Statistik Vorlesung

28. November 2025
Arbeitszeit: 100 Minuten

VORNAME:		MATR.NR.:	
NACHNAME:			

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

VERBOTEN: **alle sonstigen Unterlagen, Handys**

Aufgabe 1 und 2: Für jede vollständig richtig gelöste Frage gibt es 2 Punkte. Es gibt keine Teilpunkte und keine Minuspunkte.

Lösungswege müssen nachvollziehbar aufgeschrieben sein.

Aufgabe	max. Punkte	erreichte Punkte	Teilsumme
1	10		
2	10		
9	14		
3	10		
4	13		
5	08		
6	12		
7	11		
8	12		
Summe	100		
Note			

1. (10 Punkte)

a) Markieren Sie, ob die folgende Aussage zutrifft oder nicht, und kennzeichnen Sie die jeweils passende Begründung!

Bei einem Boxplot liegen typischerweise die Hälfte aller Beobachtungen außerhalb der Box.

◯ Richtig ◯ Falsch

Begründung:

- ☐ Die Box reicht vom ersten zum dritten Quartil und enthält daher die anderen 50 % der Werte des Datensatzes.
- ☐ Außerhalb der Box liegen nur (potentielle) Ausreißer.
- ☐ Die Aussage stimmt nur unter der Bedingung, dass die Daten symmetrisch sind.
- ☐ Die Aussage stimmt nur unter der Bedingung, dass die Daten normalverteilt sind.

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

Bei der Untersuchung zweier Merkmale ergibt sich, dass die Ausprägungen des einen Merkmals reelle Zahlen sind, während bei den Ausprägungen des zweiten Merkmals nur die Gleichheit oder Verschiedenheit festgestellt werden kann. Das geeignete Zusammenhangsmaß ist in diesem Fall der ___(1)___, weil das ___(2)___ Skalenniveau bei der Wahl des Zusammenhangsmaßes maßgebend ist.

(1)	(2)
☐ Pearson-Korrelationskoeffizient	☐ niedrigere
☐ Spearman-Korrelationskoeffizient	☐ gleiche
☐ korrigierte Kontingenzkoeffizient	☐ höhere

c) Eine Zufallsvariable X sei $N(25, 4)$-verteilt. Kennzeichnen Sie die beiden richtigen Aussagen!

- ◯ $P(X = 25) = 0{,}5$.
- ◯ $P(X = 25) = 0$.
- ◯ $P(X = 25)$ ist die Dichte der Standardnormalverteilung an der Stelle 25.
- ◯ Die Verteilung von X ist symmetrisch zu $x = 0$.
- ◯ $F(25) = \Phi(0)$.

d) Markieren Sie, ob die folgende Aussage zutrifft oder nicht, und kennzeichnen Sie die jeweils passende Begründung!

Die Behauptung, X sei $B(15; 0{,}3)$-verteilt, wird mit dem Binomialtest überprüft.

◯ Richtig ◯ Falsch

Begründung:

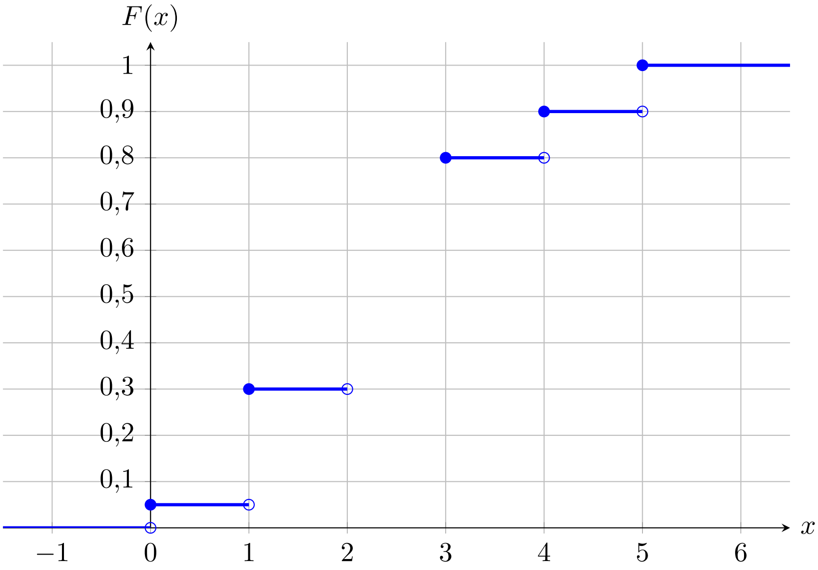
- ☐ Der Binomialtest testet die Binomialverteilung.
- ☐ Der Binomialtest testet $H_1 : p = 0{,}3$.
- ☐ Ein anderer Test testet auf Binomialverteilung.

e) Ergänzen Sie die Textlücke im nachstehenden Satz so, dass jedenfalls eine wahre Aussage entsteht!

Im Rahmen einer Regressionsanalyse sollen die Modellannahmen überprüft werden. Wenn in der entsprechenden Grafik bei der Analyse der ☐ kein systematisches Muster zu erkennen ist, deutet dies darauf hin, dass die Regression korrekt spezifiziert wurde.

Lösung: (R, 1), korr K niedrigere, (2, 5), (F, 3), Residuen

2. (10 Punkte) Im Rahmen einer Umfrage auf dem Herbstmarkt wurden 40 Besucherinnen und Besucher gefragt, wie häufig sie in dieser Saison schon den Maronistand besucht haben. Die Ergebnisse dieser Umfrage werden in folgender Verteilungsfunktion dargestellt.



a) Die Umfrage ergab, dass 12 Personen den Maronistand genau zweimal besucht haben. Ergänzen Sie die dargestellte Verteilungsfunktion so, dass auch diese Information berücksichtigt wird.

b) Wie viel Prozent der Befragten haben den Maronistand mehr als dreimal besucht? $\boxed{}$

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

5 % der Befragten haben den Maronistand ___(1)___ ___(2)___ besucht.

(1)	(2)
☐ genau	☐ einmal
☐ weniger als	☐ zweimal
☐ mehr als	☐ dreinmal

d) Ergänzen Sie die Textlücke im nachstehenden Satz so, dass jedenfalls eine wahre Aussage entsteht!

25 % der befragten Personen gaben an, genau $\boxed{}$-mal den Maronistand besucht zu haben.

e) Zwei weitere Personen werden befragt und geben an genau 3 Mal den Maronistand besucht zu haben. Wie verändert sich die Verteilungsfunktion, wenn die Stichprobe um die Antworten dieser beiden Personen ergänzt wird? Kreuzen Sie die zutreffende Aussage an!

◯ Die Verteilungsfunktion bleibt unverändert, da der Wert 3 bereits vorkam.

◯ Die Verteilungsfunktion bleibt für alle Werte $x < 3$ unverändert.

◯ Die Verteilungsfunktion steigt nur exakt bei $x = 3$ und bleibt für alle $x > 3$ konstant.

◯ Die Verteilungsfunktion steigt für alle $x \geq 3$ und kann somit auch Werte größer 1 annehmen.

◯ Keine der Antwortmöglichkeiten ist richtig.

Lösung: 0,6 für $x \in [2; 3[$, 20, weniger als einmal, 1, 5

3. „Wie viel Zeit wenden Sie in einer Woche für Hausarbeit auf?" Folgende Daten wurden erhoben:

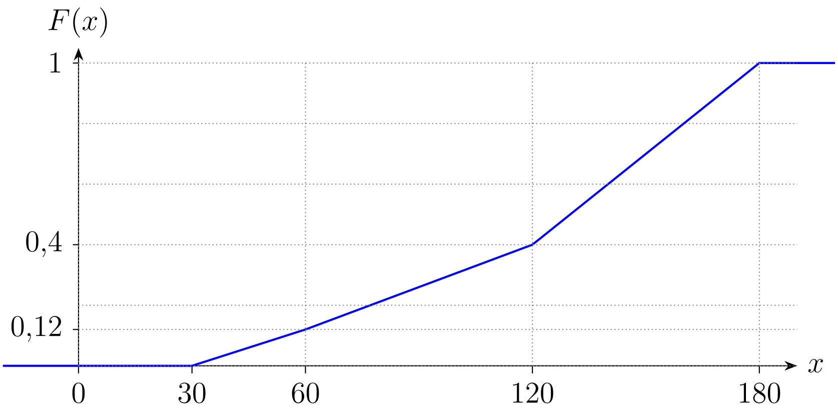
Zeit für Hausarbeit (in Minuten) [von; bis [	[30; 60[	[60; 120[	[120; 180]
Absolute Häufigkeit	36	84	180

a) (5 Punkte) Zeichnen Sie die geeignete Verteilungsfunktion! Bestimmen und interpretieren Sie den Wert $F(120)$!

b) (5 Punkte) Berechnen Sie näherungsweise das arithmetische Mittel sowie die Standardabweichung!

Ausführung Beispiel 3: Lösung:

a)



$F(120) = 0{,}4$; 40 % wenden höchstens 2 Stunden pro Woche für Hausarbeit auf.

b) $\overline{x} = 120{,}6$; $\sigma = 38{,}297$

4. a) (3 Punkte) Ein Forscher misst an einem Tag bei wolkenlosem Himmel am Schöckl die Strahlungsintensität I der Sonne in Abhängigkeit der Uhrzeit T und erhält folgende Ergebnisse:

Uhrzeit T	7	9	11	13	15	17
Intensität I (Watt/m^2)	103	345	571	619	367	132

Aus diesen Daten wurde der Pearson-Korrelationskoffizient berechnet und ergab: $korr(X, Y) = 0{,}065$.

Der Forscher behauptet in der Folge, dass die Strahlungsintensität der Sonne nicht von der Tageszeit abhängig ist. Welcher grundlegende Fehler könnte bei der Interpretation der Daten gemacht worden sein?

b) (10 Punkte) Dieser Forscher vermutet auch einen Zusammenhang zwischen Sonneneinstrahlung und der Leistung bei der Stromerzeugung und hat folgende Daten ermittelt:

Sonneneinstrahlung (Watt/m^2)	600	200	300	500
Leistung (Watt)	175	45	90	130

Ist nun diese Vermutung gerechtfertigt? Berechnen Sie eine geeignete Kennzahl und interpretieren Sie diese!

Ausführung Beispiel 4:

Lösung:

a) Intensitäten steigen nicht linear im Tagesverlauf mit der Uhrzeit.

b) Ja, gerechtfertigt. Die Daten sprechen nicht gegen einen linearen Zusammenhang.
$\rho = 0{,}986$; Starker positiver linearer Zusammenhang. Je intensiver die Sonnenstrahlung desto tendenziell höher ist die Leistung.

5. Zwei Studienfreunde treten bei einer Multiple-Choice-Klausur an, die aus 15 Fragen mit den Antwortkategorien „Ja“ und „Nein“ besteht. Die zwei Studienfreunde verfolgen bei der Klausur unterschiedliche Strategien: Student A hat nichts gelernt und rät die Antworten auf die Fragen; Student B hat sich gut vorbereitet und kann daher jede Frage mit einer Wahrscheinlichkeit von 80 % korrekt beantworten.

   a) (2 Punkte) Mit welcher Wahrscheinlichkeit beantwortet Student A die ersten drei Fragen der Klausur richtig?

   b) (3 Punkte) Wie viele Pfade des Wahrscheinlichkeitsbaumes gibt es, bei denen Student A genau drei Klausurfragen falsch beantwortet? Berechnen Sie die Anzahl, das Zeichnen des Wahrscheinlichkeitsbaumes ist nicht notwendig.

   c) (3 Punkte) Mit welcher Wahrscheinlichkeit beantwortet Student B trotz guter Vorbereitungen genau ein Drittel der Klausurfragen falsch?

Ausführung Beispiel 5:

Lösung:

a) 0,125

b) 455

c) 0,103

6. Ein Unternehmen stellt Glühbirnen her, deren Lebensdauer (in Stunden) normalverteilt mit einem Erwartungswert von $\mu = 1.000$ Stunden und einer Standardabweichung von $\sigma = 100$ Stunden ist.

   a) (4 Punkte) Skizzieren Sie die Dichtefunktion dieser Normalverteilung, tragen Sie die beiden Parameter, den Erwartungswert und die Standardabweichung, ein! Erklären Sie zudem, wie diese beiden Parameter die Form und Lage der Normalverteilung beeinflussen!

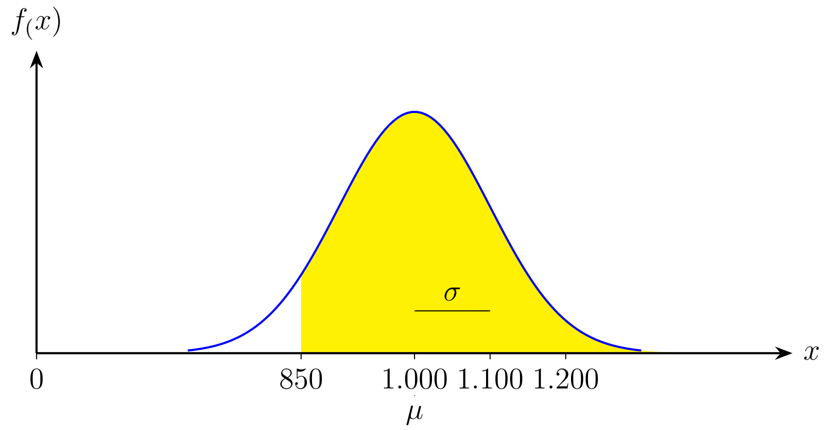
   b) (3 Punkte) Wie groß ist die Wahrscheinlichkeit, dass eine zufällig ausgewählte Glühbirne eine Lebensdauer von mehr als 850 Stunden hat? Berechnen Sie die Wahrscheinlichkeit und tragen Sie das Ergebnis in die skizzierte Dichtefunktion ein!

   c) (5 Punkte) Die Glühbirnen werden in Aktion in Dreierpackungen angeboten. Ein Kunde möchte Glühbirnen kaufen, die eine Lebensdauer von mindestens 850 Stunden haben. Mit welcher Wahrscheinlichkeit haben in der Dreierpackung mindestens zwei der drei Glühbirnen eine Lebensdauer von jeweils mehr als 850 Stunden?

Ausführung Beispiel 6:

Lösung:

a)



Eine Änderung von μ verschiebt die Verteilung entlang der x-Achse, ohne dass sich die Form der Verteilung verändert. Eine Änderung von σ beeinflusst die Breite und Höhe der Verteilung.

b) $P(X > 850) = 0{,}933$

c) 0,987

7. Bei einer Erhebung zu den Auswirkungen der Steuerreform wurden insgesamt 1.400 Personen befragt. Davon gaben 380 Personen an, durch die Reform profitiert zu haben.
   a) (4 Punkte) Bestimmen Sie ein zweiseitiges 90-%-Konfidenzintervall für den Anteil der Personen, die durch die Reform profitiert haben!
   b) (2 Punkte) Wird die Aussage „30 % aller Personen werden durch die Steuerreform profitieren“ durch das in a) berechnete Konfidenzintervall unterstützt? Begründen Sie Ihre Antwort!
   c) (5 Punkte) Wie groß müsste die Stichprobe sein, um den Anteil der Personen, die durch die Reform profitiert haben, mit einem zweiseitigen 90-%-Konfidenzintervall mit einer Genauigkeit von ± 1 Prozentpunkt zu schätzen, wenn es keine Vorabinformation über den Anteil gibt?

   Ausführung Beispiel 7:

   Lösung:

   a) $[0{,}252; 0{,}291]$
   b) Die Aussage „30 % aller Personen profitieren“ liegt damit außerhalb des Intervalls und wird somit nicht gestützt.
   c) $n = 6.766$

8. Ein Hersteller von Elektroautos möchte prüfen, ob eine neue Software-Version die Reichweite der Fahrzeuge verbessert hat.

   Bei einer Stichprobe von 8 Autos mit der **alten** Software ergab sich eine durchschnittliche Reichweite pro Akkuladung von $\overline{x}_1 = 407{,}8$ km bei einer Stichprobenstandardabweichung von $s_1 = 4{,}50$ km.

   Bei einer Stichprobe von 6 Autos mit der **neuen** Software wurde eine durchschnittliche Reichweite pro Akkuladung von $\overline{x}_2 = 414$ km bei einer Stichprobenstandardabweichung von $s_2 = 5{,}38$ km gemessen.

   Gehen Sie davon aus, dass die Messwerte Realisierungen unabhängiger Normalverteilungen sind.

   a) (2 Punkte) Es wurde ein Test durchgeführt, um zu prüfen, ob die Reichweite mit der **alten** Software signifikant über 400 km liegt. Angenommen, die Grenze des kritischen Bereiches lautet 1,895. Wie groß ist die Wahrscheinlichkeit für den Fehler 1. Art?

   b) (10 Punkte) Lässt sich mit einem geeigneten Test ($\alpha = 0{,}01$) zeigen, dass die **neue** Software-Version die durchschnittliche Reichweite gegenüber der **alten** Software-Version tatsächlich erhöht hat? Gehen Sie dabei davon aus, dass sich die Varianz durch die Software-Änderung nicht verändert hat.

   Ausführung Beispiel 8:

   Lösung:

   a) $t_{7;1-\alpha} = 1{,}895 \Rightarrow \alpha = 0{,}05$

   b) Zweistichproben-t-Test; $t_0 = 2{,}35$, $K =]2{,}681; \infty[$ H_1 nicht bestätigt. Nein.

9. Die drei Studierenden Agnes, Bernd und Eva bekommen das folgende Beispiel zur Statistikprüfung:
Ein Unternehmen möchte den monatlichen Absatz eines Produkts (in Tausend Stück) mithilfe zweier Einflussfaktoren erklären

   - x_1: Preis des Produkts (in Euro)
   - x_2: Werbeausgaben (in Tausend Euro)

   und verwendet das Modell

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

   mit y_i = Absatz (in Tsd. Stück).

   Vier Beobachtungen wurden erfasst:

Beobachtung	Preis x_1	Werbung x_2	Absatz y
1	2	1	2
2	3	3	4
3	4	4	6
4	5	6	9

   a) Agnes rechnet das Beispiel und erhält: $\hat{\beta}_0 = -2$, $\hat{\beta}_1 = 1{,}5$ und $\hat{\beta}_2 = 0{,}5$.
      i. (5 Punkte) Bestimmen Sie die Residuen für alle Beobachtungen und die Summe der quadrierten Residuen (SSE)!
      ii. (2 Punkte) Bernd erhält bei seiner Berechnung andere Werte für die Koeffizienten $\hat{\beta}_i$. Mit Hilfe dieser Werte berechnet er SSE fehlerfrei und erhält den Wert 0,25.
      Wir wissen, dass einer von den beiden die Koeffizienten richtig berechnet hat. Wer war es? Begründen Sie Ihre Antwort!
   b) (3 Punkte) Eva schummelt und weiß, dass $\hat{\beta}_1 = 1{,}5$ und $\hat{\beta}_2 = 0{,}5$ richtig ist und das Residuum der ersten Beobachtung 0,25 ist.
   Berechnen Sie mit Hilfe dieser Information den korrekten Wert für $\hat{\beta}_0$!
   c) (4 Punkte) Agnes schafft die Prüfung nicht und rechnet das Beispiel noch einmal. Zusätzlich zum Koeffizienten $\hat{\beta}_1 = 1{,}5$ berechnet sie auch dessen Standardfehler und erhält 1,803.
      i. Welchen Testwert erhält sie, wenn Sie testen möchte, ob sich β_1 signifikant von null unterscheidet?
      ii. Ist das Merkmal „Preis des Produkts“ auf dem 5-%-Niveau signifikant?

   Hinweis: Die drei Unterpunkte 9a, 9b und 9c sind unabhängig voneinander lösbar.

   Ausführung Beispiel 9:

   Lösung:

   a) i. $\hat{r}_i$: 0,5   0   0   0,5 SSE : 0,5
      ii. Da die Summe der quadrierten Residuen bei Bernd kleiner ist, hat hat er die Koeffizienten richtig berechnet.
   b) Es gilt $\hat{\beta}_0 = 2 - (1{,}5 \cdot 2 + 0{,}5 \cdot 1 + 0{,}25) = -1{,}75$.
   c) i. $t_0 = \frac{1{,}5}{1{,}803} \approx 0{,}832$
      ii. $K =]-\infty; -12{,}706[\cup]12{,}706; \infty[;\ t_0 \notin K$
      Das Merkmal „Preis des Produkts“ ist auf dem 5-%-Niveau nicht signifikant.