

Wiederholung, Ergänzung, Erklärung & Intuition:
Statistik I & II für Studierende der Wirtschaftswissenschaften
(Ludwig-Maximilians-Universität München)

Autoren:

Matthias Aßenmacher* Ann-Kathrin Köpple† Christoph Luther‡
Patricia Haro§ Maximilian Mandl¶

Stand: July 6, 2021

Dieses Dokument enthält die Lösungen zu den Beispielaufgaben.

*Institut für Statistik, LMU München; Kontakt bei Fragen & Anregungen: matthias@stat.uni-muenchen.de

†Studentische Hilfskraft (SoSe20), Institut für Statistik, LMU München

‡Studentische Hilfskraft (WiSe 18/19 - SoSe 20), Institut für Statistik, LMU München

§Studentische Hilfskraft (WiSe 18/19 - WiSe 19/20), Institut für Statistik, LMU München

¶Institut für medizinische Informationsverarbeitung, Biometrie und Epidemiologie, LMU München

This work is licensed under a Attribution-NonCommercial 4.0 International (CC BY-NC 4.0)

License details: <https://creativecommons.org/licenses/by-nc/4.0/>

Contents

1	Grundbegriffe, Skalenniveaus, Datenerhebung	5
1.1	Aufgaben	5
2	Häufigkeitsverteilungen, (univariate) grafische Darstellung	7
2.1	Aufgaben	7
3	Lagemaße	8
3.1	Aufgaben	8
4	Streuungsmaße	9
4.1	Aufgaben	9
5	Konzentrationsmaße	10
5.1	Aufgaben	10
6	Zusammenhangsmaße	11
6.1	Aufgaben	11
7	Lineare Einfachregression	13
7.1	Aufgaben	13
8	Indizes	15
8.1	Aufgaben	15
9	Zeitreihen	16
9.1	Aufgaben	16
10	R-Einführung Teil I	17
11	Kombinatorik	18
11.1	Aufgaben	18
12	Wahrscheinlichkeitsrechnung	20
12.1	Aufgaben	20
13	Zufallsvariablen	22
13.1	Aufgaben	22
14	Spezielle Verteilungen	23
14.0.1	Aufgaben I	23
14.0.2	Aufgaben II	25
14.0.3	Aufgaben III	26

15 Grenzwertsätze und Approximationen von Verteilungen	27
15.1 Aufgaben	27
16 Schätzen	28
16.1 Aufgaben	28
17 Testtheorie	29
17.1 Aufgaben	29
18 Lineare Regression II	31
18.1 Aufgaben	31
19 R-Einführung Teil II	33

Statistik I

1 Grundbegriffe, Skalenniveaus, Datenerhebung

1.1 Aufgaben

1. Datenerhebung: Welche Aussagen sind wahr?

- (a) Längsschnittdaten werden stets durch die Erhebungsmethode "Beobachtung" gewonnen. ☐
- (b) Bei Längsschnittdaten müssen die zeitlichen Abstände identisch sein. ☐
- (c) Längsschnittdaten entstehen durch die wiederholte Erhebung von Querschnittdaten. ☒
- (d) Von einer Zeitreihe spricht man, wenn ein Merkmal der Reihe nach zu festgelegten Zeitpunkten an unterschiedlichen Untersuchungseinheiten beobachtet wird. ☐
- (e) Keine der obigen Aussagen ist wahr. ☐

2. Welche Aussagen zu Skalenniveaus sind wahr?

- (a) Wenn man die Anzahl der möglichen Ausprägungen zählen kann, ist ein Merkmal metrisch skaliert. ☐
- (b) Ein ordinalskaliertes Merkmal besitzt mehr mögliche Ausprägungen als ein nominalskaliertes Merkmal. ☐
- (c) Im Gegensatz zu ordinalskalierten Merkmalen kann man metrische Merkmale ordnen. ☐
- (d) Nominalskalierte Merkmale können auch stetig sein. ☐
- (e) Keine der obigen Aussagen ist wahr. ☒

3. Welche Aussagen zu Skalenniveaus sind wahr?

- (a) Die Existenz eines natürlichen Nullpunkts ist Voraussetzung für eine Absolutskala. ☒
- (b) Die „Intervallskala“ hat einen höheren Informationsgehalt als die „Verhältnisskala“. ☐
- (c) Geschlecht ist ein Beispiel für ein ordinalskaliertes Merkmal. ☐
- (d) Quasistetig bedeutet, dass ein Merkmal theoretisch diskret ist, aber wie ein stetiges Merkmal behandelt wird. ☐
- (e) Keine der obigen Aussagen ist wahr. ☐

2 Häufigkeitsverteilungen, (univariate) grafische Darstellung

2.1 Aufgaben

1. Häufigkeitsbegriffe und Visualisierungen: Welche Aussagen sind wahr?

- (a) Für stetige oder quasistetige Merkmale sollten bei der Häufigkeitsauszählung Klassen gebildet werden. ☒
- (b) Absolute Häufigkeiten summieren sich stets zu 1 auf. ☐
- (c) Balken-/Säulendiagramme sind gut für nominale/ordinale Merkmale geeignet. ☒
- (d) Kreisdiagramme sind für ordinale Merkmale eher ungeeignet. ☒
- (e) Keine der obigen Aussagen ist wahr. ☐

2. Welche Aussagen zu Histogrammen sind wahr?

- (a) Bei ungleichen Klassenbreiten ist die Höhe des Balkens proportional zur rel. Häufigkeit. ☐
- (b) Histogramme sind gut für nominale/ordinale Merkmale geeignet. ☐
- (c) Ein Histogramm mit konstanter Klassenbreite ist dasselbe wie ein Säulendiagramm. ☐
- (d) Je gröber die Klasseneinteilung, desto höher die einzelnen Balken. ☐
- (e) Keine der obigen Aussagen ist wahr. ☒

3. Welche Aussagen sind zur Empirische Verteilungsfunktion sind wahr?

- (a) Für ein stetiges Merkmal gilt: $H(x < 5) = H(x \leq 4)$. ☐
- (b) Die kumulierten relativen Häufigkeiten summieren sich zu 1. ☐
- (c) Für diskrete Merkmale: Bei der häufigsten Beobachtung ist die emp. Verteilungsfunktion am größten. ☐
- (d) Für diskrete Merkmale: Bei der häufigsten Beobachtung macht die emp. Verteilungsfunktion den größten Sprung. ☒
- (e) Keine der obigen Aussagen ist wahr. ☐

3 Lagemaße

3.1 Aufgaben

1. Unterschiede zwischen Mittelwert und Median?

- a) Mittelwert ist robuster ggü. Ausreißern ☐
- b) Median ist robuster ggü. Ausreißern ☒
- c) Keine ☐

2. Der Median ..

- a) .. liegt immer genau in der Mitte der Box. ☐
- b) .. entspricht dem 50%-Quantil. ☒
- c) .. entspricht dem 2. Quartil. ☒
- d) .. ist wichtig dafür, zu berechnen wann ein Wert ein Ausreißer ist. ☐

3. Welche Mittelung ist geeignet, um den durchschnittlichen Anstieg der Transferausgaben in der Fußballbundesliga zu ermitteln?

- a) Arithmetisches Mittel ☐
- b) Geometrisches Mittel ☒
- c) Harmonisches Mittel ☐
- d) Alle drei machen Sinn ☐

4 Streuungsmaße

4.1 Aufgaben

1. Bei welcher Maßzahl werden hohe Abweichungen vom Mittelwert stärker gewichtet?

- a) MAD ☐
- b) Varianz ☒
- c) Bei beiden gleich stark ☐

2. Welche Aussagen zur Streuungszerlegung sind wahr?

- a) Die Varianz innerhalb der Gruppen ist immer größer als zwischen den Gruppen. ☐
- b) Man kann die Varianz innerhalb und zwischen den Gruppen einfach addieren um die Gesamtvarianz zu erhalten. ☒
- c) Es gibt Sonderfälle, bei denen die Streuung zwischen den Gruppen der Gesamtstreuung entspricht. ☒
- d) Es muss immer eine Streuung innerhalb der Gruppen vorliegen. ☐

3. Der Verschiebungssatz ..

- a) .. erleichtert die Berechnung des arithmetischen Mittels. ☐
- b) .. kann auch bei gruppierten Daten verwendet werden. ☒
- c) .. dient zur Berechnung des arithmetischen Mittels der quadrierten Daten. ☐
- d) .. benötigt das arithmetische Mittel der quadrierten Daten. ☒

4. Welche der folgenden Aussagen zum Variationskoeffizienten sind wahr?

- a) Der Variationskoeffizient ermöglicht den Vergleich von Streuungen von Merkmalen, die in verschiedenen Einheiten gemessen werden. ☒
- b) Der Variationskoeffizient ermöglicht den Vergleich von Streuungen von Merkmalen, die in verschiedenen Größenordnungen liegen. ☒
- c) Für die Berechnung des Variationskoeffizienten müssen beide Merkmale in der gleichen Einheit vorliegen. ☐
- d) Zur Berechnung des Variationskoeffizienten benötigt man den Median. ☐

5 Konzentrationsmaße

5.1 Aufgaben

1. Welche Aussagen bzgl. Gini & Lorenzkurve sind wahr?

- a) Die absolute Merkmalssumme ist unerheblich für den Gini. ☒
- b) Höherer Gini bedeutet (global) steilere Lorenzkurve. ☐
- c) Der Gini ist uneingeschränkt geeignet um die Konzentration in zwei Gruppen zu vergleichen. ☐
- d) Erhalten alle Merkmalsträger dieselbe prozentuale Steigerung ihres (absoluten) Teils der Merkmalssumme, so verändert sich der Gini nicht. ☒

2. Welche Aussagen bzgl. des Herfindahl-Index sind wahr?

- a) Der Herfindahl-Index ist uneingeschränkt geeignet um die Konzentration in zwei Gruppen zu vergleichen. ☐
- b) Falls sich die Merkmalssumme ändert, können definitive Aussagen über der Änderung des Herfindahl-Index getroffen werden. ☐
- c) Falls sich die Verteilung Merkmalssumme ändert, können definitive Aussagen über der Änderung des Herfindahl-Index getroffen werden. ☒
- d) Höherer Herfindahl-Index bedeutet ungleichere Verteilung. ☒

3. Der Gini für gruppierte Daten ist nur identisch zum "normalen" Gini, falls ..

- a) .. alle Gruppen gleich groß sind. ☐
- b) .. absolute Gleichverteilung herrscht. ☐
- c) .. Gleichverteilung innerhalb der Gruppen herrscht. ☒
- d) .. die Anzahl der Gruppen kleiner als 10 ist. ☐

6 Zusammenhangsmaße

6.1 Aufgaben

1. Welche Aussagen bzgl. relativem Risiko & Odds Ratio sind wahr?

- a) Bei beiden Maßzahlen werden zwei Gruppen verglichen. ☒
- b) Der Odds Ratio kann auf Basis von relativen Risiken berechnet werden. ☒
- c) Beim relativen Risiko werden zwei Risikomerkmale verglichen. ☐
- d) Ein Odds Ratio < 0 bedeutet eine geringere Chance in der ersten Gruppe. ☐

2. Die unter Unabhängigkeit erwarteten absoluten Häufigkeiten ..

- a) .. müssen stets ganzzahlig sein. ☐
- b) .. können ohne Kenntnis der gemeinsamen Verteilung berechnet werden. ☒
- c) .. sind identisch zu der bedingten Verteilung. ☐
- d) .. sind maximal so hoch wie die tatsächlich beobachteten Häufigkeiten. ☐

3. Welche der folgenden Aussagen über Zusammenhangsmaße für nominale Merkmale sind wahr?

- a) Mit Cramers V sind Zusammenhänge für Kontingenztafeln von verschiedener Dimension und mit unterschiedlichem n vergleichbar. ☒
- b) Φ besitzt einen kleineren Wertebereich als χ^2 . ☒
- c) Um Kontingenztafeln mit dem korrigierten Kontingenzkoeffizienten vergleichen zu können muss deren Dimension gleich sein. ☐
- d) Kein Zusammenhangsmaß für nominale Merkmale kann negative Werte annehmen. ☒

4. Welche der folgenden Aussagen sind wahr?

- a) Rang-basierte Zusammenhangsmaße sind bei metrischen Merkmalen nicht anwendbar. ☐
- b) Das Prinzip der Kon-/Diskordanz kann auch bei nominalen Merkmalen angewendet werden. ☐
- c) Zusammenhangsmaße für nominale Merkmale können nicht bei ordinalen oder metrischen Merkmalen verwendet werden. ☐
- d) Zusammenhangsmaße für nominale Merkmale können keine Richtung des Zusammenhangs angeben. ☒

5. Bindungen in Y ..

- a) .. sprechen für einen negativen Zusammenhang. ☐
- b) .. haben keinen Einfluss auf den Wert von γ . ☒
- c) .. erhöhen den Wert von Kendalls τ_b . ☐
- d) .. haben einen Einfluss auf den Wert von Kendalls/Stuarts τ_c . ☐

7 Lineare Einfachregression

7.1 Aufgaben

1. Welche der folgenden Aussagen sind wahr?

- a) Regressionskoeffizient & Korrelationskoeffizient haben die gleiche Aussagekraft. ☐
- b) Man kann bereits aus dem Korrelationskoeffizienten auf das Vorzeichen des Regressionskoeffizienten schließen. ☒
- c) Höherer Korrelationskoeffizient, bedeutet automatisch auch höherer Regressionskoeffizient. ☐
- d) Sowohl Korrelation als auch Regressionskoeffizient haben einen Wertebereich von -1 bis 1. ☐

2. R^2 bei der linearen Einfachregression ..

- a) .. ist immer kleiner/gleich dem Korrelationskoeffizienten. ☒
- b) .. kann Werte von -1 bis 1 annehmen. ☐
- c) .. beschreibt den Anteil der erklärten Streuung der Zielgröße durch die Einflussgröße. ☒
- d) .. deutet bei kleinen Werten auf eine eher schlechte Modellgüte hin. ☒
- e) .. beschreibt den Anteil der erklärten Streuung des Modells. ☐

3. Was ist bei der Interpretations der geschätzten Koeffizienten im Regressionsmodell wichtig?

- a) Stets nur den Absolutbetrag interpretieren. ☐
- b) Der Intercept ist der erwartete Wert der Zielgröße wenn die Einflussgröße ihren Durchschnittswert annimmt. ☐
- c) Interpretation des Steigungsparameters pro Einheit. ☒
- d) Der Intercept ist nicht immer sinnvoll interpretierbar. ☒
- e) Vorhersagen sollte man nur für Werte der Einflussgröße durchführen, die auch so (ähnlich) in der Stichprobe vorkommen. ☒

4. Welche Aussagen bzgl. kategorialen Regressoren sind korrekt?

- a) Sowohl Dummy- als auch Effekt-Kodierung führen zu gleichen Zahl an Dummy-Variablen. ☒
- b) Die Interpretation der geschätzten Koeffizienten ist bei Dummy- & Effekt-Kodierung identisch. ☐
- c) Der Intercept ist weder bei Dummy- noch bei Effekt-Kodierung interpretierbar. ☐

- d) Bei einer höheren Anzahl an verschiedenen Kategorien ist die Dummy-Kodierung sinnvoller. ☐

5. Interpretation des Intercepts bei Dummy-Kodierung:

- a) Der Intercept entspricht dem erwarteten Wert der Zielgröße bei Vorliegen der Referenzkategorie. ☒
- b) Eine Änderung der Referenzkategorie hat eine Änderung der geschätzten Koeffizienten für alle Dummy-Variablen zur Folge. ☒
- c) Eine Änderung der Referenzkategorie hat (potenziell) eine Änderung der Anpassungsgüte (R^2) zur Folge. ☐
- d) Die Referenzkategorie ist immer auf natürliche Art & Weise vorgegeben. ☐

8 Indizes

8.1 Aufgaben

1. Welche der folgenden Aussagen über Indexzahlen sind wahr?

- a) Zur Umbasierung (Veränderung Basisjahr) werden lediglich die Indexzahlen und nicht die Rohdaten benötigt. ☒
- b) Zur Verkettung von Indexzahlen werden die Rohdaten benötigt ☐
- c) Zur Umbasierung müssen sich alle bereits vorliegenden Indexzahlen auf das gleiche Basisjahr beziehen. ☒
- c) Indexzahlen werden für Mengen und Preise getrennt berechnet. ☒

2. Welche Aussagen bzgl. der verschiedenen Indizes sind wahr?

- a) Der Preisindex nach Laspeyres ist stets größer als der nach Paasche. ☐
- b) Sind die Mengen in Berichts- und Basisperiode gleich, so sind die Preisindizes nach Laspeyres und Paasche ebenfalls identisch. ☒
- c) Bei konstanten Preisen sind sowohl der Mengenindex nach Laspeyres als auch der nach Paasche gleich 1. ☐
- d) Der Mengenindex nach Paasche gewichtet die Mengen mit den Umsatzanteilen aus der Berichtsperiode. ☐

3. Welche Aussagen bzgl. "Spezieller Probleme" sind wahr?

- a) Bei der Substitution muss für mindestens eine Periode der Preis für beide Güter beobachtet werden. ☒
- b) Bei der Erweiterung muss für das neue Produkt auch eine Menge in der Basisperiode bekannt sein. ☐
- c) Bei der Substitution werden die Preissteigerungen des neuen Produkts einfach auf das alte übertragen. ☒
- d) Subindizes können durch Gewichtung mit Mengenanteilen zu einem Gesamtindex kombiniert werden. ☐

9 Zeitreihen

9.1 Aufgaben

1. Welche Aussagen bzgl. gleitender Durchschnitte sind wahr?

- a) Durch k wird festgelegt, ob der die Ordnung der gl. Durchschnitt gerade oder ungerade ist. ☐
- b) Bei einem gl. Durchschnitt ungerader Ordnung fallen am Rand der Zeitreihe mehr Werte weg als bei gerader Ordnung. ☐
- c) Je höher die Ordnung, desto mehr Werte fallen am Rand der Zeitreihe weg. ☒
- d) Bei einem gl. Durchschnitt gerader Ordnung gehen (bei gleichem k) mehr Zeitpunkte in die Berechnung mit ein als bei ungerader Ordnung. ☐

2. Welche Aussagen bzgl. des Zeitreihenmodells sind wahr?

- a) Mit gleitenden Durchschnitten kann die Trendkomponente geschätzt werden. ☒
- b) Mit gleitenden Durchschnitten kann die Saisonkomponente geschätzt werden. ☐
- c) Jede Zeitreihe besitzt eine Trend- und eine Saisonkomponente. ☐
- d) Welche Ordnung für die gleitenden Durchschnitte gewählt wird, hängt von der Saisonkomponente ab. ☒

10 R-Einführung Teil I

11 Kombinatorik

11.1 Aufgaben

1. Welche Aussagen bzgl. Permutationen sind wahr?

- a) Die Reihenfolge der Elemente kann eine Rolle spielen, muss es aber nicht. ☐
- b) Permutation wird eine mögliche Anordnung von Elementen in einer bestimmten Reihenfolge genannt. ☒
- c) Die Anzahl der Permutationen ohne Reihenfolge und ohne Wiederholung berechnet man mit $n!$ ☐
- d) Die Anzahl der möglichen Permutation mit Wiederholung ist größer als ohne Wiederholungen. ☐

2. Welche Aussagen bzgl. Kombinationen sind richtig?

- a) Die Betrachtung *ohne Wiederholung und ohne Reihenfolge* hat eine größere Anzahl an Kombinationsmöglichkeiten als die Betrachtung *mit Wiederholung und ohne Reihenfolge*. ☐
- b) Wenn die Reihenfolge bei der Kombination mit einbezogen werden soll, dann wird die Anzahl an möglichen Kombinationen größer. ☒
- c) Bei Betrachtung von Kombinationen *ohne Wiederholung und mit Reihenfolge* erhält man eine höhere Anzahl an Möglichkeiten als bei der Permutation ohne Wiederholung ☐
- d) Bei der Kombination mit Wiederholung und mit Reihenfolge gibt es auf dem "ersten Platz" n verschiedene Möglichkeiten, auf dem "zweiten Platz" $n - 1$, usw. ☐

3. Ein Zahlenschloss besteht aus 4 Rädern mit den Zahlen von 0 bis 9. Welche Aussage ist richtig?

- a) Um die Anzahl der möglichen Kombinationen zu berechnen benutzt man die Permutation mit Wiederholung. ☐
- b) Um die Anzahl der möglichen Kombinationen zu berechnen benutzt man die Kombination mit Wiederholung und mit Reihenfolge. ☒
- c) Um die Anzahl der möglichen Kombinationen zu berechnen benutzt man die Kombination mit Wiederholung und ohne Reihenfolge. ☐
- d) Um die Anzahl der möglichen Kombinationen zu berechnen benutzt man die Kombination ohne Wiederholung und ohne Reihenfolge. ☐
- e) Es gibt 10.000 verschiedene Kombinationen. ☒

4. Bei einem Basketballspiel laufen nacheinander 5 Spieler auf das Spielfeld. Man berechnet die Anzahl der Möglichkeiten für das Einlaufen mit ...

- a) der Kombination ohne Wiederholung und mit Reihenfolge. ☐
- b) der Permutation ohne Wiederholung. ☒
- c) der Permutation mit Wiederholung. ☐
- d) der Kombination ohne Wiederholung und ohne Reihenfolge. ☐
- e) der Kombination mit Wiederholung und mit Reihenfolge. ☐

5. Bei einem Rosenzüchter gibt es 14 verschiedene Rosenarten und man möchte ein Strauß mit 20 Rosen. Wie viele unterschiedliche mögliche Sträuße gibt es?

- a) $2,432902008 \times 10^{18}$ ☐
- b) $5,73166440 \times 10^8$ ☒
- c) $8,366825543 \times 10^{22}$ ☐
- d) $8,71782912 \times 10^{10}$ ☐

12 Wahrscheinlichkeitsrechnung

12.1 Aufgaben

1. Eine Mutter kauft für den Kindergeburtstag 20 Luftballons, 10 blaue und 10 rote. Welche Aussagen sind richtig?

- a) Ein sicheres Ereignis wäre entweder die Farbe blau oder rot als erstes aufzublasen. ☒
- b) Ein unmögliches Ereignis wäre die Farbe gelb als erstes auszublasen. ☒
- c) Das Komplementärereignis \bar{A} von Ereignis A "*Einen blaue Luftballon als erstes aufblasen*" ist "*Keinen Luftballon aufblasen*". ☐
- d) Bei der Frage, welchen Luftballon die Mutter als erstes aufbläst, gibt es hier zwei Elementarereignisse. ☒
- e) Das Ereignis "*Blauen Luftballon als erstes aufblasen*" ist ein Elementarereignis. ☒

2. Welche Aussagen zur Laplaceschen Wahrscheinlichkeit sind richtig?

- a) Bei der Laplaceschen Wahrscheinlichkeit kann der Ereignisraum unendlich sein, solange die Ereignisse gleich wahrscheinlich sind. ☐
- b) Die einzige Voraussetzung für ein Laplacesche Wahrscheinlichkeit ist, dass die Ereignisse gleich wahrscheinlich sind. ☐
- c) Mit der Anzahl der für A günstigen Fälle und der Anzahl aller möglichen Ereignisse berechnet man die Laplacesche Wahrscheinlichkeit. ☒
- d) Bei der Laplaceschen Wahrscheinlichkeit muss ein Zufallsexperiment zugrunde liegen. ☒

3. Welche Aussagen zur Wahrscheinlichkeitsrechnung sind wahr?

- a) Die Wahrscheinlichkeit eines unmöglichen Ereignisses ist die leere Menge. ☐
- b) Die Wahrscheinlichkeit der Schnittmenge zweier disjunkter Ereignisse ist die Summe der Einzelwahrscheinlichkeiten abzüglich der Schnittmenge der beiden Ereignisse. ☐
- c) Wenn man die Wahrscheinlichkeit des Ereignisses A kennt, kann man auch die Wahrscheinlichkeit des Komplementärereignisses \bar{A} berechnen. ☒
- d) Die Wahrscheinlichkeit eines Ereignisses A kann Zahlen zwischen -1 und 1 annehmen. ☐
- e) Ist B eine Teilmenge von A , so ist die Wahrscheinlichkeit von B kleiner oder gleich der Wahrscheinlichkeit von A . ☒

4. Welche Aussagen zur bedingten Wahrscheinlichkeit sind wahr?

- a) Den Satz der totalen Wahrscheinlichkeit verwendet man, wenn man die unbedingte Wahrscheinlichkeit eines Ereignisses berechnen möchte. ☒
- b) Man kann den Satz von Bayes verwenden, um eine bedingte Wahrscheinlichkeit zu berechnen. ☒
- c) Die bedingte Wahrscheinlichkeit $P(E|G)$ bedeutet die Wahrscheinlichkeit des Ereignisses E , unter der Bedingung, dass das Ereignis G nicht eingetreten ist. ☐
- d) Wenn die bedingte Wahrscheinlichkeit $P(A|B)$ ungleich der bedingten Wahrscheinlichkeit $P(B|A)$, dann sind A & B disjunkt. ☐

5. Welche Aussagen zur stochastischen Unabhängigkeit sind wahr?

- a) Wenn die bedingte Wahrscheinlichkeit $P(A|B)$ ungleich der bedingten Wahrscheinlichkeit $P(B|A)$, dann sind A & B unabhängig. ☐
- b) Ist die bedingte Wahrscheinlichkeit $P(B|G)$ ungleich der Wahrscheinlichkeit $P(B|\bar{G})$, dann liegt keine stochastische Unabhängigkeit zwischen B und G vor. ☒
- c) Ist die bedingte Wahrscheinlichkeit $P(B|G)$ ungleich der Wahrscheinlichkeit $P(B)$, dann liegt keine stochastische Unabhängigkeit zwischen B und G vor. ☐
- d) Liegt stochastische Unabhängigkeit vor, so berechnet man die gemeinsame Wahrscheinlichkeit als die Summe der beiden Einzelwahrscheinlichkeiten. ☐

13 Zufallsvariablen

13.1 Aufgaben

Welche Aussagen zu den Zufallsvariablen sind richtig?

- a) Der fünffache Münzwurf ist eine diskrete Zufallsvariable. ☒
- b) Mit Zufallsvariablen können Ergebnisse von Zufallsexperimenten beschrieben werden, die noch nicht durchgeführt wurden. ☒
- c) Der Träger der Zufallsvariable Würfelwurf eines 24-seitigen Würfels ist $T = \{1, 2, 3, \dots, 24\}$ ☒
- d) Eine stetige Zufallsvariable hat abzählbar endlich viele mögliche Ergebnisse. ☐

14 Spezielle Verteilungen

14.0.1 Aufgaben I

1. Welche Aussagen bzgl. der Gleichverteilung sind richtig?

- a) Bei der diskreten Gleichverteilung muss es immer zwei mögliche Ausgänge geben, die die gleiche Wahrscheinlichkeit haben. ☐
- b) Die Augenzahl bei dem Wurf eines Würfels mit 88 Seiten folgt einer Gleichverteilung. ☒
- c) Das Ergebnis des Wurfs einer unfairen Münze mit Kopf und Zahl ist bernoulliverteilt. ☒
- d) Das Ergebnis des Wurfs einer unfairen Münze mit Kopf und Zahl ist gleichverteilt. ☐

2. Welche Aussagen bzgl. der Bernoulliverteilung und Binomialverteilung sind wahr?

- a) Die Binomialverteilung besteht aus n abhängigen und identische Bernoulliexperimenten. ☐
- b) Der Erwartungswert der Bernoulliverteilung ist die Wahrscheinlichkeit, dass das Ereignis eintritt, selbst. ☒
- c) Bei der Bernoulliverteilung und bei der Binomialverteilung gibt es jeweils nur zwei mögliche Ergebnisse. ☐
- d) Die Verteilungsfunktion der Binomialverteilung steigt von 0 bis 1. ☒

3. Gegeben ist die Rechnung $P(X = 4) = 0,3 \cdot 0,7^3$ Welche Aussagen sind wahr?

- a) Hierbei handelt es sich um die geometrische Verteilung. ☒
- b) Der Erwartungswert der Verteilung beträgt hier ca. 3,33 ☒
- c) Mit dieser Gleichung berechnet man die Wahrscheinlichkeit, dass das gewünschte Ereignis bei der dritten Wiederholung eintritt. ☒
- d) Die Wahrscheinlichkeit für das Ereignis von Interesse liegt bei 0,7. ☐

4. Welche Aussagen bezüglich der geometrischen und hypergeometrischen Verteilung sind wahr?

- a) Je größer die Wahrscheinlichkeit eines Ereignisses ist, desto länger die erwartete Dauer, bis es zum ersten Mal auftritt. ☐
- b) Die hypergeom. Verteilung verwendet man, wenn man wissen will, mit welcher Wahrscheinlichkeit ein Ereignis nach wie vielen Wiederholungen eines Zufallsexperiments eintritt. ☐
- c) Mit der hypergeometrische Verteilung berechnet man die Wahrscheinlichkeit unter der Voraussetzung, dass die gezogenen Elemente wieder zurückgelegt werden. ☐

5. Welche Aussagen zur Poissonverteilung sind richtig?

- a) Sie eignet sich zur Berechnung der Wahrscheinlichkeit für das Eintreten (ja/nein) eines Ereignisses. ☐
- b) Das λ muss größer gleich 0 sein. ☒
- c) Die Varianz und der Erwartungswert sind immer gleich groß. ☒
- d) Das λ gibt die Intensitätsrate an, das heißt je größer λ ist, desto häufiger wird das Ereignis wahrscheinlich auftreten. ☒

6. Welche Aussagen zur Multinomialverteilung sind richtig?

- a) Bei der Multinomialverteilung ist es wichtig zu wissen, wie groß n ist. ☒
- b) Die Multinomialverteilung ist die Verallgemeinerung der Bernoulliverteilung mit mehr als zwei möglichen Ereignissen. ☐
- c) k gibt Anzahl der verschiedenen möglichen Ergebnisse an. ☒
- d) Die k Wahrscheinlichkeiten müssen sich nicht zwingend zu 1 aufaddieren. ☐

14.0.2 Aufgaben II

1. Welche Aussagen zur stetigen Gleichverteilung sind richtig?

- a) Jede mögliche Ausprägung kommt mit einer Wahrscheinlichkeit von 0 vor. ☒
- b) Der Unterschied zwischen der stetigen und diskreten Gleichverteilung ist, dass die stetige Gleichverteilung unendlich viele verschiedene Werte annehmen kann. ☒
- c) Die Dichtefunktion ist dasselbe wie die Verteilungsfunktion. ☐
- d) Die Dichtefunktion bei der stetigen Gleichverteilung ist eine steigende Gerade. ☐
- e) Die Dichtefunktion von $X \sim U(2, 6)$ verläuft auf der Höhe 0,25. ☒

2. Welche Aussagen zur Exponentialverteilung sind richtig?

- a) Die zukünftige Wartezeit ist abhängig von der davor schon verstrichenen Zeit. ☐
- b) Bei der Exponential-, wie auch der geometrischen Verteilung geht es darum, wann das gewünschte Ereignis zum ersten/nächsten Mal auftritt. ☒
- c) Die Exponentialverteilung ist nur für $X \geq 0$ definiert. ☒
- d) Ein zentraler Begriff bei der Exponentialverteilung ist die Gedächtnislosigkeit. ☒

3. Welche Aussagen zur Normalverteilung / Chi-Quadrat-Verteilung sind wahr?

- a) Die Dichtefunktion ist bei beiden symmetrisch um den Erwartungswert. ☐
- b) Der Durchschnitt von i.i.d. normalverteilten Zufallsvariablen ist normalverteilt. ☒
- c) Um die Dichtefunktion der Normalverteilung aufzustellen benötigt man nur σ und μ . ☒
- d) Die Verteilungsfunktion der Normalverteilung hat bei μ ihr Maximum. ☐
- e) Für die Chi-Quadrat-Verteilung werden die Quadrate von mehreren standardnormalverteilten Zufallsvariablen gebildet und aufsummiert. ☒

4. Welche Aussagen über die t-Verteilung sind richtig?

- a) Die t-Verteilung kann immer durch die Standardnormalverteilung approximiert werden. ☐
- b) Je größer die Anzahl der Freiheitsgrade, desto näher kommt sie der Standardnormalverteilung. ☒
- c) Die Wahrscheinlichkeitsmasse an den Rändern ist bei der t-Verteilung geringer als bei der Standardnormalverteilung. ☐
- d) Je weniger Freiheitsgrade, desto ähnlicher ist die t-Verteilung der Standardnormalverteilung. ☐

14.0.3 Aufgaben III

1. Welche Aussagen sind richtig?

- a) Eine Verteilungsfunktion geht immer von -1 bis 1. ☐
- b) Der Wert einer Verteilungsfunktion an der Stelle $x = 5$ ist $P(X \leq 5)$. ☒
- c) Die Werte der Verteilungsfunktion entsprechen kumulierten Wahrscheinlichkeiten. ☒
- d) Die Wahrscheinlichkeitsfunktion ist bei diskreten Zufallsvariablen optisch dasselbe wie die Verteilungsfunktion bei stetigen Zufallsvariablen. ☐

2. Welche Aussagen sind richtig?

- a) Die Verteilungsfunktion ist monoton steigend. ☒
- b) Die Verteilungsfunktionen von stetigen Zufallsvariablen sind Treppenfunktionen im Gegensatz zu den diskreten Zufallsvariablen. ☐
- c) Eine Eigenschaft der Verteilungsfunktion ist, dass die Fläche unterhalb der Funktion immer 1 ergibt. ☐
- d) Die Summe aller Einzelwahrscheinlichkeiten bei einer stetigen Zufallsvariable ergibt 1. ☐
- e) Eine Eigenschaft der Dichtefunktion einer stetigen ZV ist, dass die Funktionswerte immer größer gleich 0 sind. ☒

3. Welche Aussagen sind richtig?

- a) Um die Verteilungsfunktion zu erhalten, leitet man die Dichtefunktion ab. ☐
- b) Die Fläche unter der Dichtefunktion ist die Quantilsfunktion. ☐
- c) Die Quantilsfunktion ist die Umkehrfunktion der Verteilungsfunktion. ☒
- d) Die Ableitung der Verteilungsfunktion ist die Dichte. ☒

15 Grenzwertsätze und Approximationen von Verteilungen

15.1 Aufgaben

1. Welche Aussagen bzgl. den Grenzwertsätzen ist wahr?

- a) Je größer der Stichprobenumfang, desto größer wird der Einfluss von Ausreißern. ☐
- b) Mit zunehmendem Stichprobenumfang nähert sich die beobachtete relative Häufigkeit der theoretischen Wahrscheinlichkeit an. ☒
- c) Egal welche Verteilung vorliegt, nähert sich der Mittelwert von i.i.d. Zufallsvariablen bei zunehmenden Stichprobenumfang der Normalverteilung an. ☒

2. Welche Aussagen bzgl. Approximationen sind richtig?

- a) Die hypergeom. Verteilung kann direkt durch die Normalverteilung approximiert werden. ☐
- b) Bei der Approximation der Poissonverteilung durch die Normalverteilung wird das λ für das μ und das σ einfach übernommen. ☒
- c) Die Poissonverteilung darf immer durch die Normalverteilung approximiert werden. ☐
- d) Die hypergeometrische Verteilung $H(9,120,40)$ darf durch die Binomialverteilung approximiert werden. ☐

16 Schätzen

16.1 Aufgaben

Welche Aussagen ist bzgl. Punkt- & Intervallschätzer sind richtig?

- a) Punktschätzer sind immer richtig und wahrheitsgetreu. ☐
- b) Ein Konfidenzintervall gibt den Bereich vor, in dem der wahre Parameter immer zu finden ist. ☐
- c) Man kann die Länge eines Intervalls variieren, je nach dem mit welcher Wahrscheinlichkeit der wahre Parameter durch das Intervall überdeckt werden soll. ☒
- d) Die Länge eines Intervalls hängt von α , σ^2 , der Verteilung und dem Stichprobenumfang ab. ☒
- e) Ein 99%-KI besagt, dass zu 99% mein Schätzer richtig ist. ☐

17 Testtheorie

17.1 Aufgaben

1. Welche Aussagen zum p-Wert sind richtig?

- a) Der p-Wert gibt die Wahrscheinlichkeit an, mit der der beobachtete Wert richtig ist. ☐
- b) Der p-Wert gibt $P(X)$ an. ☐
- c) Die Hypothese, die man nachweisen will befindet sich in der Alternativhypothese H_1 . ☒
- d) Wenn der p-Wert kleiner als α ist, dann lehnt man H_0 ab. ☒
- e) Bei einem beidseitigen Test gibt es immer zwei Ablehnbereiche. ☒
- f) Der p-Wert gibt uns darüber Auskunft, mit welcher Wahrscheinlichkeit unser Wert oder ein Wert, der noch weiter von der Nullhypothese entfernt ist, auftritt, unter der Annahme, dass H_0 stimmt. ☒

2. Welche Aussagen über die Hypothesentests sind wahr?

- a) Mit Hypothesentest kann man entscheiden, ob eine Nullhypothese richtig ist. ☐
- b) Bei einem einfachen t-Test gibt es immer nur einen Ablehnbereich. ☐
- c) Der Unterschied zwischen dem einfachen Gauss-Test und einfachen t-Test ist die unbekannte (wahre) Varianz beim t-Test. ☒
- d) Die Teststatistik beim Gauss-Test ist im Gegensatz zum t-Test normalverteilt. ☒

3. Welche Aussagen über die Hypothesentests sind wahr?

- a) Mit dem F-Test kann man das Verhältnis von Varianzen zweier unabhängigen Zufallsvariablen testen. ☒
- b) Beim doppelten Gauss-Test handelt es sich im Gegensatz zum einfachen Gauss-Test um das Verhältnis der Varianzen zweier unabhängiger, normalverteilter ZV. ☒
- c) Beim doppelten Gauss-Test sind die Varianzen unbekannt. ☐
- d) Der Unterschied zwischen dem doppelten t-Test und dem Welch-Test ist, dass beim doppelten t-Test die Varianzen bekannt sind und beim Welch-Test diese unbekannt sind. ☐
- e) Die Besonderheit beim Paired t-Test ist, dass es sich hierbei um einen Test mit zwei abhängigen normalverteilten ZV handelt. ☒
- f) Der F-Test ist immer zweiseitig. ☐

4. Welche Aussagen über die Hypothesentests sind wahr?

- a) Mit dem Chi-Quadrat-Anpassungstest kann man testen, ob die Verteilung zweier ZV unterschiedlich ist. ☐
- b) Der Vorteil beim Mann-Whitney-U-Test ist, dass die ZV nicht mehr unbedingt normalverteilt sein müssen. ☒
- c) Der Kolmogorov-Smirnov-Anpassungstest prüft, ob sich zwei ZV in ihrer Verteilung unterscheiden. ☒
- d) Beim Chi-Quadrat-Unabhängigkeitstest wird die Unabhängigkeit zweier Merkmale überprüft, indem die beobachteten Häufigkeiten mit den zu erwarteten Häufigkeiten (ohne Abhängigkeit) verglichen werden. ☒

5. Welche Aussagen über die Hypothesentests sind wahr?

- a) Der Ablehnbereich ist beim Gauss-Test (bei identischem α) größer als beim t-Test. ☒
- b) Beim Gauss-Test liegt etwas mehr Unsicherheit zugrunde als beim t-Test. ☐
- c) Bei größeren Stichproben nähern sich die Ergebnisse des t-Tests und des Gauss-Tests an. ☒
- d) Beim Welch-Test sind wir mit unserer Entscheidung ein wenig sicherer als beim doppelten t-Test, da wir nur eine unbekannte Varianz schätzen müssen. ☐

18 Lineare Regression II

18.1 Aufgaben

1. Welche Aussagen über die einfache lineare Regression sind wahr?

- a) Der Erwartungswert der Residuen sollte per Annahme gleich 1 sein. ☐
- b) Homoskedastizität bedeutet konstante Varianz der Fehlertermen. ☒
- c) In Statistik II trifft man die Annahme, dass die Fehlerterme Zufallsvariablen sind. ☒
- d) In Statistik II kann man keine Rückschlüsse auf die Grundgesamtheit ziehen, da schlussendlich alles nur Annahmen sind und nichts sicher ist. ☐

2. Welche Aussagen über die einfache lineare Regression sind wahr?

- a) Mit steigenden x-Werten, sollte auch die Varianz der Residuen steigen. ☐
- b) Die angenommene Verteilung der Fehlerterme lautet $N(0, \sigma^2)$. ☒
- c) Bei der Effektkodierung wählt man im Gegensatz zur Dummykodierung keine Referenzkategorie. ☐
- d) Bei der Effektkodierung nehmen bei Vorliegen der Referenzkategorie alle Dummys den Wert -1 an. ☒

3. Welche Aussagen bezüglich der multiplen linearen Regression sind richtig?

- a) Eine Eigenschaft des KQ-Schätzers ist, dass er erwartungstreu ist. ☒
- b) Bei der Effektkodierung wird eine andere Anzahl an Dummy-Variablen benötigt als bei der Dummy-Kodierung. ☐
- c) Im Falle von $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4}$ hat die t -Verteilung für den Signifikanztest 5 Freiheitsgrade. ☐
- d) $y = \beta_0 + \beta_1 \sqrt{x} + \beta_2 x^2 + \beta_3 x^3$ ist ein lineares Modell. ☒

4. Welche Aussagen bezüglich des Signifikanztest und der KI-Intervalle sind richtig?

- a) Mit dem Signifikanztest für einen Parameter kann man überprüfen, ob die Nullhypothese signifikant richtig ist. ☐
- b) Da in der Nullhypothese angenommen wird, dass β_j gleich 0 ist, muss man diesen hypothetischen Wert bei Berechnung der Teststatistik praktisch nicht weiter beachten. ☒
- c) Das Konfidenzintervall liegt symmetrisch um den Schätzer. ☒

- d) Wenn $t = -4,8$ ist und $t_{n-p-1, 1-\alpha/2} = 4,5$, dann kann die Nullhypothese $H_0 : \beta_j = 0$ nicht abgelehnt werden. ☐
- e) Eine Teststatistik die t_{194} -verteilt ist, stammt aus einem Modell mit $n = 200$ und 6 geschätzten Koeffizienten. ☒

5. Welche Aussagen über das Bestimmtheitsmaß und den Overall F-Test sind richtig?

- a) Wenn die Nullhypothese abgelehnt wird, dann haben alle Regressoren einen statistisch signifikanten Erklärungswert für die Zielvariable Y . ☐
- b) Bei $R^2 = 1$ spricht man von einer Nullanpassung. ☐
- c) Bei $R^2 = 0$ wird die Streuung der Zielgröße vom Regressionsmodell komplett erklärt. ☐
- d) Ist $R^2 = 0$, dann wird H_0 nicht abgelehnt. ☒
- e) Ist mein berechneter Wert f negativ, dann wird H_0 abgelehnt. ☐

19 R-Einführung Teil II