

# Statistik Vorlesung

25. September 2024

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.

Berechnungen müssen nachvollziehbar aufgeschrieben sein.

Aufgabe	max. Punkte	erreichte Punkte
1	12	
2	10	
3	15	
4	20	
5	32	
6	11	
Summe	100	
Note		

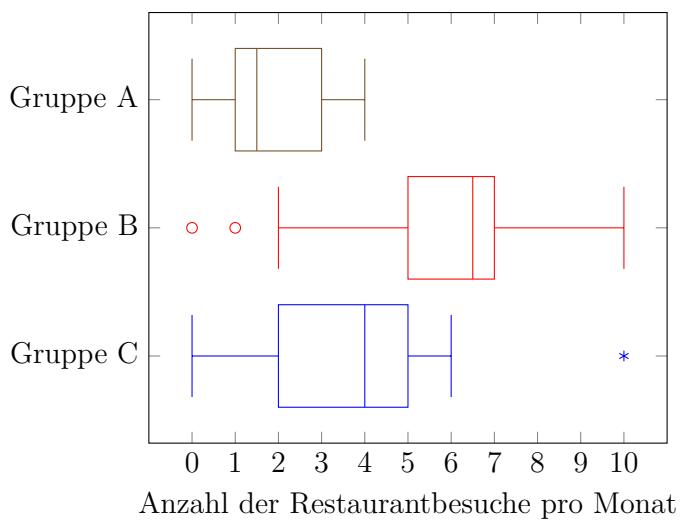
1. (12 Punkte)

- a) Der Übergang von einer metrischen Skala auf eine Ordinalskala bedeutet im Allgemeinen einen Informationsverlust.
- Richtig       Falsch
- b) Der Kontingenzkoeffizient kann zur Messung des Zusammenhangs zwischen einem nominalen und einem ordinalen Merkmal verwendet werden.
- Richtig       Falsch
- c) Der zentrale Grenzwertsatz besagt, dass die Summe von unabhängigen Zufallsvariablen gegen eine Normalverteilung konvergiert.
- Richtig       Falsch
- d) Als Verteilungsfunktion  $F(x)$  einer stetigen Zufallsvariable bezeichnet man die Funktion, die die Wahrscheinlichkeit dafür angibt, dass die Zufallsvariable  $X$  genau den Wert  $x$  annimmt.
- Richtig       Falsch
- e) Ist der  $p$ -Value kleiner als die Irrtumswahrscheinlichkeit  $\alpha$ , so ist die  $H_1$  bestätigt.
- Richtig       Falsch
- f) Die Regressionsgerade bei der linearen Regression geht stets durch mindestens zwei Punkte des Streudiagramms.
- Richtig       Falsch

Lösung: R R R F R F

2. (10 Punkte)

Bei einer Studie wurden 600 Personen in drei Gruppen (A, B und C) eingeteilt und befragt, wie häufig sie pro Monat ein Restaurant besuchen. Eine Auswertung der Umfrage sieht wie folgt aus:



- a) In den Gruppen A, B und C wurden jeweils 200 Personen befragt.  
 Richtig       Falsch       Nicht ablesbar
  
- b) Die Hälfte der Personen in Gruppe A besucht weniger als zweimal pro Monat ein Restaurant.  
 Richtig       Falsch       Nicht ablesbar
  
- c) Personen der Gruppe C besuchen durchschnittlich viermal pro Monat ein Restaurant.  
 Richtig       Falsch       Nicht ablesbar
  
- d) Der Quartilsabstand in der Gruppe C beträgt 6.  
 Richtig       Falsch       Nicht ablesbar
  
- e) Personen in der Gruppe B besuchen im Durchschnitt signifikant häufiger ( $\alpha \leq 0,1$ ) ein Restaurant als Personen aus der Gruppe C.  
 Richtig       Falsch       Nicht ablesbar

Lösung: N R N F N

3. Die Statistik Austria hat erhoben, dass es am Jahresende 2022 in Österreich 759.000 Personen mit registrierter Behinderung gab. In der folgenden Tabelle finden Sie die Personen [in Tausend] gegliedert nach Alter und Geschlecht.

Alter in Jahren	[0; 25[	[25; 45[	[45; 65[	[65; 85[	[85; 95]
Männer	18	32	110	141	41
Frauen	12	27	97	174	107
Personen insgesamt	30	59	207	315	148

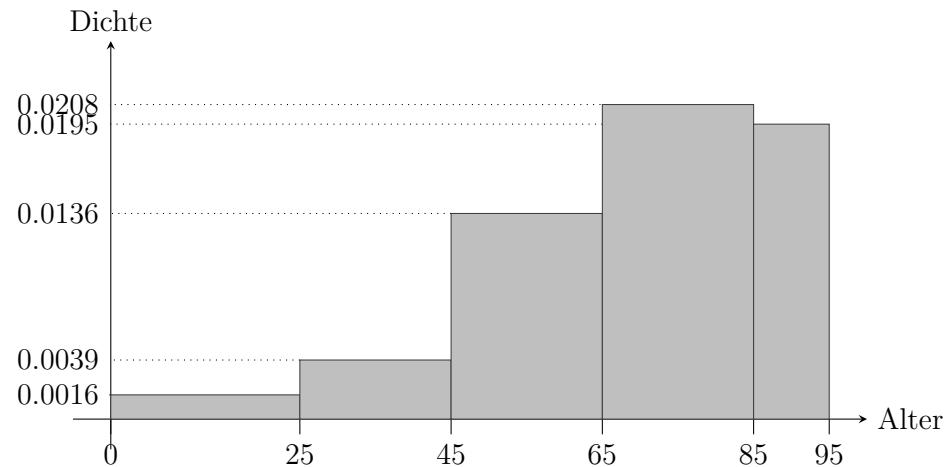
- a) (5 Punkte) Zeichnen Sie ein Histogramm für das Merkmal Alter der registrierten behinderten Personen!
- b) (7 Punkte) Bestimmen Sie Mittelwert und Standardabweichung des Merkmals Alter der registrierten behinderten Personen!
- c) (2 Punkte) Sie möchten wissen, ob es einen Zusammenhang zwischen dem Merkmal Geschlecht und dem Merkmal Alter gibt. Welche Kennzahl ist dafür geeignet? Begründen Sie! Sie müssen diese Kennzahl nicht berechnen!
- d) (1 Punkt) Angenommen die Kennzahl aus c) beträgt 0,268. Interpretieren Sie diesen Wert im gegebenen Kontext!

Ausführung Beispiel 3:

Ausführung Beispiel 3:

Lösung:

a)



- b) Mittelwert: 66,891; Standardabweichung: 19,130
- c) korrigierter Kontingenzkoeffizient, da Geschlecht nominal
- d) Der Zusammenhang zwischen dem Merkmal Geschlecht und dem Merkmal Alter ist schwach.

4. a) Julia bucht einen 4-tägigen Kurzurlaub auf Teneriffa. Sie hat im Internet recherchiert, dass dort im September durchschnittlich an 27 Tagen die Sonne scheint, an den restlichen drei Tagen ist es bewölkt oder es regnet. Das Wetter an einem Tag ist vom Wetter an einem anderen Tag unabhängig.
- i. (2 Punkte) Berechnen Sie die Wahrscheinlichkeit, dass Julia während ihres gesamten Urlaubs nur Sonnentage genießen kann!
  - ii. (4 Punkte) Wie hoch ist die Wahrscheinlichkeit, dass es in Julias Urlaub an maximal einem Tag bewölkt ist oder regnet? Erstellen Sie ein Baumdiagramm und berechnen Sie die Wahrscheinlichkeit!
- b) Der Tourismusverband von Teneriffa hat erhoben, dass die durchschnittlichen Ausgaben von Touristen für ihren Urlaub normalverteilt sind. Die durchschnittlichen Ausgaben betragen 1.500 Euro pro Person, die Standardabweichung liegt bei 200 Euro.
- i. (8 Punkte) Berechnen Sie die Wahrscheinlichkeit, dass ein zufällig ausgewählter Tourist zwischen 1.400 und 1.700 Euro ausgibt. Zeichnen Sie die diesem Beispiel entsprechende Dichtefunktion, tragen Sie die Parameter der Verteilung ein und kennzeichnen Sie das Wahrscheinlichkeitsintervall!
  - ii. (2 Punkte) Wie hoch ist die Wahrscheinlichkeit, dass Julia genau 1.500 Euro für ihren Teneriffaurlaub ausgibt? Begründen Sie Ihre Antwort!
  - iii. (4 Punkte) Wie viel geben die sparsamsten 20 % der Teneriffa-Touristen maximal für ihren Urlaub aus? Tragen Sie das berechnete Ergebnis ebenfalls in die Dichtefunktion von i. ein!

Ausführung Beispiel 4:

Ausführung Beispiel 4:

Lösung:

- a) 0,6561; 0,9477
- b) 0,533; 0; 1.331,68

5. a) (6 Punkte) Laut einer Umfrage können von 90 befragten Personen 63 Personen nicht stricken. Berechnen Sie ein zweiseitiges 95 %-Konfidenzintervall für den Anteil der Personen, die nicht stricken können und interpretieren Sie dieses!

b) (12 Punkte) Aufgrund dieses Ergebnisses aus a) bietet ein Handarbeitsgeschäft wieder Strickkurse an Werktagen an. Die Anmeldezahlen sind in der folgenden Tabelle zusammengefasst:

Wochentag	Montag	Dienstag	Mittwoch	Donnerstag	Freitag	Samstag
Anmeldungen	42	69	53	48	63	55

Überprüfen Sie zum Testniveau  $\alpha = 10\%$  die Behauptung, dass alle Wochentage gleich beliebt sind! Formulieren Sie die Hypothesen und führen Sie den entsprechenden Test durch! Interpretieren Sie Ihr Ergebnis!

c) (14 Punkte) Stricken ist bekanntlich auch eine heilsame und entspannende Tätigkeit, die viele Vorteile für die Gesundheit hat. Um dies zu überprüfen, wurde beim Strickkurs von jeweils 9 Personen vor dem Kurs bzw. nach dem Kurs der Ruhepuls [Schläge/Minute] gemessen und das Ergebnis in folgender Tabelle zusammengefasst:

Person	1	2	3	4	5	6	7	8	9
Ruhepuls vor dem Kurs	74	87	78	81	69	72	79	84	75
Ruhepuls nach dem Kurs	66	75	85	69	78	58	79	79	67

Testen Sie ohne Annahme der Normalverteilung! Lässt sich zeigen, dass der Ruhepuls nach dem Strickkurs niedriger ist?

- i. Welcher Test ist anzuwenden? Begründen Sie Ihre Entscheidung!
- ii. Formulieren Sie die Hypothesen dieses Tests!
- iii. Berechnen Sie den Testwert!
- iv. Bestimmen Sie den  $p$ -Value!
- v. Wie ist zu einem Testniveau von  $\alpha = 5\%$  zu entscheiden?

Ausführung Beispiel 5:

Ausführung Beispiel 5:

**Lösung:**

- a)  $[0,605; 0,795]$  Mit einer Wahrscheinlichkeit von 95 % liegt der wahre Anteil an Personen, die nicht stricken können, in diesem Intervall.

b)  $H_0 :$

$X$	Montag	Dienstag	Mittwoch	Donnerstag	Freitag	Samstag
$P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$H_1 :$  Die Zufallsvariable  $X$  ist anders verteilt.

$t_0 = 9,236; K = ]9,236; \infty[$

$t_0$  nicht in  $K$ ;  $H_0$  beibehalten.

Die Annahme diskreter Gleichverteilung kann aufrecht erhalten bleiben.

- c) Wilcoxon-Vorzeichenrangtest

$H_1 : d_{0,5} < 0 H_0 : d_{0,5} \geq 0$

$t_0 = -1,540$

p-Value: 0,062

$H_1$  nicht bestätigt.

6. Die Maronistände haben wieder Einzug gehalten! Um dem Ansturm um die Mittagszeit gerecht zu werden, möchte ein Betreiber die leeren Tüten rechtzeitig vorbereiten. Er geht davon aus, dass je nach Temperatur [in °C] die Anzahl der verkauften Tüten [in Stück, zwischen 11:00 und 13:00 Uhr] schwanken wird. Dazu hat er folgende Aufzeichnungen aus dem vorigen Jahr herangezogen:

Temperatur	0	-5	-3	4
Anzahl verkaufter Maronitüten	40	60	50	30

Die Regressionsgerade wurde aus dem Modell wie folgt geschätzt:  $\hat{Y} = 41,739 - 3,261x$  Der Standardfehler der Regression wurde ebenfalls geschätzt:  $\hat{\sigma} = 2,331$ .

- a) (3 Punkte) Berechnen Sie den Standardfehler des Koeffizienten  $\hat{\beta}_1 (\hat{\sigma}_{\hat{\beta}_1})$  und interpretieren Sie diesen!
- b) (8 Punkte) Kann die Temperatur in °C als ein für den Absatz unverzichtbares erklärendes Merkmal angesehen werden? Führen Sie den entsprechenden Test zum Niveau  $\alpha = 5\%$  durch und beantworten Sie die Frage!

Ausführung Beispiel 6:

Ausführung Beispiel 6:

Lösung:

- a) 0,344; Schwankung des geschätzten Koeffizienten  $\beta_1$
- b)  $H_0 : \beta_1 = 0$   $H_1 : \beta_1 \neq 0$   
 $t_0 = -9,488$ ;  $K = ] -\infty; -4,303 [ \cup ] 4,303; \infty [$   $t_0 \in K$ ;  $H_1$  bestätigt.  
Ja, die Temperatur kann als unverzichtbar angesehen werden.