22$	$p_{C3} = 28$	$p_{C4} = 30$

Tragen Sie bitte in die freien Felder die Werte des dreigliedrigen gleitenden Durchschnitts ein.

11	10	05	
----	----	----	--

Ein Preisindex nach Paasche

a) wird zur Deflationierung (Preisbereinigung) benutzt

b) ist z.B. der Preisindex des Privaten Verbrauchs

c) ist z.B. der Preisindex des Bruttosozialprodukts

d) alle Antworten a) bis c) sind richtig.

11	11	01	
----	----	----	--

Bei einem gleitenden Durchschnitt von 7 Gliedern ($p=7$) gilt folgendes (fehlende Angaben ergänzen, bzw. Zutreffendes ankreuzen)

a) es fallen am Anfang und am Ende Glieder weg

b) muss man einen *zentrierten* gleitenden Durchschnitt

berechnen nicht berechnen

c) dann fällt der erste Wert der gleitenden Durchschnitts auf den-ten Zeitpunkt der Zeitreihe

d) wenn die Zeitreihe einen regelmäßigen Zyklus von genau Perioden (Gliedern, Zeitpunkten) hätte, würde man diesen Zyklus mit diesem gleitenden Durchschnitt genau ausschalten

e) Die Abweichungen der beobachteten Werte (d.h. der Zeitreihe, der Ursprungswerte) von den gleitenden Mittelwerten stellen folgendes dar:

den Trend, bzw. die "glatte Komponente" (=Trend + Konjunktur)

die sog. "saisonbereinigten Werte"

- den Zyklus
- die trendbereinigten Werte.

11	11	03c	
<p>Mit den folgenden Daten lassen sich <u>zentrierte</u> gleitende Durchschnitte von jeweils <u>vier</u> Perioden berechnen:</p> <p style="text-align: center;"> $t = 0$ $t = 1$ $t = 2$ $t = 3$ $t = 4$ $t = 5$ $y_0 = 60$ $y_1 = 50$ $y_2 = 32$ $y_3 = 46$ $y_4 = 60$ $y_5 = 34$ </p> <p>Die Werte, die man für die zentrierten gleitenden Durchschnitte erhält, lauten</p> <p>a) 48, 40, 46 beginnend mit $t=1,5$ b) 47, 45 beginnend mit $t=2$ c) 49,5 44,4 beginnend mit $t=2$ d) keine der Antworten a) bis c) ist richtig. Die richtigen Werte lauten vielmehr:</p>			

12	05	01	
<p>Bei der Geburt eines Kindes ist es in der Regel von Interesse, welches Geschlecht das Kind hat, wobei dieses bis jetzt noch immer nicht im voraus bestimmbar oder beeinflussbar ist (das Geschlecht ist also eine Zufallsvariable). Um auszurechnen, wie wahrscheinlich es genau ist, bei n Kindern genau $x \leq n$ Töchter zu haben rechne ich</p> <p>a) mit der geometrischen Verteilung b) mit der Binomialverteilung c) mit der hypergeometrischen Verteilung d) alle Antworten a) bis c) sind falsch.</p>			

12	06	02	
<p>Eine überragende Intelligenz, etwa ein IQ (Intelligenzquotient) von 137,5 oder mehr ist zwar ein erfreulicher Umstand, er kommt aber leider nur selten vor. Wie wahrscheinlich ist es, einen IQ von 137,5 oder mehr zu haben, wenn man davon ausgeht, dass der IQ normalverteilt ist mit $\mu = 100$ und $\sigma = 15$</p> <p>a) 1,75% b) 1,24% c) 0,62% d) alle Antworten a) bis c) sind falsch.</p>			

12	09	01	
<p>Die Post des Landes x plant eine Gebührenerhöhung für eine ihrer Dienstleistungen, die bisher von 70 vH der Bevölkerung in Anspruch genommen wurde. Eine Stichprobe vom Umfang $n = 2500$ ergab, dass bei der höheren Gebühr immerhin noch 65 vH der Bevölkerung diese Leistung nachfragen werden.</p> <p>Die Post schließt daraus, daß sich die Nachfrager preisunelastisch verhalten und freut sich auf die Mehreinnahmen durch die Gebührenerhöhung: Eine Gebührenerhöhung vorzunehmen und dabei zu erleben, daß es doch nicht zu Mehreinnahmen kommt, weil die Kundschaft davonläuft (ausweicht und die Dienstleistung nicht mehr in Anspruch nimmt)</p> <p>a) ist der Fehler erster Art</p> <p>b) ist der Fehler zweiter Art</p> <p>c) ist in jedem ein Fehler, weil sich ja gezeigt hat, daß statt 70% nur noch 65% die Leistung nachfragen (das hat also mit Fehler erster oder zweiter Art gar nichts zu tun)</p> <p>d) alle Antworten sind falsch, weil man auf dieses Problem die Wahrscheinlichkeitsrechnung gar nicht anwenden kann.</p>			

12	10	01a	
<p>Angenommen, es gelte durch eine Umfrage die durchschnittlichen Ausgaben der Konsumenten für ein bestimmtes Gut festzustellen, wobei die Ausgaben mit einer Genauigkeit von ± 10 DM und einer Sicherheit von 99% zu schätzen sind. Es ist davon auszugehen, dass die Standardabweichung der Ausgaben in der Grundgesamtheit 50 DM beträgt. Wie viele Konsumenten sind danach mindestens zu befragen?</p> <p>a) nur 68 (genau: 67,6); b) 135,29;</p> <p>c) 166; d) 96</p> <p>c) über 200; e) über 300</p> <p>d) kann man nicht bestimmen, es fehlen Angaben.</p>			

12	03	02	
<p>A und B seien disjunkte Ereignisse mit $P(A), P(B) > 0$. Dann gilt</p> <p>a) $P(A B) = P(B A) = 0$</p> <p>b) $P(A)P(B) \neq 0$</p> <p>c) $P(A \cup B) = P(A) + P(B)$</p> <p>d) Alle Antworten a) bis d) sind richtig.</p>			

13	08	01a	
<i>wieder ein Problem mit Vampiren:</i>			
Über den Zusammenhang zwischen der nächtlichen Außentemperatur und der Anfahrzeit eines Vampirs zum nächsten Opfer bestand an fünf Tagen der folgende Zusammenhang			
	Außentemperatur (°C)	Fahrzeit (Minuten)	
	X	Y	
	-20	60	
	-10	50	
	± 0	40	
	+10	30	
	+20	20	
	Summe	0	200
a) Könnte man diese Daten auch in Form einer zweidimensionalen Häufigkeitstabelle darstellen und wie sähe dann diese Tabelle aus? (wenn nötig auf die Rückseite schreiben)			
b) Bestimmen Sie die Regressionsgerade sowie den Korrelationskoeffizienten r und das Bestimmtheitsmaß! (die notwendigen und für die Beurteilung wichtigen Nebenrechnungen am besten auf der Rückseite) (Hinweis: Für die Varianzen gilt $s_x^2 = 200$ und $s_y^2 = 200$)			
c) Machen Sie bitte auf der Rückseite dieses Blattes eine Skizze des Streudiagramm mit der eingezeichneten Regressionsgeraden			
d) Wenn nicht X, die Temperatur, die unabhängige Variable ist, sondern Y, die Fahrzeit und entsprechend nicht Y, sondern X die zu schätzende Größe ist (die abhängige Variable), wie groß ist dann der Korrelationskoeffizient und die Bestimmtheit?			

14	09	03	
<p>Dipl. Kfm. K aus E buchte eine besonders preisgünstige Bildungsreise zu den ägyptischen Pyramiden, wobei jedoch zu bedenken ist, daß sich 1. die Pyramiden in unmittelbarer Nähe zur Wüste befinden, und dass sich 2. die intensive Beschäftigung mit Statistik auf K's Geisteszustand etwas ungünstig ausgewirkt hat. Die Reise führte leider dazu, dass sich K auf unbestimmte Zeit in einem wenig belebten Gebiet befand. Man möge davon ausgehen, dass eine solche unangenehme Situation infolge einer Bildungsreise nur eine Wahrscheinlichkeit von 0,05% (also $\pi = 0,0005$) hat, kann aber auch, da so ein Vorfall äußerst unangenehm ist, eine höhere Wahrscheinlichkeit annehmen.</p>			
<p>a) Formulieren Sie die Null- und Alternativhypothese und geben sie in Worten an, worin hier der Fehler 1. Art und der Fehler 2. Art besteht:</p>			
<p>b) Angenommen, es gelte tatsächlich stets (für beliebig viele Reisen dieser Art) $\pi = 0,0005$. Wie wahrscheinlich ist es dann, daß man bei 50 Bildungsreisen</p> <ol style="list-style-type: none">1. höchstens einmal2. mindestens einmal <p>ein solches tragisches Ende erlebt?</p>			
<p>c) Angenommen man habe eine Stichprobe von 400 Bildungsreisen und es sei nur einmal ein derartiger Vorfall passiert. Kann man dann sagen, die oben aufgestellte Hypothese sei mit $\alpha = 5\%$ einseitig getestet zu verwerfen? Führen Sie den entsprechenden Test durch (Berechnungen evtl. auf der Rückseite) und kreuzen Sie bitte hier Ihre Entscheidung an:</p> <p>z- Wert aus der Tabelle (z_α)</p> <p>Wert der Prüfgröße:</p>			
<p>Entscheidung: Nullhypothese <input type="checkbox"/> annehmen <input type="checkbox"/> ablehnen (verwerfen)</p>			