Statistik Vorlesung

19. März 2026

Arbeitszeit: 100 Minuten

VORNAME:		MATR.NR.:	
NACHNAME:			

ERLAUBT: **Skriptum des Instituts, Taschenrechner gemäß Liste**

VERBOTEN: **alle sonstigen Unterlagen, Handys**

Aufgaben 8 und 9: Für jede vollständig richtig gelöste Frage gibt es 2 Punkte. Es gibt keine Teilpunkte und keine Punkteabzüge.

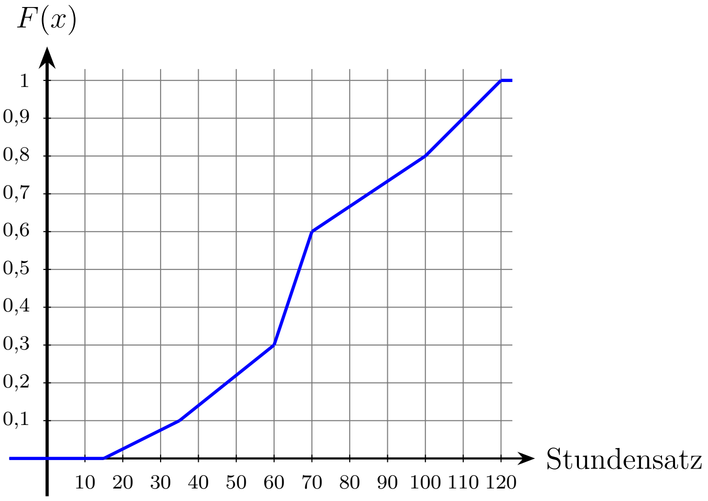
Lösungswege müssen nachvollziehbar aufgeschrieben sein.

Aufgabe	max. Punkte	erreichte Punkte	Teilsumme
1	07		
2	10		
3	12		
4	11		
5	15		
6	15		
7	10		
8	10		
9	10		
Summe	100		
Note			

1. Bei einer Erhebung der WKO wurden 300 Freelancer im Bereich IT nach ihrem Stundensatz gefragt. Um den Datenschutz zu wahren, mussten die befragten Personen keine genauen Angaben machen, sondern nur angeben, in welche Kategorie Sie fallen:

Stundensatz in €	absolute Häufigkeiten
15–35	
35–60	
60–70	
70–100	
100–120	

Die entsprechende approximierende empirische Verteilungsfunktion sieht zu den erhobenen Daten wie folgt aus:



a) (2 Punkte) Bestimmen Sie mit Hilfe der obigen Graphik die absoluten Häufigkeiten für die einzelnen Kategorien. Sie können Ihre Ergebnisse direkt in die obige Tabelle einfügen.

b) (2 Punkte) Bestimmen Sie $F(100)$ und interpretieren Sie den Wert im gegebenen Kontext.

c) (3 Punkte) Berechnen Sie näherungsweise den Median für die Daten.

Lösung:

a) (30, 60, 90, 60, 60)

b) $F(100) = 0{,}8$ bedeutet, dass 80 Prozent einen Stundensatz von 100 € oder weniger haben.

c) 66,667

2. In einer Studie soll untersucht werden, ob das während einer Prüfung konsumierte Getränk einen Einfluss auf den Prüfungserfolg hat. Die Ergebnisse sind in der untenstehenden Kontingenztabelle dargestellt.

	Prüfung bestanden	Prüfung nicht bestanden	**Summe**
Kaffee	60	40	
Wasser		30	
Energy Drink	30		40
Summe	120		

a) (2 Punkte) Ergänzen Sie die Kontingenztabelle.

b) (8 Punkte) Bestimmen Sie eine geeignete Kennzahl für den Zusammenhang zwischen dem während der Prüfung konsumierten Getränk und dem Prüfungserfolg und interpretieren Sie Ihr Ergebnis.

Lösung:

a)

	Prüfung bestanden	Prüfung nicht bestanden	**Summe**
Kaffee	60	40	100
Wasser	30	30	60
Energy Drink	30	10	40
Summe	120	80	200

b) 0,246 Zwischen dem während der Prüfung konsumierten Getränk und dem Prüfungserfolg besteht ein schwacher Zusammenhang.

3. Unter den Unternehmen, die einen Firmenkredit bei einer Bank aufnehmen und später Insolvenz anmelden, gibt es 30 % Kleinunternehmen (K) und 40 % Mittelstandsunternehmen (M). Der Rest sind Großkonzerne (G).

   In der Gruppe der Kleinunternehmen sind folgende Wahrscheinlichkeiten bekannt: Die Wahrscheinlichkeit beträgt 25 %, dass zum Zeitpunkt der Insolvenz eine Kreditsumme von weniger als 50.000 Euro aushaftet. Sie beträgt 40 % für eine Kreditsumme zwischen 50.000 Euro und 150.000 Euro. Der restliche Anteil entfällt auf Kreditsummen, die mit mindestens 150.000 Euro aushaften.

   Es gibt keine aushaftenden Kreditsummen über 500.000 Euro

   a) (4 Punkte) Vervollständigen Sie die Tabelle der bivariaten Verteilung der Unternehmensgruppen und der Höhe der aushaftenden Kreditsumme (in Tausend Euro).

	[0; 50[	[50; 150[	[150; 500]	Summe
K				
M		0,18	0,14	
G		0,09	0,12	
Summe				

   b) (2 Punkte) Wie lautet die Verteilung der ausfallenden Kredite in der Gruppe der Großkonzerne?

   c) (3 Punkte) Wie hoch ist die erwartete Höhe der im Insolvenzfall aushaftenden Kreditsumme näherungsweise?

   d) (3 Punkte) Wie hoch ist die erwartete Höhe der im Insolvenzfall aushaftenden Kreditsumme näherungsweise in der Gruppe der Großkonzerne?

Lösung:

a)

	[0; 50]	[50; 150]	[150; 500]	Summe
K	0,075	0,12	0,105	0,3
M	0,08	0,18	0,14	0,4
G	0,09	0,09	0,12	0,3
Summe	0,245	0,39	0,365	1

b)

X in Tsd	[0; 50]	[50; 150]	[150; 500]	Summe
$P(X)$	0,3	0,3	0,4	1

c) 163.750

d) 167.500

4. Eine stetige Zufallsgröße X ist gleichverteilt im Intervall $[0; M]$. Außerdem ist bekannt, dass $F(50) = 0{,}2$. (*Hinweis*: M ermitteln.)

   a) (3 Punkte) Bestimmen Sie das 70-%-Quantil.

   b) (3 Punkte) Bestimmen Sie $P(X \geq 40)$ und $P(X = 40)$.

   c) (3 Punkte) Veranschaulichen Sie die Aufgaben a) und b) in einer Skizze und zeichnen Sie auch $F(50)$ ein.

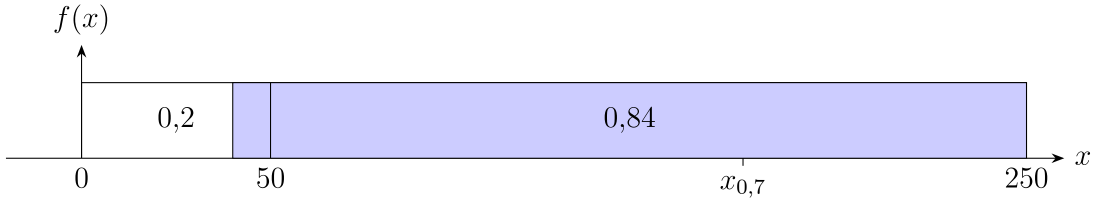
   d) (2 Punkte) Wie groß ist die Standardabweichung?

   Lösung:

   a) 175

   b) 0,84; 0

   c)



   d) 72,169

5. 28 Personen wurden befragt, wie viel sie für Schokoladeosterhasen ausgeben. Es wurde ein Mittelwert von Euro 25,– bei einer Stichprobenstandardabweichung von Euro 7,– berechnet. Die Ausgaben können als normalverteilte Zufallsgröße betrachtet werden.

   a) (5 Punkte) Bestimmen Sie ein zweiseitiges 95-%-Konfidenzintervall für den Erwartungswert und interpretieren Sie dieses.

   b) (10 Punkte) Lässt sich mit diesen Daten die Vermutung: „Der Erwartungswert der Ausgaben liegt bei höchstens Euro 23,–“, aufrechterhalten? Formulieren Sie die Hypothesen und führen Sie den entsprechenden Test zum Testniveau $\alpha = 1$ % durch. Was liefert dieser Test als Antwort?

   Lösung:

   a) $[22{,}285; 27{,}715]$ Mit einer statistischen Sicherheit von 95 % kann darauf vertraut werden, dass die Ausgaben für Schokoladeosterhasen im Mittel in diesem Intervall liegen.

   b) $H_0 : \mu \leq 23$ und $H_1 : \mu > 23$, $t_0 = 1{,}512$, $K = [2{,}473; \infty]$, $t_0 \notin K$, H_0 beibehalten. Die Vermutung lässt sich aufrechterhalten.

6. Unterscheidet sich der Konsum von Osterpinzen zwischen den Bundesländern Steiermark und Wien? Dazu wurde erhoben, wie viele Osterpinzen (in Stück à 240 g) Haushalte in der Steiermark und in Wien in der Osterzeit kaufen:

Steiermark (X)	4	8	5	3	6	2	7
Wien (Y)	9	15	7	9	4	13	12

Testen Sie ohne Annahme von Normalverteilung, ob in der Steiermark signifikant weniger Osterpinzen verzehrt werden als in Wien.

a) (2 Punkte) Welcher Test kommt zur Anwendung? Begründen Sie Ihre Entscheidung.

b) (2 Punkte) Formulieren Sie die Hypothesen.

c) (6 Punkte) Berechnen Sie die Testgröße.

d) (3 Punkte) Bestimmen Sie den p-Wert.

e) (2 Punkte) Werden in der Steiermark signifikant weniger Osterpinzen verzehrt als in Wien? Wählen Sie ein Signifikanzniveau und begründen Sie Ihre Antwort.

Lösung:

a) Wilcoxon-Rangsummentest, Test auf Lage, 2 Stichproben, unverbunden, nicht normalverteilt

b) $H_0 : c \leq 0$, $H_1 : c > 0$

c) $t_0 = -2{,}364$

d) 0,009

e) H_1 bestätigt. Ja.

7. In einer Studie soll untersucht werden, ob die Anzahl der Schlafstunden pro Nacht einen Einfluss auf die Punktzahl in einem Konzentrationstest hat. Es wird ein einfaches lineares Regressionsmodell der Form

$$Y = \beta_0 + \beta_1 x + \varepsilon$$

angenommen, wobei

- Y: Punktzahl in einem standardisierten Konzentrationstest,
- X: durchschnittliche Schlafdauer pro Nacht (in Stunden),
- ε: Störterm.

Es wurden $n = 7$ Personen untersucht und deren Schlafdauer gemessen. Auf Basis dieser Daten ergeben sich die geschätzten Regressionskoeffizienten $\hat{\beta}_0 = -1{,}2$, $\hat{\beta}_1 = 2{,}1$. Die geschätzte Standardabweichung des Störterms beträgt $\hat{\sigma} = 2{,}1$.

Weiters wurde $\sum\limits_{i=1}^{7} (x_i - \overline{x})^2 = 28$ berechnet.

a) (2 Punkte) Berechnen Sie die Standardabweichung des Koeffizienten $\hat{\beta}_1$.

b) (5 Punkte) Testen Sie zum Niveau $\alpha = 0{,}05$, ob die Schlafdauer einen signifikanten Einfluss auf die Punktzahl im Konzentrationstest hat.

c) (3 Punkte) Betrachten Sie zwei Personen:

Person A schläft im Durchschnitt 5 Stunden pro Nacht.

Person B schläft im Durchschnitt 8 Stunden pro Nacht.

Bestimmen Sie mit Hilfe der geschätzten Regressionsgeraden die Differenz der erwarteten Punktzahlen und interpretieren Sie diese im Kontext der Studie.

Lösung:

a) 0,397

b) $H_0 : \hat{\beta}_1 = 0$, $H_1 : \hat{\beta}_1 \neq 0$, $t_0 = 5{,}292$, $K =]-\infty; -2{,}571[\cup]2{,}571; \infty[$, $t_0 \in K$; Ja.

c) 6,3; Im Rahmen des geschätzten linearen Regressionsmodells wird erwartet, dass eine Person, die durchschnittlich 8 Stunden pro Nacht schläft, im Konzentrationstest etwa 6,3 Punkte mehr erzielt als eine Person, die durchschnittlich 5 Stunden pro Nacht schläft.

8. (10 Punkte)

a) Ergänzen Sie die Textlücken im nachstehenden Satz durch Ankreuzen des jeweils zutreffenden Ausdrucks aus den nachstehenden Tabellen so, dass jedenfalls eine wahre Aussage entsteht.

Der Rangkorrelationskoeffizient ist geeignet, den Zusammenhang zwischen einem ____ 1 ____ und einem ____ 2 ____ Merkmal zu messen.

1	2
☐ nominalen	☐ dichotomen
☐ dichotomen	☐ metrischen
☐ ordinalen	☐ nominalen

b) Gegeben sei eine diskrete Zufallsvariable X mit Verteilungsfunktion $F_X(x)$:
Wie lässt sich die Wahrscheinlichkeit $P(X = 2 \text{ oder } X = 3)$ mithilfe von $F_X(x)$ ermitteln? Kreuzen Sie die zutreffende Aussage an.

- ◯ $F_X(2) + F_X(3)$
- ◯ $F_X(3) - F_X(2)$
- ◯ $F_X(3)$
- ◯ $F_X(3) - F_X(1)$
- ◯ Nicht über $F_X(x)$ lösbar.

c) Ein Taxiunternehmen beschreibt den Fahrpreis (in Euro) in Abhängigkeit von der gefahrenen Strecke (in km) und der Tageszeit (Tag/Nacht) mit folgender Regressionsgleichung: $\hat{Y} = 5 + 2 \cdot \mathit{Strecke} + 3 \cdot \mathit{Nacht}$. Demnach beträgt die Grundgebühr in der Nacht (also für eine Strecke von 0 km) ☐ Euro.

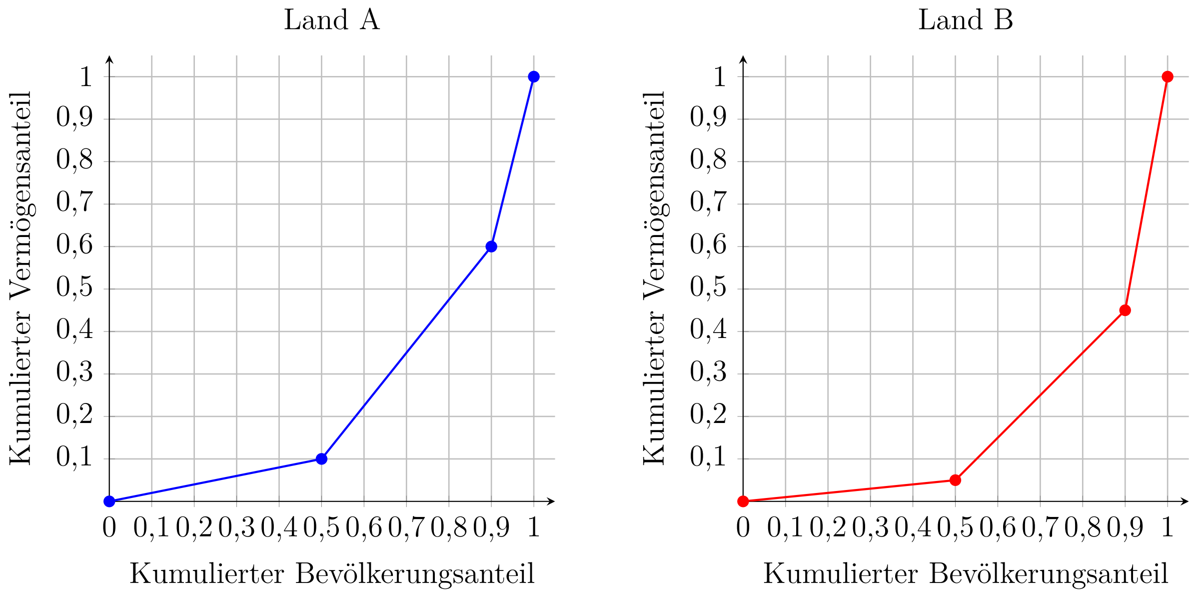
d) Für eine normalverteilte Zufallsvariable $X \sim N(\mu; \sigma)$ erfolgt die Standardisierung durch $Z = \frac{X-\mu}{\sigma}$. Welches Quantil z der Standardnormalverteilung entspricht einem Wert x der Zufallsvariable X, der genau drei Standardabweichungen oberhalb des Erwartungswerts μ liegt? ☐

e) Kreuzen Sie die beiden zutreffenden Aussagen an.

- ◯ Wird die Nullhypothese nicht verworfen, gilt sie als bewiesen.
- ◯ Ein kleiner p-Wert spricht gegen die Nullhypothese.
- ◯ Wird das Siginifikanzniveau α verkleinert, wird der kritische Bereich kleiner.
- ◯ Wird das Signifikanzniveau α erhöht, sinkt die Wahrscheinlichkeit eines Fehlers 1. Art.
- ◯ Ist der p-Wert größer als α, so liegt der Testwert im kritischen Bereich.

Lösung:
(ordinalen, metrischen); 4; 8; 3; (2, 3)

9. (10 Punkte) Für zwei Länder sind die Lorenzkurven für das Vermögen der Bevölkerung gegeben:



Beantworten Sie die folgenden Fragen:

a) Welchen Anteil am Gesamtvermögen besitzt die Mittelschicht in Land A? ☐

b) Der Anteil der Mittelschicht an der Gesamtbevölkerung beträgt in beiden Ländern ☐ %.

c) Ergänzen Sie die Textlücke im nachstehenden Satz, indem Sie die zutreffende Aussage ankreuzen:

Der Anteil am Gesamtvermögen, den die unteren 50 % der Bevölkerung besitzen, ist in Land A ☐ in Land B.

◯ geringer als ◯ höher als ◯ gleich groß wie

d) Kreuzen Sie die zutreffende Aussage an:

◯ Die Gini-Koeffizienten beider Länder sind gleich groß.
◯ Der Gini-Koeffizient von Land A ist größer.
◯ Der Gini-Koeffizient von Land B ist größer.

e) Kreuzen Sie die richtige Aussage an.

Die reichsten 10 % der Bevölkerung besitzen mehr als 50 % des Gesamtvermögens.

◯ nur in Land A
◯ nur in Land B
◯ in beiden Ländern
◯ in keinem Land

Lösung:
0,5; 40; 2; 3; 2