

Statistik Vorlesung

11. Mai 2023

Dauer der Prüfung: 100 Minuten

ZUNAME:			
VORNAME:		MATR.NR.:	

ERLAUBT: Skriptum des Instituts, nicht-graphikfähiger Taschenrechner

VERBOTEN: alle sonstigen Unterlagen, graphikfähiger Taschenrechner, Handys

Bei den Single-Choice-Fragen bringt eine richtige Antwort 2 Punkte und eine falsche 1 Punkt Abzug.
Es gibt keine negative Punktemitnahme in ein anderes Beispiel.

Aufgabe	max. Punkte	erreichte Punkte
1	12	
2	10	
3	20	
4	20	
5	30	
6	8	
Summe	100	
Note		

1. (12 Punkte)

- a) Die folgenden beiden Datensätze haben den gleichen Median: (49, 42, 18, 19, 1, 1, 0) und (41, 2, 50, 4, 3, 48, 18).

Richtig Falsch

- b) Je näher der Pearson'sche Korrelationskoeffizient bei 1 liegt, desto größer wird die Steigung der Regressionsgeraden.

Richtig Falsch

- c) Die Zufallsgröße „Augenzahl eines fairen Würfels“ ist diskret gleichverteilt mit $x_i = i$ und $p_i = \frac{1}{6}$ für $i = 1 \dots 6$

Richtig Falsch

- d) Eine Wahrscheinlichkeit $P(x_1 \leq X \leq x_2)$ ist bei einer Verteilungsfunktion als Fläche zwischen dem Graphen der Funktion, der x -Achse und den jeweiligen Grenzen (x_1 und x_2) gegeben.

Richtig Falsch

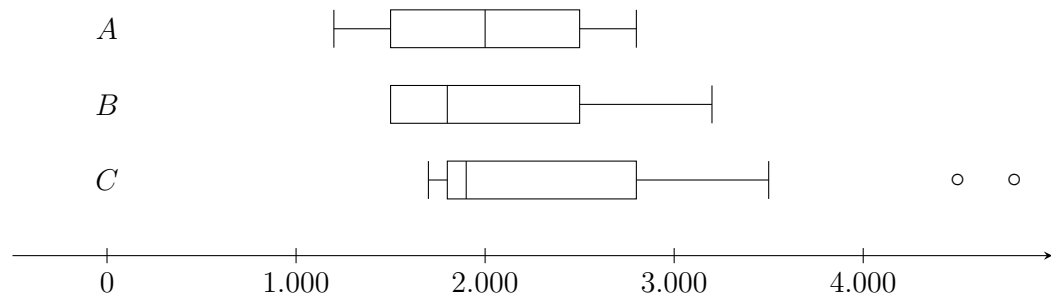
- e) Ist beim Gauß-Test die Testgröße groß (z.B. $t_0 > z_\gamma$), ist signifikant bestätigt, dass die Daten keiner Normalverteilung unterliegen.

Richtig Falsch

- f) Mit der Methode der kleinsten Fehlerquadrate wird die Regressionsgerade so geschätzt, dass die Summe der quadrierten Residuen minimiert wird.

Richtig Falsch

2. (10 Punkte) Ein Unternehmen zeichnet die Tagesumsätze seiner drei Filialen A , B und C ein Jahr lang auf und stellt diese graphisch dar.



- a) In Filiale C gibt es außergewöhnlich große Umsätze, die möglicherweise Ausreißer sind.
 Richtig Falsch
- b) Mindestens 50 % der Umsätze in Filiale A und mindestens 75 % der Umsätze in Filiale B betragen 2.000 Euro oder mehr.
 Richtig Falsch
- c) Die Spannweite ist bei den Daten der Filiale B genauso groß wie bei den Daten der Filiale C .
 Richtig Falsch
- d) Das arithmetische Mittel ist bei den Daten der Filiale C kleiner als der Median.
 Richtig Falsch
- e) Die Umsätze in Filiale C sind linksschief.
 Richtig Falsch

3. Für ihre Bachelorarbeit zum Thema Ernährungsverhalten interviewt Linda sechs Personen. Neben diesen ausführlichen, qualitativen Interviews wurden u.a. auch folgende (quantitative) Daten der sechs Personen erhoben: monatliches Einkommen (in Euro), monatliche Ausgaben für Lebensmittel (in Euro), Stellenwert von gesunder Ernährung (1 = sehr wichtig, 2 = eher wichtig, 3 = eher nicht wichtig, 4 = überhaupt nicht wichtig).

Interviewte Personen	1	2	3	4	5	6
monatliches Einkommen (in Euro)	2.400	1.650	3.100	2.800	2.350	4.800
monatliche Ausgaben (in Euro)	810	580	820	1.300	780	2.800
Stellenwert von gesunder Ernährung	2	4	3	1	2	1

- a) (10 Punkte) Welche Skalenniveaus haben die Merkmale? Bestimmen Sie das am besten geeignete Zusammenhangsmaß zwischen den Ausgaben für Lebensmittel und dem Stellenwert von gesunder Ernährung! Interpretieren Sie das Ergebnis ausführlich!
- b) (10 Punkte) Geben Sie die Gleichung der Regressionsgeraden zur Beschreibung der Abhängigkeit der Ausgaben für Lebensmittel vom Einkommen an!

Ausführung Beispiel 3:

Ausführung Beispiel 3:

4. a) In einem Casino wird unter anderem ein einfaches Würfelspiel angeboten. Es wird einmal mit einem fairen Würfel gewürfelt. Bei einem Einsatz von 10 Euro erhalten Sie Ihren Einsatz zurück, wenn eine 1 gewürfelt wird oder Sie erhalten 26 Euro, wenn eine 6 gewürfelt wird, sonst erhalten Sie nichts.
1. (5 Punkte) Berechnen Sie Erwartungswert und Varianz für den Gewinn!
 2. (2 Punkte) Angenommen das Ereignis A ist definiert als $\{5, 6\}$. Berechnen Sie für die Augenzahl X mit den Ausprägungen $x_i = 1, \dots, 6$ die bedingten Wahrscheinlichkeiten $P(X = x_i|A)$ für eine mögliche Augenzahl x_i und geben Sie diese bedingten Wahrscheinlichkeiten in Tabellenform an!
 3. (3 Punkte) Berechnen Sie die bedingte Erwartung für den Gewinn dieses Spiels, wenn Sie wissen, dass $X \in A$.
- b) Die Anzahl der Personen, die dieses Würfelspiel in einer Stunde spielen, ist Poisson-verteilt. Im Durchschnitt spielen es zehn Personen pro Stunde.
1. (3 Punkte) Wie groß ist die Wahrscheinlichkeit, dass in einer Stunde vier oder mehr Personen dieses Spiel spielen?
 2. (4 Punkte) Ein Casino-Abend hat sieben Stunden. Wie groß ist die Wahrscheinlichkeit, dass an einem Abend in jeder der sieben Stunden vier oder mehr Spieler dieses Spiel spielen?
 3. (3 Punkte) Wie ist die Anzahl der Spieler an einem Casino-Abend verteilt? Begründen Sie ausführlich!

Ausführung Beispiel 4:

Ausführung Beispiel 4:

5. a) (13 Punkte) Der Muttertag steht vor der Tür, aber laut einer Befragung von 250 Personen, wissen 105 Personen nicht, dass der Muttertag in Österreich immer auf den zweiten Sonntag im Mai fällt.
1. Berechnen Sie ein zweiseitiges 95-%-Konfidenzintervall für den Anteil der Personen, die nicht wissen, wann der Muttertag ist und interpretieren Sie dieses!
 2. Es wird vor allem der Generation-unter-30 nachgesagt, dass sie nicht weiß, wann Mütter ihren Ehrentag haben. Um dieser Frage nachzugehen, soll ein 95-%-Konfidenzintervall mit einer gewünschten Länge von 0,07 berechnet werden. Allerdings liegen keine Informationen über den Stichprobenanteil vor. Welchen Stichprobenumfang empfehlen Sie?
- b) (17 Punkte) Immer wieder ist davon zu lesen, dass besonders Männer den Muttertag für sich entdeckt haben. Um zu überprüfen, ob Männer tatsächlich mehr als die Frauen für Geschenke für ihre Mütter ausgeben, wurden jeweils 7 Personen dazu befragt und das Ergebnis in der folgenden Tabelle zusammengefasst:

Männer Ausgaben in Euro	65	93	57	72	46	97	88
Frauen Ausgaben in Euro	42	69	53	48	73	65	34

Erfahrungsgemäß können die Ausgaben für den Muttertag nicht als Realisation einer normalverteilten Zufallsgröße angesehen werden. Lässt sich zeigen, dass Männer mehr für die Geschenke zum Muttertag ausgeben als Frauen?

1. Welcher Test ist anzuwenden? Begründen Sie Ihre Entscheidung!
2. Formulieren Sie die Hypothesen dieses Tests!
3. Berechnen Sie den Testwert!
4. Bestimmen Sie den p -Value!
5. Wie ist zu einem Testniveau von 5 % zu entscheiden? Interpretieren Sie!
6. Ändert sich Ihre Antwort aus 5., wenn das Testniveau 1 % beträgt?

Ausführung Beispiel 5:

Ausführung Beispiel 5:

6. (8 Punkte) Regina arbeitet während ihres Studiums als Kellnerin in einem Innenstadtlokal. Bei dieser Tätigkeit ist das Trinkgeld ein wesentlicher Teil ihres Einkommens. Die nachfolgende Tabelle zeigt vier Rechnungsbeträge und das erhaltene Trinkgeld. Regina möchte mit diesen Daten ein Regressionsmodell ermitteln, um daraus das Trinkgeld für einen beliebigen Rechnungsbetrag zu schätzen.

Rechnungsbetrag in Euro	20	25	30	50
Trinkgeld in Euro	2	3	4	7

Regina hat auch bereits die Koeffizienten der Regressionsgerade berechnet:
Achsenabschnitt $a = -1,12$ und Steigung $b = 0,164$.

- Berechnen Sie die Residuen!
- Wie hoch ist gemäß dieses Regressionsmodells das zu erwartende Trinkgeld für einen Rechnungsbetrag von 15 Euro?
- Wie hoch ist die Varianz des Störterms?

Ausführung Beispiel 6: