

Statistik Vorlesung

4. Juli 2023

Dauer der Prüfung: 100 Minuten

ZUNAME:			
VORNAME:		MATR.NR.:	

ERLAUBT: Skriptum des Instituts, nicht-graphikfähiger Taschenrechner

VERBOTEN: alle sonstigen Unterlagen, graphikfähiger Taschenrechner, Handys

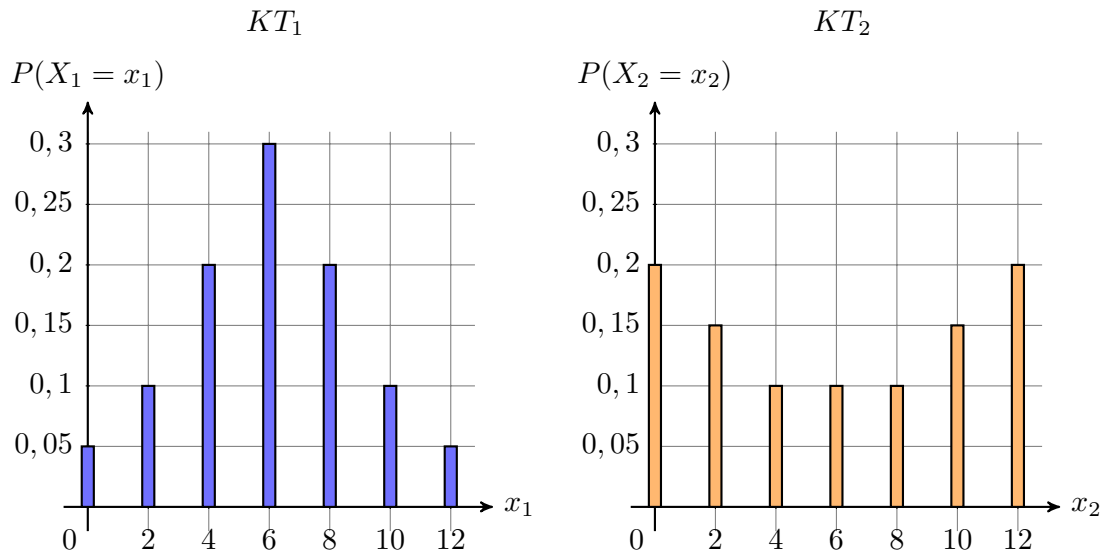
Bei den Single-Choice-Fragen bringt eine richtige Antwort 2 Punkte und eine falsche 1 Punkt Abzug.
Es gibt keine negative Punktemitnahme in ein anderes Beispiel.

Aufgabe	max. Punkte	erreichte Punkte
1	12	
2	10	
3	20	
4	20	
5	26	
6	12	
Summe	100	
Note		

1. (12 Punkte)

- a) Bei klassierten Daten können die Klassenmittelpunkte mit den Klassenmittelwerten identisch sein.
 Richtig Falsch
- b) Die Menge des verzehrten Fleisches pro Jahr dividiert durch die Bevölkerungsanzahl ist eine Kennzahl, bei der Babys eine unbeteiligte Masse darstellen.
 Richtig Falsch
- c) Die approximierende Verteilungsfunktion einer stetigen Zufallsvariablen ist unstetig.
 Richtig Falsch
- d) Die Dichtefunktion $f(x)$ jeder stetigen Zufallsvariablen hat einen Wendepunkt.
 Richtig Falsch
- e) Ist beim Gauß-Test mit $H_0 : \mu \leq \mu_0$ der Testwert groß (z.B. $t_0 > z_\gamma$), ist signifikant bestätigt, dass die Daten höchstens gleich dem vermuteten μ_0 sind.
 Richtig Falsch
- f) Wenn der F -Test in der Regression signifikant ist, ist zumindest eine erklärende Variable sinnvoll.
 Richtig Falsch

2. (10 Punkte) In der Statistik-Übung werden im Semester zwei Kurztests durchgeführt, bei denen man jeweils bis zu 12 Punkte erreichen kann. Die beiden Tests KT_1 und KT_2 werden auf die Anzahl richtig gelöster Beispiele hin untersucht. Die Zufallsvariablen X_1 und X_2 beschreiben die Anzahl der beim jeweiligen Test erreichten Punkte. Die beiden nachstehenden Abbildungen zeigen jeweils die Verteilungen der beiden Variablen X_1 und X_2 . Welche der folgenden Aussagen sind richtig?



- a) Beide Zufallsvariablen X_1 und X_2 sind binomialverteilt.
 Richtig Falsch
- b) Bei KT_1 erreichen mehr Teilnehmer mehr als 6 Punkte als bei KT_2 .
 Richtig Falsch
- c) Die Mittelwerte der beiden Zufallsvariablen sind gleich.
 Richtig Falsch
- d) Die Standardabweichungen der beiden Zufallsvariablen sind gleich: $\sigma_{X_1} = \sigma_{X_2}$.
 Richtig Falsch
- e) Beide Verteilungen sind symmetrisch.
 Richtig Falsch

3. Ein Konzern ist weltweit tätig, er hat vier Niederlassungen in Europa und sechs in Südamerika. Die jährlichen Umsätze (in Mio. €) der Niederlassungen sind in der folgenden Tabelle aufgelistet:

Europa	219	138	530	321		
Südamerika	358	189	222	471	123	250

Der Umsatzmittelwert der südamerikanischen Niederlassungen beträgt 268,833 Mio. €, die Umsatz-Standardabweichung für Südamerika beträgt 114,754 Mio. €.

- (7 Punkte) Bestimmen Sie den Median, den Mittelwert und die Standardabweichung der Umsätze für Europa! Berechnen Sie eine geeignete Kennzahl zur Klassifizierung der Streuung! Wo ist die Streuung der Umsätze höher, in Europa oder in Südamerika?
- (5 Punkte) Berechnen Sie den durchschnittlichen konzernweiten Umsatz sowie die Standardabweichung des Umsatzes für den gesamten Konzern mit den Formeln für gruppierte Daten.
- (4 Punkte) Zeichnen Sie ein geeignetes Diagramm, in dem der Umsatz aller Niederlassungen des Konzerns wie folgt in Klassen zusammengefasst dargestellt wird: 0 bis 200 Mio. €, über 200 bis 500 Mio. €, über 500 bis 800 Mio. €.
- (4 Punkte) Zeichnen Sie eine geeignete Verteilungsfunktion mit der Klasseneinteilung aus c) und kennzeichnen Sie darin näherungsweise den Umsatz, der von 10 % der Niederlassungen überschritten wird.

Ausführung Beispiel 3:

Ausführung Beispiel 3:

4. a) (10 Punkte) Endlich hat nun doch der Sommer begonnen und die Grillsaison ist eröffnet! Bei einem Sommerfest werden 3 verschiedene Griller verwendet, um Koteletts zu grillen: 30 % der Koteletts werden auf einem Holzkohlegriller, 45 % auf einen Gasgriller und die restlichen Koteletts auf einem Elektrogriller gegrillt. Aus Kostengründen stammen insgesamt 87,6 % der angebotenen Koteletts aus konventioneller Tierhaltung, die übrigen aus Bio-Tierhaltung. Der Anteil der Biokoteletts beim Holzkohlegriller beträgt 12 % und beim Elektrogriller 10 %.
1. Stellen Sie den Sachverhalt in einem Baumdiagramm dar.
 2. Wie hoch ist der Bio-Anteil beim Gasgriller?
 3. Jemand bekommt ein Bio-Kotelett serviert. Mit welcher Wahrscheinlichkeit wurde dieses auf einem Holzkohlegriller gegrillt?
 4. Wie wahrscheinlich ist es, rein zufällig ein Kotelett aus konventioneller Tierhaltung serviert zu bekommen, das nicht auf einem Holzkohlegriller gegrillt wurde?
- b) (4 Punkte) Was wäre ein Sommerfest ohne Erfrischungsgetränke? Unter anderem wird Zitronenlimonade in Bechern angeboten. Die Füllmenge ist normalverteilt mit einem Erwartungswert von 200 ml und einer Standardabweichung von 25 ml.
1. Bestimmen Sie den Wert der Verteilungsfunktion an der Stelle 190!
 2. Bestimmen Sie das 0,2-Quantil dieser Verteilung!
- c) (6 Punkte) Ein Kellner serviert Becher mit Zitronenlimonade auf einem Tablett. Darauf haben 8 Becher Platz. Das Gewicht des Tablett ist normalverteilt mit einem Erwartungswert von 390 g und einer Standardabweichung von 10 g. Das Gewicht eines gefüllten Bechers ist ebenfalls normalverteilt mit einem Erwartungswert von 200 g und einer Standardabweichung von 25 g. Wie groß ist die Wahrscheinlichkeit, dass ein Tablett mit 8 gefüllten Bechern zwischen 1900 g und 2100 g wiegt?

Ausführung Beispiel 4:

Ausführung Beispiel 4:

5. In einer repräsentativen Umfrage unter 1.000 Touristen wurde u.a. ihre Herkunft (inländische/ausländische Gäste) und ihre Reisedauer (übernachtende Touristen/Tagesbesucher) erhoben:

Ausländische Gäste gesamt	420
Übernachtende Touristen	325
Tagesbesucher	95
Inländische Gäste gesamt	580
Übernachtende Touristen	415
Tagesbesucher	165

- a) (13 Punkte)
1. Bestimmen Sie ein 99-%-Konfidenzintervall für den Anteil der inländischen Gäste und interpretieren Sie dieses!
 2. Ein Institut, das mit derselben Stichprobe arbeitet, gibt an, der Anteil ausländischer Gäste läge zwischen 40 und 44 Prozent. Welches Konfidenzniveau hat dieses Institut unterstellt?
- b) (13 Punkte) Gibt es einen Zusammenhang zwischen den beiden Merkmalen Herkunft und Reisedauer? Wählen Sie einen geeigneten Test aus, führen Sie diesen durch ($\alpha = 0,01$) und interpretieren Sie Ihr Ergebnis!

Ausführung Beispiel 5:

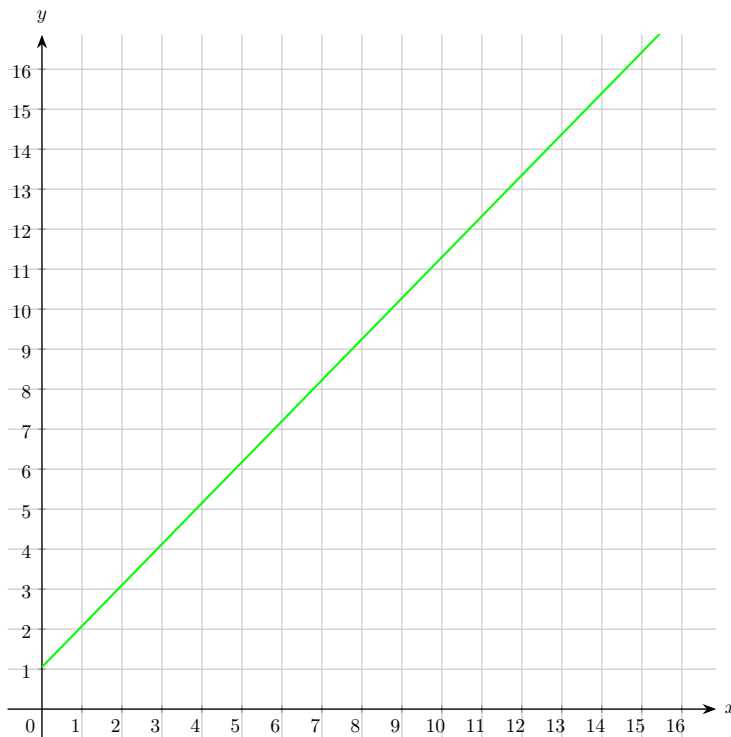
Ausführung Beispiel 5:

6. (12 Punkte) Um die Abhängigkeit des Weizenertes von der Düngemittelmenge zu schätzen, führt ein Agrarbetrieb eine lineare Regressionsanalyse durch. Für die statistische Untersuchung werden fünf gleich große Ackerbauflächen verwendet. Die Untersuchung brachte folgendes Ergebnis (ha ... Hektar = 10.000 m²):

Düngemittel/ha in kg (X)	1	3	5	9	12
Weizenertag/ha in Tonnen (Y)	3	5	5	7	16

Folgende Werte wurden aus dem Datensatz bereits ermittelt: $\sigma_x = 4$, $\sigma_y = 4,578$, $\bar{y} = 7,2$, $Cov(X,Y) = 16,4$.

- (3 Punkte) Stellen Sie die Gleichung der Regressionsgeraden auf!
- (2 Punkte) Geben Sie eine Schätzung an, wie hoch der erreichte Weizenertag/ha sein wird, wenn pro ha im Schnitt 10 kg Düngemittel eingesetzt werden! Wie hoch ist die geschätzte Zu- bzw. Abnahme des Weizenertes in Tonnen/ha, falls die Düngemittelmenge um 1 kg/ha erhöht wird?
- (2 Punkte) Welcher geschätzte Weizenertag würde sich bei einem Düngemiteleinsatz von 40 kg/ha ergeben? Ist diese Schätzung sinnvoll? Begründen Sie!
- (3 Punkte) Im nachstehenden Koordinatensystem wurde die in a) ermittelte Regressionsgerade eingezeichnet. Tragen Sie den beobachteten Punkt mit den Werten $x = 9$ und $y = 7$ in die Zeichnung ein! Kennzeichnen Sie den geschätzten Wert, berechnen Sie den Fehler und kennzeichnen Sie diesen in der Graphik!
- (2 Punkte) Das Bestimmtheitsmaß entspricht bei der einfachen Regression dem quadrierten Korrelationskoeffizienten. Berechnen Sie das Bestimmtheitsmaß und erläutern Sie die Bedeutung des erhaltenen Wertes!



Ausführung Beispiel 6:

Ausführung Beispiel 6: