

Aufgabenteil zur Klausur zum
Modul 2.1 im BA-Studiengang
„Politik- und Verwaltungswissenschaft“
und zum
Modul 3 im BA-Studiengang „Soziologie“

Termin: 9. März 2009, 14.00 - 18.00 Uhr

Prüfer: apl. Prof. Dr. H.-J. Mittag

Der Aufgabenteil wird *nicht* eingesammelt. Es wird aber empfohlen, dass Sie Ihre Antworten eintragen und dieses Dokument zu Hause aufbewahren. So können Sie später mit der Musterlösung vergleichen.

Hinweise zur Bearbeitung der Klausur:

1. Bitte lesen Sie diese Hinweise vollständig und aufmerksam durch, bevor Sie mit der Bearbeitung beginnen.
2. Bitte legen Sie für die Identitätskontrolle Ihren Personalausweis und die Anmeldebestätigung neben die Klausurunterlagen. Im Klausorraum sind das Rauchen und die Benutzung von Mobiltelefonen nicht gestattet.
3. Die Klausur besteht aus 21 Aufgaben, die in 240 Minuten zu bearbeiten sind. Bitte kontrollieren Sie sofort, ob Sie
 - ein vollständiges Klausurexemplar mit 21 Aufgaben,
 - eine Formelsammlung (mit Glossar) zum Kurs 33209 und
 - einen LOTSE-Markierungsbogenerhalten haben.
4. Bevor Sie mit der Bearbeitung der Klausuraufgaben beginnen, füllen Sie bitte den Identifikationsteil des Markierungsbogens aus. Tragen Sie dort Ihren Namen, Ihre Matrikelnummer, Ihre Anschrift sowie das Datum ein und unterschreiben Sie.
5. Für die Bewertung der Klausur sind ausschließlich Ihre Markierungen auf dem LOTSE-Markierungsbogen ausschlaggebend. Erfahrungen haben gezeigt, daß Sie spätestens 20 Minuten vor Abgabe der Klausur mit dem Markieren beginnen sollten. Kontrollieren Sie ganz am Schluss noch einmal Ihre Markierungen, bevor Sie den Markierungsbogen abgeben. Es wird dringend empfohlen, die Markierungen auf dem Lotsebeleg mit einem Bleistift und nicht mit einem Kugelschreiber durchzuführen, weil Sie dann während der Klausur mit einem weichen Radiergummi auch noch Antworten ändern können.
6. Bei jeder Aufgabe ist die maximal erreichbare Anzahl der Punkte angegeben. Insgesamt können Sie 100 Punkte erreichen. Mit 50 Punkten haben Sie die Klausur auf jeden Fall bestanden.
7. Die Klausur besteht aus 16 Multiple-Choice-Aufgaben (je 5 Punkte) und 5 numerischen Aufgaben (je 4 Punkte). Sind die *numerischen Aufgaben* richtig beantwortet, erhalten Sie die volle Punktzahl, ansonsten werden 0 Punkte vergeben. Bei den Multiple-Choice-Aufgaben sind 5 Aussagen vorgegeben, von denen x zutreffend sind mit $x \geq 1$. Zutreffende Aussagen sind zu markieren, falsche Aussagen sind nicht zu markieren. Die Bewertung der *Multiple-Choice-Aufgaben* erfolgt nach folgendem Prinzip:
 - Für jede korrekte Markierung / Nicht-Markierung werden 20 % der pro Aufgabe erzielbaren Punkte vergeben, für jede inkorrekte Markierung / Nicht-Markierung werden 20 % der bei der Aufgabe maximal erreichbaren Punktzahl abgezogen. Es werden aber pro Aufgabe mindestens 0 Punkte vergeben (Ausschluss negativer Salden). Aufgaben, bei denen Sie keine Markierung vornehmen, gelten als nicht bearbeitet und werden mit 0 Punkten bewertet.

Nachstehend ist angedeutet, wie die Lösungen für die ersten 10 Multiple-Choice-Aufgaben aussehen könnten (die Eintragungen sind allerdings hier willkürlich vorgenommen worden und dienen nur Illustrationszwecken):

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|
| <input checked="" type="checkbox"/> A | A | <input checked="" type="checkbox"/> A | <input checked="" type="checkbox"/> A | <input checked="" type="checkbox"/> A | A | A | A | <input checked="" type="checkbox"/> A | A |
| <input checked="" type="checkbox"/> B | B | B | <input checked="" type="checkbox"/> B | B | <input checked="" type="checkbox"/> B | B | B | <input checked="" type="checkbox"/> B | B |
| C | <input checked="" type="checkbox"/> C | <input checked="" type="checkbox"/> C | <input checked="" type="checkbox"/> C | C | C | <input checked="" type="checkbox"/> C | C | <input checked="" type="checkbox"/> C | <input checked="" type="checkbox"/> C |
| D | D | <input checked="" type="checkbox"/> D | <input checked="" type="checkbox"/> D | D | <input checked="" type="checkbox"/> D | D | <input checked="" type="checkbox"/> D | <input checked="" type="checkbox"/> D | D |
| E | E | E | <input checked="" type="checkbox"/> E | <input checked="" type="checkbox"/> E | E | E | <input checked="" type="checkbox"/> E | E | <input checked="" type="checkbox"/> E |

- Bewertungsbeispiel für eine Multiple-Choice-Aufgabe mit 5 zu vergebenden Punkten: Sind A und C richtig und es wurden A, C und D markiert, gibt es 3 Punkte (Für A, B, C und E gibt es jeweils 1 Punkt, also insgesamt 4 Punkte, weil die Markierung von A und C korrekt und die Nicht-Markierung von B und E ebenfalls korrekt ist. Die Antwort D ist inkorrekt markiert worden, so dass $4 - 1 = 3$ Punkte übrig bleiben).
8. Für Zwischenrechnungen können Sie die Rückseiten der Aufgabenblätter verwenden.
 9. Als Hilfsmittel ist eine *Formelsammlung (mit Glossar)* zum Kurs „Statistik“ (Kurs 33209) zugelassen, die mit der Klausur verteilt wird und ebenfalls nicht eingesammelt wird. Auch die Benutzung eines *Taschenrechners* ist gestattet, sofern dieser
 - nicht programmierbar ist und keine Texte oder Formeln speichern kann,
 - über keine alphanumerische Tastatur verfügt und nicht drahtlos mit anderen Geräten kommunizieren kann.
 10. Abzugeben ist nur der Lotse-Markierungsbogen. Der Aufgabenteil und die Formelsammlung werden nicht eingesammelt. Da der Aufgabenteil nicht eingesammelt wird, können Zwischenrechnungen grundsätzlich nicht in die Bewertung eingehen.
 11. Täuschungen und Täuschungsversuche können zum Ausschluss von der Klausur und zur Bewertung mit „nicht ausreichend (5,0)“ führen.
 12. Alle Klausurteilnehmer erhalten von der FernUniversität eine Benachrichtigung, auf dem die erreichte Punktzahl und die Note vermerkt sind. Die maschinelle Auswertung nimmt erfahrungsgemäß einen Zeitraum von bis zu 8 Wochen in Anspruch. Sehen Sie daher in diesem Zeitraum bitte von Nachfragen zum Klausurergebnis ab.

Viel Erfolg bei der Klausurbearbeitung!

Aufgabe 1 (Aussagenlogik)

(5 Punkte)

In der nachstehenden Wahrheitstabelle stehen im Tabellenkopf der ersten beiden Spalten zwei Aussagen a und b , die wahr (w) oder falsch (f) sind. In den beiden ersten Spalten sind alle möglichen Fälle bezüglich des Wahrheitsgehalts der beiden Aussagen angegeben (beide Aussagen „w“, nur eine Aussage „w“, beide Aussagen „f“).

| a | b | P1: $a \vee \neg b$ | P2: $\neg a$ | K: $\neg a \wedge \neg b$ |
|-----|-----|---------------------|--------------|---------------------------|
| w | w | | | |
| w | f | | | |
| f | w | | | |
| f | f | | | |

Bei den letzten drei Spalten ist nur der Tabellenkopf ausgefüllt. Hier sind zwei Prämissen $P1$ und $P2$ angegeben, die sich aus a und b ableiten. Im Kopf der letzten Spalte steht ein aus den Prämissen abgeleiteter logischer Schluss K , der allerdings noch auf seinen Wahrheitsgehalt zu überprüfen ist. Das Zeichen \neg bedeutet die Negation einer Aussage, \wedge (Konjunktion) beinhaltet ein logisches „und“ (zwei Aussagen gelten gleichzeitig), während das Zeichen \vee (Disjunktion) ein logisches „oder“ darstellt (von zwei Aussagen gilt mindestens eine).

Welche der folgenden Aussagen sind richtig ? (x aus 5)

- A) Wenn die Aussagen a und b beide wahr sind, sind auch die Prämissen $P1$ und $P2$ erfüllt, also beide wahr.
- B) Wenn die Aussage a wahr und b falsch ist, ist $P1$ erfüllt, nicht aber $P2$.
- C) Wenn die Aussage a falsch und b wahr ist, ist $P2$ erfüllt, nicht aber $P1$.
- D) Wenn die Aussagen a und b beide falsch sind, sind sowohl $P1$ als auch $P2$ erfüllt.
- E) Wenn $P1$ und $P2$ wahr sind, ist auch K wahr, d. h. die Konklusion ist korrekt.

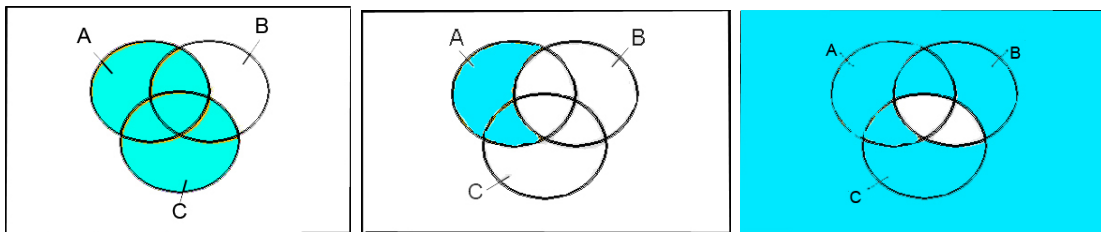
Hinweis: Ergänzen Sie am besten zunächst die 12 fehlenden Werte der obigen Tabelle, d. h., setzen sie jeweils w oder f ein. Die ausgefüllte Tabelle geht zwar nicht in die maschinelle Bewertung ein, erleichtert es Ihnen aber, die zutreffenden der fünf vorgegebenen Antwortalternativen zu finden.

Aufgabe 2 (Operationen mit Mengen)

(5 Punkte)

Zur Veranschaulichung von Mengenoperationen lassen sich Venn-Diagramme heranziehen. Diese bestehen aus einem Rechteck, in dem Mengen A , B , C , ... als Kreise oder Ellipsen dargestellt sind. Das Rechteck repräsentiert eine Grundgesamtheit, von der die eingezeichneten Mengen Teilmengen sind.

Nachstehend sind drei Venn-Diagramme abgebildet, die sich auf die Verknüpfung von drei Mengen A , B und C beziehen. Die Verknüpfungen erfolgen über die Symbole \cup (Vereinigung von Mengen) und \cap (Schnittmengenbildung) oder \setminus (Differenzmengenbildung). Die Komplementärmenge von A , B und C werden mit \bar{A} , \bar{B} und \bar{C} bezeichnet.



Welche der folgenden Aussagen, die sich alle auf die dunkel markierten Flächen beziehen, sind richtig? (x aus 5)

- A) Im ersten Venn-Diagramm ist die Vereinigungsmenge der beiden Mengen A und C dargestellt, also $A \cup C$.
- B) Im zweiten Venn-Diagramm ist die Schnittmenge aus A und der Komplementärmenge von B dargestellt, also $A \cap \bar{B}$.
- C) Im zweiten Venn-Diagramm ist die Differenzmenge von A und B dargestellt, also $A \setminus B$.
- D) Im dritten Venn-Diagramm ist die Vereinigungsmenge von \bar{B} und \bar{C} dargestellt, also $\bar{B} \cup \bar{C}$.
- E) Im dritten Venn-Diagramm ist die Komplementärmenge von $B \cap C$ dargestellt.

Aufgabe 3 (Nominaldefinition)

(5 Punkte)

Die folgenden Aussagen beziehen sich auf die Nominaldefinition, also auf die in den Erfahrungswissenschaften übliche Form des Definierens.

Welche Aussagen sind zutreffend? (x aus 5)

- A) Eine Nominaldefinition legt fest, dass ein bestimmter Begriff (das Definiendum) mit einem anderen Begriff (dem Definiens) gleichbedeutend ist.
- B) Eine Nominaldefinition kann niemals falsch oder richtig sein.
- C) Eine Nominaldefinition hat keinen empirischen Informationsgehalt.
- D) Eine Nominaldefinition berücksichtigt alle Merkmale des zu definierenden Objekts oder Sachverhalts.
- E) Eine Nominaldefinition kann dem Zweck dienen, Begriffe auf die Bedürfnisse eines Forschungsprojekts zuzuschneiden und eine intersubjektive Kontrolle des Forschungsprozesses zu ermöglichen.

Aufgabe 4 (Messen / Verfahren der Datenerhebung)

(5 Punkte)

Welche der folgenden Aussagen sind richtig ? (x aus 5)

- A) Die Validität eines Messinstruments charakterisiert, inwieweit ein Messinstrument bei wiederholter Messung die gleichen Messwerte liefert.
- B) Wenn man allgemeine Bevölkerungsumfragen als offene Online-Befragungen organisiert, ist mit erheblichen Verzerrungen zu rechnen.
- C) Die Randomized Response Technik ist ein Verfahren der Datenerhebung, bei dem in Interviews vollständige Anonymität der Befragten gewährleistet werden kann.
- D) Die „Total-Design-Methode“ ist ein Ansatz der Datenerhebung, der vor allem darauf abzielt, die Rücklaufquote bei schriftlichen Befragungen zu erhöhen.
- E) Das Quotenauswahlverfahren ist ein Verfahren der Zufallsauswahl, das Kostenvorteile gegenüber anderen Verfahren der Datenerhebung bietet.

Aufgabe 5 (Befragungen)

(5 Punkte)

Welche der folgenden Aussagen sind richtig ? (x aus 5)

- A) Bei einem Telefoninterview ist der Einfluss von Einflüssen, die in der Person des Interviewers liegen, geringer als bei einem persönlichen Interview (Face-to-Face-Interview).
- B) Bei Befragungen wird manchmal die Methode der Klumpenauswahl herangezogen. Diese stellt ein zweistufiges Auswahlverfahren dar, bei dem auf der ersten Stufe eine Zufallsauswahl von Teilmengen einer Grundgesamtheit erfolgt und auf der zweiten Stufe eine Untersuchung aller Elemente der zufällig ausgewählten Teilmengen.
- C) Das narrative Interview ist ein Beispiel für eine strukturierte Befragung.
- D) Bei Interviews sind Antwortverzerrungen möglich, die z. B. in der Art der Frageformulierung begründet sein können.
- E) Non-Response, also Antwortausfälle durch Antwortverweigerung oder Nichterreichbarkeit bei Befragungen, kann die Repräsentativität einer Stichprobe stark beeinträchtigen, also zu verzerrten Stichproben führen.

Aufgabe 6 (Merkmalsklassifikationen)

(5 Punkte)

Ein Konzern will im Rahmen von Neubauplanungen auch Informationen über den Bedarf an Parkplätzen und Kinderbetreuungsplätzen für seine Mitarbeiter gewinnen. Hierfür werden alle Beschäftigten gebeten, Angaben zu folgenden Merkmalen zu machen:

- 1 Anzahl der im Haushalt lebenden Kinder
- 2 Alter der im Haushalt lebenden Kinder
- 3 Verkehrsmittel, das für die Fahrt zur Arbeitsstätte überwiegend genutzt wird (z. B. eigener PKW, Fahrrad, Bus, ...)
- 4 Entfernung zwischen Wohnung und Arbeitsstätte
- 5 Einschätzung des vom Arbeitgeber aktuell vorgehaltenen Kinderbetreuungsangebots (1 = sehr gut, ... , 6 = völlig unzureichend).

Welche der folgenden Aussagen sind richtig ? (x aus 5)

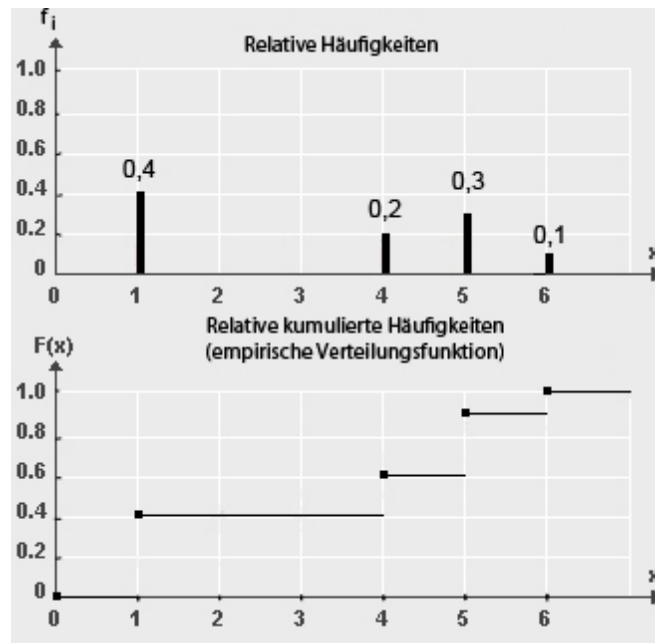
- A) Die Merkmale 1, 2 und 4 sind metrisch skaliert.
- B) Das Merkmal 3 ist nominalskaliert, Merkmal 5 hingegen ordinalskaliert.
- C) Metrisch skalierte Merkmale lassen sich – unter Informationsverlust – auch auf einer Ordinalskala messen.
- D) Bei ordinalskalierten Merkmalen ist die Differenzenbildung zulässig, nicht aber bei nominalskalierten Merkmalen.
- E) Merkmal 1 ist ein Beispiel für ein diskretes Merkmal, Merkmal 4 für ein stetiges Merkmal.

Hinweis: Der Begriff „metrische Skala“ wird als Oberbegriff für „Intervallskala“, „Verhältnisskala“ und „Absolutskala“ verwendet.

Aufgabe 7 (Häufigkeitsverteilungen)

(5 Punkte)

Die nachstehende Abbildung zeigt im oberen Teil die relative Häufigkeitsverteilung anhand eines Stabdiagramms und im unteren Teil die relative kumulierte Häufigkeitsverteilung für einen 20 Werte umfassenden Datensatz, der durch ein Würfelexperiment zustande kam (20-maliges Würfeln mit einem Würfel). Die relativen Häufigkeiten sind im Stabdiagramm auch numerisch ausgewiesen.



Welche der folgenden Aussagen sind richtig ? (x aus 5)

- A) Bei den 20 Würfeln trat 6-mal die Augenzahl 5 auf.
- B) Von den 20 Würfeln führten 12 zu Augenzahlen, die unter 5 lagen.
- C) Von den 20 Würfeln führten 14 zu Augenzahlen, die unter 5 lagen.
- D) Durch Angabe der relativen Häufigkeiten oder der kumulierten relativen Häufigkeiten ist die empirische Verteilung des Merkmals „Augenzahl X “ eindeutig bestimmt.
- E) Die relativen Häufigkeiten gehen aus den absoluten Häufigkeiten hervor, indem man letztere durch den Umfang n des Datensatzes dividiert.

Aufgabe 8 (Kenngrößen empirischer Verteilungen)

(5 Punkte)

Bei 20 Arbeitnehmern eines mittelständischen Betriebes mit 150 Beschäftigten wurde die Zeit gemessen, die sie täglich am Bildschirm verbringen (Angaben in vollen Stunden). Es ergab sich der nachstehende Datensatz:

$$3, 4, 1, 2, 2, 6, 5, 4, 3, 5, 2, 2, 5, 2, 5, 4, 1, 2, 5, 5.$$

Welche der folgenden Aussagen sind richtig ? (x aus 5)

- A) Der Datensatz besitzt zwei Modalwerte.
- B) Der Mittelwert des Datensatzes ist durch 3, 4 gegeben.
- C) Der Median des Datensatzes hat den Wert 3, 0.
- D) Streicht man bei obigem Datensatz den ersten Wert (also die erste 3), bleibt der Median unverändert.
- E) Die Spannweite des Datensatzes hat den Wert 6.

Aufgabe 9 (Visualisierung empirischer Verteilungen)

(5 Punkte)

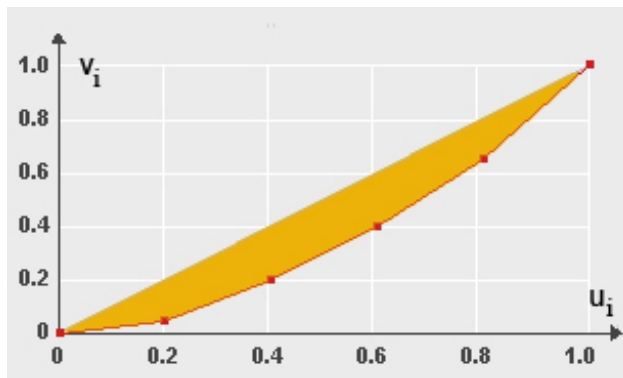
Es sei erneut der 20 Werte umfassende Datensatz aus Aufgabe 8 betrachtet. Welche der folgenden Aussagen sind richtig ? (x aus 5)

- A) Die durch obigen Datensatz definierte empirische Verteilung lässt sich z. B. anhand eines Stabdiagramms für absolute Häufigkeiten oder relative Häufigkeiten visualisieren. Beim Übergang von der einen zur anderen Darstellungsform ändert sich nur die Skalierung der Ordinatenachse.
- B) Die empirische Verteilung des Datensatzes lässt sich ohne Informationsverlust auch anhand eines Boxplots veranschaulichen.
- C) Ein Boxplot (einfachste Variante) veranschaulicht 5 Charakteristika des Datensatzes. Der Median ist genau in der Mitte der Box eingezeichnet.
- D) Anfang und Ende der Box sind durch das untere resp. das obere Quartil des Datensatzes bestimmt.
- E) Wenn man die Zeitangaben für alle 150 Beschäftigten hätte und zwar mit höherer Genauigkeit (Angaben in Minuten statt in vollen Stunden), könnte man die Daten zu Klassen zusammenfassen (z. B. Gruppierung der Daten zu 1-Stunden-Bereichen) und die Klassenbesetzungshäufigkeiten anhand eines Histogramms darstellen.

Aufgabe 10 (Konzentrationsmessung)

(5 Punkte)

In einer Region konkurrieren fünf Energieversorgungsunternehmen. Es seien $x_1 = 10$, $x_2 = 30$, $x_3 = 40$, $x_4 = 50$ und $x_5 = 70$ die Umsätze dieser Firmen im letzten Geschäftsjahr (Umsätze jeweils in Millionen Euro). Die nachstehende Abbildung zeigt die auf der Basis dieser Umsatzdaten errechnete Lorenzkurve (Polygonzug). Die Stützpunkte (u_i, v_i) der Lorenzkurve sind auf der Lorenzkurve betont. In der Tabelle neben der Grafik sind die Abszissenwerte u_i der Lorenzkurve schon eingetragen.



| i | u_i | v_i |
|-----|-------|-------|
| 0 | 0 | 0 |
| 1 | 0,2 | v_1 |
| 2 | 0,4 | v_2 |
| 3 | 0,6 | v_3 |
| 4 | 0,8 | v_4 |
| 5 | 1 | 1 |

Welche der folgenden Aussagen sind richtig ?

(x aus 5)









- A) Der Ordinatenwert v_4 des Stützpunkts $(0,8; v_4)$ der Lorenzkurve hat den Wert $v_4 = 0,65$.
- B) Der Gini-Koeffizient liefert Aussagen des Typs „ x % der Merkmalsträger teilen sich y % der Merkmalssumme“ (Messung relativer Merkmalskonzentration).
- C) Der unnormierte Gini-Koeffizient kann im Falle $n = 5$ nicht größer als 0,8 werden.
- D) Der normierte Gini-Koeffizient hat bei obigem Datensatz den Wert 0,30.
- E) Der Wert, den man für den Ordinatenwert v_2 errechnet, gibt an, welcher Anteil des Gesamtumsatzes aller 5 Energieversorger auf die beiden umsatzschwächsten Unternehmen entfällt.

Beachten Sie auch Aufgabe 41, die an die vorstehende Aufgabe direkt anknüpft.

Aufgabe 11 (Bivariate empirische Verteilungen)

(5 Punkte)

In der nachstehenden Kontingenztabelle für absolute Häufigkeiten sind Ergebnisse der Befragung einer Wählerstichprobe wiedergegeben (Daten des ZDF-Politbarometers vom Februar 1998). Die Häufigkeiten beziehen sich auf $n = 931$ befragte Personen, die ihre Parteipräferenz für den fiktiven Fall einer am nächsten Sonntag bevorstehenden Bundestagswahl geäußert hatten. Die Tabelle zeigt die Ergebnisse differenziert nach Geschlecht und jeweils mit Angabe der Randverteilungen für die Merkmale „Parteipräferenz X “ und „Geschlecht Y “.

| | | Ausprägungen von Y | | Zeilensummen ↓ | |
|--------------------|--|---|---|-------------------|----------------------|
| | |  b_1 |  b_2 | | |
| Ausprägungen von X |  a_1 | 142 | 203 | 345 | Randverteilung von X |
| |  a_2 | 194 | 160 | 354 | |
| |  a_3 | 40 | 53 | 93 | |
| |  a_4 | 35 | 20 | 55 | |
| |  a_5 | 14 | 23 | 37 | |
| |  a_6 | 20 | 27 | 47 | |
| Spaltensummen → | | 445 | 486 | 931 | |
| | | Randverteilung von Y | | | |

Welche der folgenden Aussagen sind richtig ?

(x aus 5)

- A) Die Randverteilung von X ist nichts anderes als die Häufigkeitsverteilung dieses Merkmals, die sich bei Verzicht auf die Differenzierung nach Geschlecht ergibt.
- B) Aus der Tabelle ergibt sich, dass von den 931 Befragten 445 Männer waren.
- C) Der Tabelle entnimmt man, dass von den 931 Befragten 354 Personen die SPD präferierten.
- D) Aus den Werten in der Tabelle lassen sich auch bedingte relative Häufigkeiten ermitteln, z. B. die bedingte relative Häufigkeit $f_Y(b_2|a_2)$. Letztere gibt an, welcher Anteil der befragten Personen, die sich für die SPD entschieden hatten, Frauen waren.
- E) Für die in Aufgabenteil D genannte bedingte Häufigkeit errechnet man bei Rundung auf 3 Dezimalstellen den Wert 0,352, d. h. 35,2%.

Aufgabe 12 (Zusammenhangsmessung)

(5 Punkte)

Welche der folgenden Aussagen sind richtig? (x aus 5)

- A) Der Korrelationskoeffizient r nach Bravais-Pearson misst die Stärke eines linearen Zusammenhangs zwischen zwei Merkmalen X und Y .
- B) Wenn $r = 1$ ist, bedeutet dies, dass die Datenpaare $(x_1, y_1), \dots, (x_n, y_n)$ alle auf einer steigenden oder fallenden Geraden liegen.
- C) Im Falle $r = 0$ ist noch nicht ausgeschlossen, dass zwischen den Merkmalen X und Y ein nicht-linearer Zusammenhang besteht.
- D) Der Rangkorrelationskoeffizient r_{SP} nach Spearman ist ein für ordinalskalierte Merkmale anwendbares Zusammenhangsmaß, das – wie der Korrelationskoeffizient r nach Bravais-Pearson – stets Werte zwischen -1 und $+1$ annimmt.
- E) Der Rangkorrelationskoeffizient r_{SP} lässt sich auch auf metrisch skalierte Merkmale anwenden. Die in den Daten enthaltene Information wird dann aber nicht ausgeschöpft, weil bei der Berechnung von r_{SP} nur die Rangpositionen der Werte für X resp. für Y verarbeitet werden.

Aufgabe 13 (Kombinatorik)

(5 Punkte)

Eine „faire“ Münze, also eine Münze mit gleichen Eintrittswahrscheinlichkeiten für „Kopf“ und „Zahl“, wird n -mal geworfen und die Anzahl X der Ausgänge mit „Zahl“ festgestellt. Welche der folgenden Aussagen sind richtig? (x aus 5)

- A) Die Wahrscheinlichkeit dafür, im Falle $n = 3$ höchstens einmal „Zahl“ zu erhalten, ist $0,5$.
- B) Die Wahrscheinlichkeit dafür, im Falle $n = 3$ genau einmal „Zahl“ zu erhalten, ist $0,375$.
- C) Die Wahrscheinlichkeit dafür, im Falle $n = 4$ mindestens einmal „Zahl“ zu erhalten, ist kleiner als $0,9$.
- D) Die Wahrscheinlichkeit dafür, im Falle $n = 4$ mindestens einmal „Zahl“ zu erhalten, ist genau so groß wie die Wahrscheinlichkeit höchstens dreimal „Zahl“ zu erzielen.
- E) Die Anzahl X der Ausgänge mit „Zahl“ lässt sich durch eine Binomialverteilung mit den Parametern n und $p = 0,5$ modellieren.

Aufgabe 14 (Konfidenzintervalle)

(5 Punkte)

Welche der folgenden Aussagen, die sich auf Konfidenzintervalle für den Erwartungswert $\mu = E(X)$ eines normalverteilten Merkmals X mit bekannter Varianz σ^2 beziehen, sind richtig? (x aus 5)

- A) Das Konfidenzintervall für μ zum Konfidenzniveau $1 - \alpha$ wird schmaler, wenn der Stichprobenumfang n erhöht wird.
- B) Das Konfidenzintervall für μ zum Konfidenzniveau $1 - \alpha$ wird schmaler, wenn $1 - \alpha$ erhöht, also die Irrtumswahrscheinlichkeit α verkleinert wird.
- C) Der unbekannte Parameter μ kann auch außerhalb des Konfidenzintervalls liegen.
- D) Die Grenzen eines Konfidenzintervalls lassen sich als Ausprägungen von Stichprobenfunktionen interpretieren, also als Realisationen von Zufallsvariablen.
- E) Keine der vorstehenden Aussagen ist richtig.

Aufgabe 15 (Normalverteilung; Standardnormalverteilung)

(5 Punkte)

Welche der folgenden Aussagen sind richtig? (x aus 5)

- A) Die Wahrscheinlichkeit dafür, dass eine normalverteilte Zufallsvariable eine Ausprägung x mit $x \leq 0$ annimmt, beträgt 0,5.
- B) Jede normalverteilte Zufallsvariable kann anhand einer Lineartransformation in eine standardnormalverteilte Zufallsvariable überführt werden.
- C) Der Wert der Verteilungsfunktion einer standardnormalverteilten Zufallsvariablen an der Stelle 1 stimmt mit dem Wert der Verteilungsfunktion an der Stelle -1 überein.
- D) Die Wahrscheinlichkeit $P(X \leq a)$ dafür, dass eine normalverteilte Variable X eine Ausprägung hat, die nicht größer als a ist, entspricht dem Wert der Verteilungsfunktion $F(x)$ der betreffenden Normalverteilung an der Stelle $x = a$.
- E) Ist Z eine standardnormalverteilte Zufallsvariable, so ist für diese die Bedingung $P(Z \leq a) = 0,975$ erfüllt, wenn man für a den Wert $a = 1,96$ wählt.

Aufgabe 16 (Testen, Fehler beim Testen)

(5 Punkte)

Es seien n Beobachtungen für ein Merkmal gegeben. Die Werte werden als Realisationen unabhängig identisch normalverteilter Stichprobenvariablen X_1, \dots, X_n aufgefasst (Normalverteilung mit unbekanntem Erwartungswert μ und Varianz σ^2). Getestet werden soll

$$H_0 : \mu = \mu_0 \quad \text{gegen} \quad H_1 : \mu \neq \mu_0$$

und zwar zum Signifikanzniveau α (zweiseitiger Test).

Welche der folgenden Aussagen sind richtig ? (x aus 5)

- A) Wenn man die Varianz σ^2 als bekannt annimmt, kann man den standardisierten Stichprobenmittelwert $Z = (\bar{X} - \mu_0)/\sigma_{\bar{X}}$ als Prüfgröße für den Test heranziehen (Gauß-Test). Die Dichte dieser Prüfgröße ist symmetrisch bezüglich des Nullpunkts.
- B) Die Nullhypothese wird verworfen, wenn die Prüfgröße einen Wert annimmt, der innerhalb des Ablehnungsbereichs liegt. Die Grenzen des Ablehnungsbereichs hängen vom Signifikanzniveau α ab.
- C) Die fälschliche Verwerfung der Nullhypothese H_0 wird als Fehler 1. Art bezeichnet. Die Wahrscheinlichkeit für den Eintritt eines Fehlers 1. Art hat beim hier betrachteten Test den Wert α .
- D) Wenn die Prüfgröße im Annahmehbereich liegt, kann die Nullhypothese als statistisch „bewiesen“ angesehen werden, in dem Sinne, dass ihre Gültigkeit mit einer Irrtumswahrscheinlichkeit von α als gesichert angenommen werden kann.
- E) Wenn man bei obigem Test die Varianz σ^2 resp. die Standardabweichung σ nicht als bekannt voraussetzen kann und eine Schätzung $\hat{\sigma}$ heranzieht (Schätzung von σ durch die korrigierte Stichprobenstandardabweichung), ist die resultierende Prüfgröße t -verteilt. Die Grenzen des Ablehnungsbereichs für die Nullhypothese liegen dann – bei unverändertem Signifikanzniveau α – enger zusammen.

Beachten Sie auch Aufgabe 43, die an die vorstehende Aufgabe direkt anknüpft.

Aufgabe 41 (Konzentrationsmessung)

(4 Punkte)

Geben Sie den Flächeninhalt A an, der in Aufgabe 10 in der Abbildung dunkel markiert ist. Gemeint ist also der Inhalt der Fläche, die von der Lorenzkurve und der Strecke gebildet wird, die den Nullpunkt $(0; 0)$ mit dem Punkt $(1; 1)$ verbindet. Geben Sie Ihre Antwort mit drei Nachkommastellen rechtsbündig in das Antwortfeld ein. Verwenden Sie für das **Dezimalkomma** unbedingt ein **eigenes Feld**. Falls Sie also z. B. „0,36“ errechnen, tragen Sie in die letzten fünf Felder „0,360“ ein. Vergessen Sie nicht, Ihre Antwort rechtzeitig vor Ende der Klausur auf den Markierungsbogen zu übertragen.

(numerisch)

 $A =$

| | | | | | | |
|--|--|--|--|--|--|--|
| | | | | | | |
|--|--|--|--|--|--|--|

Aufgabe 42 (Kombinatorik)

(4 Punkte)

Bei der Deutschen Meisterschaft in der Disziplin 100-m-Lauf (Herren) treten im Endlauf sechs Kandidaten K_1, K_2, \dots, K_6 gegeneinander an. Wieviele Möglichkeiten für die Verteilung der ersten drei Plätze gibt es? Tragen Sie Ihr Ergebnis rechtsbündig in das Antwortfeld ein. Übertragen Sie Ihr Ergebnis rechtzeitig vor Ende der Klausur auf den Markierungsbogen.

(numerisch)

| | | | | | |
|--|--|--|--|--|--|
| | | | | | |
|--|--|--|--|--|--|

Aufgabe 43 (Gauß-Test)

(4 Punkte)

Berechnen Sie für den Gauß-Test aus Aufgabe 16 (s. dort die Aufgabenteile A-B) für den Fall $\alpha = 0,05$ zunächst die Grenzen des Annahmebereichs für die Nullhypothese und tragen Sie dann in das Antwortfeld ein, wie weit diese beiden Werte auseinander liegen. Es ist also die Länge des Annahmebereichs gesucht. Geben Sie ihre Antwort auf vier Dezimalstellen nach dem Komma genau an. Falls Sie für den Annahmebereich das Intervall $[a; b]$ errechnen mit irgendwelchen Werten a und b und für die Länge $b - a$ des Intervalls z. B. 4,523, wäre 4,5230 in die sechs Felder einzutragen. Das **Dezimalkomma** muss also auch hier unbedingt ein **eigenes Feld** belegen.

(numerisch)

| | | | | | |
|--|--|--|--|--|--|
| | | | | | |
|--|--|--|--|--|--|

Aufgabe 44 (Wahrscheinlichkeiten bei Normalverteilung)

(4 Punkte)

Bei der Serienfertigung eines bestimmten Schraubentyps ist aus der Auswertung eines Produktionsvorlaufs bekannt, dass der Durchmesser X des Schraubenkopfs normalverteilt ist mit Erwartungswert $\mu = 6,0$ (Zielwert) und einer Standardabweichung $\sigma = 0,01$ (Angaben in mm), also $X \sim N(6; 0,01^2)$. Weicht der Schraubenkopfdurchmesser bei einer Schraube um mehr als 0,02 (mm) nach oben oder nach unten vom Zielwert ab, d. h. liegt eine beobachtete Realisation x von X nicht im Intervall $[5,98; 6,02]$, so ist die Schraube für den vorgesehenen Einsatzzweck nicht brauchbar. Sie gilt dann als Ausschuss.

Wie groß ist die Wahrscheinlichkeit P dafür, dass der Durchmesser einer Schraube im Intervall $[5,98; 6,02]$ liegt, die Schraube also kein Ausschuss ist? Geben Sie das Ergebnis auf vier Stellen nach dem Dezimalkomma genau an. Verwenden Sie für das **Dezimalkomma** ein **eigenes Feld**. Falls Sie also z. B. „0,8934“ errechnen, tragen Sie in die letzten sechs Felder „0,8934“ ein. Vergessen Sie nicht, Ihre Antwort rechtzeitig vor dem Ende der Klausur auf den Markierungsbogen zu übertragen.

(numerisch)

 $P =$

| | | | | | |
|--|--|--|--|--|--|
| | | | | | |
|--|--|--|--|--|--|

Aufgabe 45 (Kleinst-Quadrat-Schätzung)

(4 Punkte)

Für zwei Merkmale X und Y stehen 6 Beobachtungswerte $(x_1, y_1), \dots, (x_6, y_6)$ zur Verfügung, nämlich die Werte $(6, 0; 3, 0)$, $(5, 0; 3, 2)$, $(7, 0; 2, 5)$, $(7, 0; 2, 3)$, $(8, 0; 2, 0)$ und $(9, 0; 2, 0)$. Für diesen Datensatz wurde nach der Methode der kleinsten Quadrate eine Regressionsgerade

$$\hat{y} = \hat{\alpha} + \hat{\beta}x.$$

bestimmt, wobei sich für $\hat{\beta}$ aus den Daten der Wert $\hat{\beta} = -0,34$ ergab. Errechnen Sie auch den Wert $\hat{\alpha}$, der sich nach der Kleinst-Quadrat-Methode ergibt. Tragen Sie Ihr Ergebnis rechtsbündig und auf zwei Stellen nach dem Dezimalstellen genau in das Antwortfeld ein. Sie würden also vier Felder benötigen, wenn etwa „5,43“ Ihre Lösung wäre, denn das **Dezimalkomma** muss auch hier wieder ein **eigenes Feld** belegen. Vergessen Sie nicht, Ihre Antwort rechtzeitig vor dem Ende der Klausur auf den Markierungsbogen zu übertragen.

(numerisch)

 $\hat{\alpha} =$

| | | | | | |
|--|--|--|--|--|--|
| | | | | | |
|--|--|--|--|--|--|