FernUniversität in Hagen

LOTSE

Aufgabenteil zur Klausur zum

Modul 2.1 im BA-Studiengang „Politik- und Verwaltungswissenschaft“

und zum

Modul 3 im BA-Studiengang „Soziologie“

Termin: **1. März 2011, 14.00 - 18.00 Uhr**

Prüfer: **apl. Prof. Dr. H.-J. Mittag**

Abzugeben in einem ausgefüllten Klausurumschlag **ist** am Ende **nur der** maschinenauswertbare **Markierungsbogen**. Den Umschlag bitte nicht zukleben. Das Aufgabenheft, die Formelsammlung und das Konzeptpapier werden *nicht* eingesammelt.

Hinweise zur Bearbeitung der Klausur:

1. Bitte lesen Sie diese Hinweise vollständig und aufmerksam durch, bevor Sie mit der Bearbeitung beginnen. Legen Sie für die Identitätskontrolle Ihren Personalausweis und die Anmeldebestätigung neben die Klausurunterlagen. Strikt untersagt sind während der Klausur das Rauchen und die Benutzung von Mobiltelefonen und anderen Geräten, die eine Verbindung zum Internet herstellen. **Bitte schalten Sie Ihre Mobiltelefone aus**!

2. Die Klausur besteht aus 23 Aufgaben und zwar 15 Multiple-Choice-Aufgaben (Antwort-Auswahl-Verfahren) mit insgesamt 75 Punkten und 8 numerischen Aufgaben mit insgesamt 25 Punkten. Die Klausurdauer beträgt 240 Minuten.

3. Bitte kontrollieren Sie sofort, ob Sie ein vollständiges Klausurexemplar mit einem **Aufgabenteil auf weißem Papier**, eine **Formelsammlung mit Glossar und angehängtem Konzeptpapier auf gelbem Papier** sowie einen **LOTSE-Markierungsbogen** erhalten haben. Bevor Sie mit der Bearbeitung der Aufgaben 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.

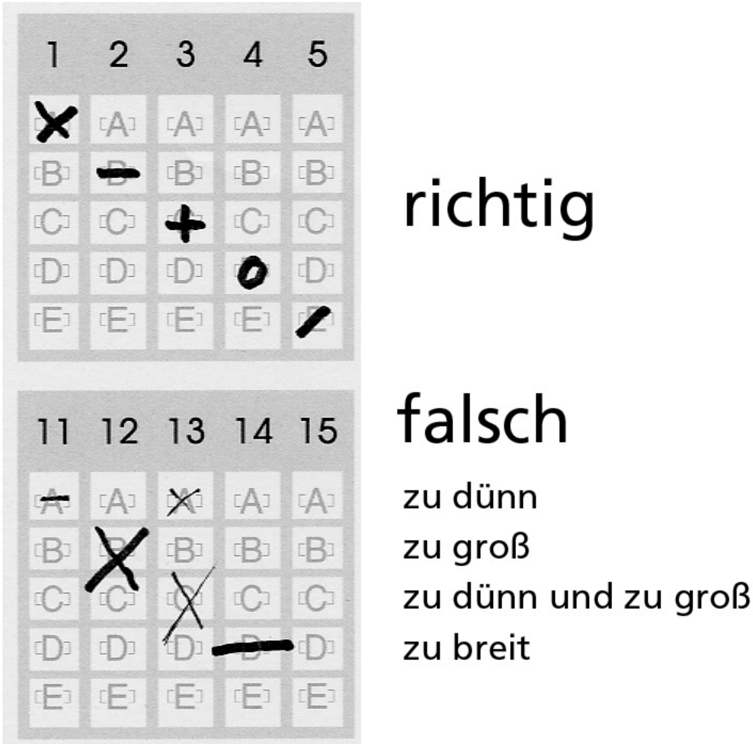
4. Für die Bewertung der Klausur sind ausschließlich Ihre Markierungen auf dem LOTSE-Markierungsbogen ausschlaggebend. Sie können nach Auswertung der Klausur online über das LOTSE-Korrektursystem einsehen, was Sie auf dem Markierungsbogen eingetragen haben und wie das System Ihre Eintragungen bewertet hat. Sie können so Ihre Antworten auch später mit der Musterlösung vergleichen.

5. 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.

6. Bei jeder Aufgabe ist die maximal erreichbare Anzahl der Punkte angegeben. Insgesamt können Sie 100 Punkte erreichen. Bei Erreichen von 50 Punkten haben Sie die Klausur auf jeden Fall bestanden.

7. Sind die *numerischen Aufgaben* richtig beantwortet, erhalten Sie die volle Punktzahl, ansonsten werden 0 Punkte oder eine Teilpunktzahl vergeben.

8. Bei den *Multiple-Choice-Aufgaben* sind **fünf Aussagen** vorgegeben, **von denen mindestens eine zutreffend ist**. Zutreffende Aussagen sind von Ihnen auf dem Markierungsbogen mit einem Strich, einem Kreuz oder einem Kreis zu kennzeichnen, falsche Aussagen sind nicht zu markieren. Wichtig ist, dass Ihre Markierungen nicht zu dünn sind und nicht in Nachbarfelder hineinreichen. Die Markierungen sind mit einem weichen Bleistift durchzuführen (empfohlen, weil von Ihnen noch änderbar) oder einem schwarzen Filzstift mittlerer Stärke. Falls Sie bei Verwendung eines Filzstifts noch Korrekturen vornehmen, müssen diese eindeutig und klar sein, damit wir sie anerkennen können. In der nachstehenden Grafik ist angedeutet, wie die Markierungen aussehen bzw. nicht aussehen sollten (die Eintragungen wurden hier willkürlich vorgenommen).



9. Die Bewertung der *Multiple-Choice-Aufgaben* erfolgt nach folgendem Prinzip: Sie erhalten

   - 1 Punkt, wenn Sie 3 der 5 vorgegebenen Antworten richtig haben,
   - 3 Punkte, wenn Sie 4 der 5 vorgegebenen Antworten richtig haben,
   - 5 Punkte, wenn Sie alle 5 Antworten richtig haben.

   Aufgaben, bei denen Sie weniger als 3 Antworten richtig haben, werden mit 0 Punkten bewertet. Aufgaben, bei denen Sie keine Markierung vornehmen, gelten als nicht bearbeitet und werden ebenfalls mit 0 Punkten bewertet. Das Verfahren berücksichtigt, dass bei geschlossenen Aufgabenformen schon durch bloßes Raten richtige Antworten erreicht werden können (siehe dazu http://www.fernuni-hagen.de/mks/lotse/gesamtbewertung.shtml).

10. Beispiel zur Bewertung der MC-Aufgaben: Sind die Aussagen A und B richtig sowie C, D und E falsch und es wurden A, B und C als richtig markiert, gibt es 3 Punkte, weil die Antworten zu A, B, D und E zutreffen.

11. Für Zwischenrechnungen können Sie das der Formelsammlung angehängte Konzeptpapier verwenden. Zwischenrechnungen gehen nicht in die Bewertung ein, weil nur der LOTSE-Bogen eingesammelt und verarbeitet wird.

12. Als Hilfsmittel ist neben der ausgeteilten **Formelsammlung** (mit Glossar) nur ein **Taschenrechner** zugelassen. Dieser darf nicht programmierbar sein und auch nicht über eine alphanumerische Tastatur verfügen. Ferner darf er keine Texte oder Formeln speichern und nicht drahtlos mit anderen Geräten kommunizieren können.

13. Täuschungen, Täuschungsversuche und andere Verstöße gegen die Prüfungsdisziplin können zum Ausschluss von der Klausur und zur Bewertung mit „nicht ausreichend (5,0)“ führen.

14. Alle Klausurteilnehmer erhalten von der FernUniversität eine Benachrichtigung, auf der die erreichte Punktzahl und die Note vermerkt sind. Die maschinelle Auswertung nimmt erfahrungsgemäß einen Zeitraum von 8 Wochen in Anspruch. Sehen Sie daher bis Ende April von Nachfragen zum Klausurergebnis ab.

Viel Erfolg bei der Klausurbearbeitung!

Multiple-Choice-Aufgaben

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: $\neg a$	P2: $a \vee \neg b$	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. Ob diese erfüllt sind, hängt davon ab, ob die Aussagen a und b wahr oder falsch sind. 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 falsch und b wahr ist, ist $P1$ erfüllt, nicht aber $P2$.

C) Wenn die Aussage a wahr und b falsch ist, sind weder $P1$ noch $P2$ erfüllt.

D) Wenn die Aussagen a und b beide falsch sind, ist $P2$ erfüllt, nicht aber $P1$.

E) Wenn $P1$ und $P2$ wahr sind, ist auch K wahr.

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 (Datengewinnung) (5 Punkte)

Welche der folgenden Aussagen sind richtig? Bei den Aufgabenteilen A und D geht es jeweils um die Beurteilung der Richtigkeit des letzten Satzes. (x aus 5)

A) Das Statistische Bundesamt (DeStatis) erfasst automatisiert die Nutzungshäufigkeit von DeStatis-Datenangeboten zu einzelnen Themenbereichen, um das Verhalten der Datennutzer zu verfolgen und auszuwerten. Diese Art der Gewinnung von Daten zum Nutzerverhalten ist ein Beispiel für nicht-reaktive Datenerhebung.

B) Wenn man allgemeine Bevölkerungsumfragen als offene Online-Befragungen organisiert, ist mit systematisch verzerrten Ergebnissen zu rechnen.

C) Das Random-Route-Verfahren ist ein Verfahren, mit dem sich Zufallsbevölkerungsstichproben ohne Rückgriff auf Namens- oder Adressdateien gewinnen lassen.

D) Bei Experimenten mit Personen, z. B. in der Psychologie, werden i. a. zwei Gruppen von Teilnehmern gebildet – eine Versuchsgruppe und eine Kontrollgruppe. Anschließend werden Einflussgrößen in beiden Gruppen planmäßig verändert.

E) „Auswahleinheit“ und „Erhebungseinheit“ sind Begriffe, die inhaltlich dasselbe bedeuten, also synonym verwendet werden können.

Aufgabe 3 (Merkmalsklassifikationen) (5 Punkte)

Welche der folgenden Aussagen sind richtig? Bei den Aufgabenteilen B und D geht es jeweils um die Beurteilung der Richtigkeit des letzten Satzes. Der aus zwei Teilaussagen bestehende letzte Satz in Aufgabenteil D gilt als richtig, wenn jede Teilaussage zutrifft. Der hier und in Aufgabenteil E verwendete Begriff „metrische Skala“ ist als Oberbegriff für „Intervallskala“, „Verhältnisskala“ und „Absolutskala“ zu verstehen. (x aus 5)

A) Operationen, die für ordinalskalierte Daten zulässig sind, sind auch für nominalskalierte Daten zulässig.

B) Bei einer Einkommenserhebung wird u. a. der Bildungsstand von Arbeitnehmern erfasst und zwar anhand des höchsten erreichten Bildungsabschlusses (Ausprägungen des Merkmals „Bildungsstand“: ohne Schulabschluss, Hauptschule, mittlere Reife, Fachhochschulreife, Abitur, akademischer Abschluss). Das Merkmal „Bildungsstand“ ist ordinalskaliert.

C) Das in Aufgabenteil B spezifizierte Merkmal „Bildungsstand“ ist diskret.

D) Die Mitarbeiter einer Firma werden gebeten, die Entfernung zwischen Wohnung und Firma anzugeben sowie das für die Fahrt zur Arbeit überwiegend benutzte Transportmedium (z. B. „PKW“ oder „Fahrrad“). Das Merkmal „Entfernung“ ist metrisch skaliert, das Merkmal „Transportmedium“ nominalskaliert.

E) Metrisch skalierte Merkmale können sowohl qualitativ als auch quantitativ sein.

Aufgabe 4 (Messen / Stichprobenverfahren) (5 Punkte)

Welche der folgenden Aussagen sind richtig? Bei Aussage D geht es um die Beurteilung des Wahrheitsgehalts des letzten Satzes. (x aus 5)

A) Die Reliabilität eines Messinstruments charakterisiert, inwieweit ein Messinstrument bei wiederholter Messung die gleichen Messwerte liefert.

B) Wenn eine Messung dem Gütekriterium „Reliabilität" genügt, genügt sie auch dem Gütekriterium der „Validität".

C) Die Klumpenauswahl ist eine zweistufige Auswahlprozedur, bei der in der ersten Stufe Teilmengen der Grundgesamtheit (sog. Klumpen) zufällig ausgewählt werden, deren Elemente dann – in der zweiten Verfahrensstufe – alle für die Datenerhebung herangezogen werden.

D) Eine Grundgesamtheit von $N = 10.000$ (10 Tausend) Personen wird bezüglich des Merkmals „Höchster erreichter Bildungsabschluss" in drei Teilpopulationen zerlegt, wobei letztere $N_1 = 2.600$, $N_2 = 5.200$ und $N_3 = 2.200$ Personen umfassen. Aus den Teilpopulationen werden dann in der zweiten Verfahrensstufe Zufallsstichproben des Umfangs $n_1 = 78$, $n_2 = 156$ resp. $n_3 = 44$ gezogen. Das damit praktizierte Auswahlverfahren repräsentiert eine geschichtete Stichprobenauswahl mit proportionaler Schichtung.

E) Das Quotenauswahlverfahren ist ein nicht-zufallsgesteuertes Verfahren zur Gewinnung einer Stichprobe, das z. B. bei Befragungen in der Markt- und Meinungsforschung Anwendung findet.

Aufgabe 5 (Definitionen) (5 Punkte)

Welche der folgenden Aussagen sind zutreffend? Bei Aussage E geht es um die Beurteilung des Wahrheitsgehalts des zweiten Satzes. (x aus 5)

A) Eine Realdefinition beinhaltet eine Aussage über Eigenschaften eines Gegenstands oder Sachverhalts.

B) Mit einer Nominaldefinition wird ein Gegenstand oder Sachverhalt (das Definiendum) durch einen anderen Gegenstand oder Sachverhalt (das Definiens) erklärt.

C) Eine Realdefinition umfasst alle Eigenschaften des Definiendums.

D) Sowohl Nominaldefinitionen als auch Realdefinitionen sind entweder richtig oder falsch.

E) Wenn man theoretische Konstrukte (z. B. die latente Variable „Leistungsmotivation") messen will, muss man zuerst eine Operationalisierung vornehmen. Letztere beinhaltet die Festlegung von Handlungsanweisungen, mit denen sich den betreffenden Variablen Ausprägungen beobachtbarer Variablen zuordnen lassen.

Aufgabe 6 (Univariate Häufigkeitsverteilungen / Kenngrößen) (5 Punkte)

Gegeben sei der folgende Datensatz für ein Merkmal X:
$5,2 \quad 6,4 \quad 4,2 \quad 4,6 \quad 4,8 \quad 3,9 \quad 6,1 \quad 7,1 \quad 4,2 \quad 7,6 \quad 6,5.$

Welche der folgenden Aussagen, die alle von diesem Datensatz ausgehen, sind richtig? Beachten Sie, dass eine aus mehreren Teilaussagen bestehende Aussage nur dann als richtig zu bewerten ist, wenn jede Teilaussage zutrifft. (x aus 5)

A) Der obige Datensatz hat einen eindeutig bestimmten Modalwert.

B) Wenn man bei obigem Datensatz den letzten Wert $(6,5)$ der Urliste um $0,6$ erhöht und gleichzeitig den ersten Wert $(5,2)$ um $0,5$ vermindert, hat dies zur Folge, dass der Mittelwert $\overline{x}$ des Datensatzes größer und der Median $\widetilde{x}$ kleiner wird.

C) Mit der in Aufgabenteil B spezifizierten Veränderung des ersten und letzten Wertes der originären Urliste verändert sich die Spannweite des Datensatzes.

D) Wenn man bei obigem Datensatz den letzten Wert $(6,5)$ streicht, werden sowohl der Mittelwert $\overline{x}$ als auch der Median $\widetilde{x}$ des Datensatzes kleiner.

E) Wenn man alle Werte des obigen Datensatzes halbiert, geht die empirische Standardabweichung s des Datensatzes auf die Hälfte ihres Ausgangswertes zurück.

Aufgabe 7 (absolute und relative Häufigkeiten) (5 Punkte)

Das Statistische Amt eines EU-Staates teilt mit, dass die Gesamtbevölkerung des betreffenden Landes Anfang 2010 bei $36,0$ Millionen gelegen habe, wobei $49,5$ % der Gesamtbevölkerung (Kinder eingerechnet) Männer seien. Ferner wird bekannt gegeben, dass der Anteil der als erwerbstätig registrierten Männer sich auf $58,0$ % belaufe, während bei den Frauen nur $44,5$ % als erwerbstätig gemeldet seien.

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

A) Die Anzahl der erwerbstätigen Frauen liegt unter $8,2$ Millionen.

B) Die Anzahl der erwerbslosen Männer liegt zwischen $7,4$ und $7,5$ Millionen.

C) Der Anteil der erwerbstätigen Männer an der Gesamtbevölkerung beträgt mehr als $29,5$ %.

D) Die Gesamtzahl der Erwerbstätigen liegt über $18,7$ Millionen.

E) Wenn man die Anzahl der Erwerbstätigen und die der Erwerbslosen zunächst für die Personen männlichen Geschlechts und danach auch für die Personen weiblichen Geschechts addiert, erhält man die absolute Randverteilung für das binäre Merkmal „Geschlecht“.

Aufgabe 8 (Konzentrationsmessung) (5 Punkte)

Die folgende Aufgabe ist adaptiert aus Toutenburg / Schomaker / Wißmann (2006), Arbeitsbuch zur deskriptiven und induktiven Statistik, Springer Verlag, Heidelberg:

An einem Gymnasium werden fünf Schüler einer Mittelstufenklasse nach ihrem monatlichen Taschengeld befragt. Dabei ergaben sich folgende Werte in Euro:

$x_1 = 20$, $x_2 = 40$, $x_3 = 50$, $x_4 = 65$ und $x_5 = 80$.

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

A) Wenn man auf der Basis des obigen Datensatzes die Lorenzkurve zeichnet – also einen Polygonzug, der den Nullpunkt mit den Punkten $(0,2; v_1)$, $(0,4; v_2)$, $(0,6; v_3)$, $(0,8; v_4)$ und $(1; 1)$ verbindet – so nimmt diese Kurve an der Stelle $0,4$ einen Wert an, der zwischen $0,21$ und $0,22$ liegt.

B) Der Wert, den man für den Ordinatenwert v_2 errechnet, gibt an, welcher Anteil des gesamten monatlichen Taschengeldes aller 5 Schüler auf die beiden Schüler mit dem geringsten monatlichen Taschengeld entfällt.

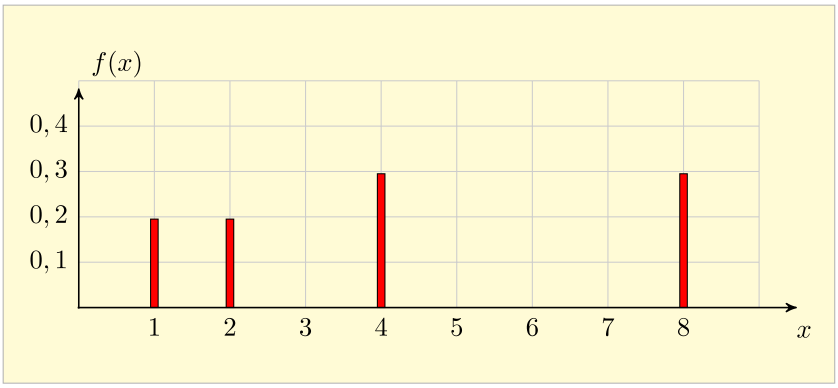
C) Der unnormierte Gini-Koeffizient hat bei obigem Datensatz einen Wert G, der unter $0,23$ liegt.

D) Der normierte Gini-Koeffizient hat bei obigem Datensatz einen Wert G^*, der zwischen $0,26$ und $0,27$ liegt.

E) Wenn man bei dem Datensatz $x_1 = 20$, $x_2 = 40$, $x_3 = 50$, $x_4 = 65$ und $x_5 = 80$ alle Werte verdoppelt, bleibt der Wert des normierten und auch des unnormierten Gini-Koffizienten unverändert.

Aufgabe 9 (Wahrscheinlichkeits- und Verteilungsfunktion) (5 Punkte)

Die nachstehende Abbildung zeigt die Wahrscheinlichkeitsfunktion $f(x)$ einer diskreten Zufallsvariablen X, die vier Ausprägungen $x_1 = 1$, $x_2 = 2$, $x_3 = 4$ und $x_4 = 8$ aufweist. Die Ausprägungen x_1 und x_2 weisen jeweils die Eintrittswahrscheinlichkeit $0,2$ auf, x_3 und x_4 je die Eintrittswahrscheinlichkeit $0,3$.



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

A) Die Verteilungsfunktion $F(x)$ der diskreten Zufallsvariablen X nimmt für $x = 2$ den Wert $0,4$ an.

B) Die Verteilungsfunktion $F(x)$ von X nimmt für $x = 3$ den Wert $0,4$ an.

C) Die Verteilungsfunktion $F(x)$ von X ist nur bis $x = 8$ definiert.

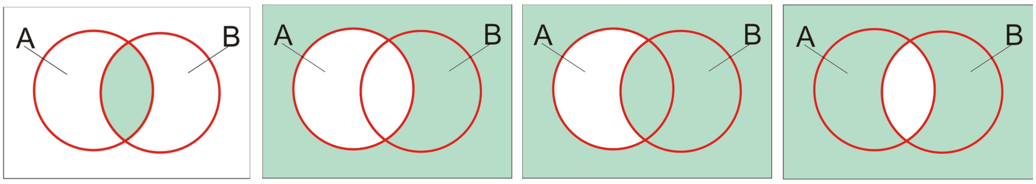
D) Der Erwartungswert $E(X)$ der Zufallsvariablen X ist größer als $3,9$, aber kleiner als $4,1$.

E) Wenn man die Zufallsvariable X der Lineartransformation $Y = X + 1$ unterzieht, so ist die Varianz der transformierten Variablen Y identisch mit der der ursprünglichen Variablen X.

Aufgabe 10 (Venn-Diagramme) (5 Punkte)

Zur Veranschaulichung von Ereignissen oder von Mengen lassen sich Venn-Diagramme heranziehen. Diese bestehen aus einem Rechteck, in dem die Ereignisse als Kreise oder Ellipsen dargestellt sind. Das Rechteck repräsentiert eine Grundgesamtheit, von der die eingezeichneten Mengen Teilmengen sind. Es bezeichnen $\overline{A}$ und $\overline{B}$ die Komplementärmengen von A und B, $A \cap B$ deren Schnittmenge und $A \cup B$ die Vereinigungsmenge von A und B.

Nachstehend sind vier Venn-Diagramme abgebildet, die sich auf die Verknüpfung zweier Ereignisse oder Mengen A und B beziehen.



Welche der folgenden Aussagen, die sich auf die obigen Diagramme beziehen, sind richtig? (x aus 5)

A) Das erste Venn-Diagramm (von links gezählt, also in der üblichen Leserichtung) veranschaulicht anhand der dunkler gefärbten Fläche die Schnittmenge von A und B, also $A \cap B$.

B) Die dunkel gefärbte Fläche im zweiten Venn-Diagramm stellt die Komplementärmenge $\overline{A}$ von A dar.

C) Im dritten Venn-Diagramm ist durch die dunkler gefärbte Fläche die Vereinigungsmenge aus $\overline{A}$ und B dargestellt, also $\overline{A} \cup B$.

D) Bildet man aus den beiden im zweiten und dritten Venn-Diagramm dargestellten Mengen die Schnittmenge, so resultiert die Menge, die im zweiten Venn-Diagramm dargestellt ist.

E) Im vierten Venn-Digramm ist durch die dunkler gefärbte Fläche die Schnittmenge der Komplementärmengen von A und B dargestellt, also $\overline{A} \cap \overline{B}$.

Aufgabe 11 (Wahrscheinlichkeiten, bedingte Wahrscheinlichkeiten) (5 Punkte)

An einer gemeinsamen Statistikklausur, die nur für Studierende der BA-Studiengänge „Politik- und Verwaltungswissenschaft“ (POL) und „Soziologie“ (SOZ) angeboten wird, haben in Berlin 80 Personen teilgenommen, davon 45 mit Studienfach POL. Nach der Klausurauswertung stellte sich heraus, dass 56 Studierende bestanden haben. Von den erfolgreichen Klausurteilnehmern waren 32 Studierende des Fachs POL.

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

A) Die Wahrscheinlichkeit dafür, dass eine nach Abschluss der Auswertungen aus dem Stapel mit allen 80 Klausuren zufällig herausgegriffene Klausur mit „nicht bestanden“ bewertet wurde, ist kleiner als $0,3$.

B) Die Wahrscheinlichkeit dafür, dass eine zufällig herausgegriffene Klausur mit „bestanden“ bewertet und einem Studierenden des Fachs POL zuzuordnen ist, ist größer als $0,38$.

C) Die bedingte Wahrscheinlichkeit dafür, dass eine der mit dem Ergebnis „bestanden“ bewerteten Klausuren die Klausur eines POL-Studierenden ist, ist größer als $0,55$.

D) Wenn man für die beiden Merkmale „Studienfach“ (Ausprägungen „POL“ und „SOZ“) und „Klausurleistung“ (Ausprägungen „bestanden“ und „nicht bestanden“) eine Vierfeldertafel für absolute Häufigkeiten mit den beiden Randverteilungen anlegt, so besteht die Randverteilung jedes Merkmals aus zwei Elementen, deren Summe jeweils 80 ergibt.

E) Wenn man für die beiden in Aufgabenteil D genannten Merkmale „Studienfach“ und „Klausurleistung“ eine Vierfeldertafel für relative Häufigkeiten anlegt, wieder mit den beiden Randverteilungen, so besteht die Randverteilung des Merkmals „Studienfach“ aus zwei Werten, von denen einer der Wert $0,4375$ ist.

Aufgabe 12 (Stetige Verteilungen) (5 Punkte)

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

A) Für eine standardnormalverteilte Zufallsvariable Z gilt, dass eine Realisation von Z mit Wahrscheinlichkeit $0,05$ außerhalb des Intervalls $[-1,96; 1,96]$ liegt.

B) Ist X eine mit Erwartungswert 1 und Varianz 4 normalverteilte Zufallsvariable, so besitzt die Wahrscheinlichkeit $P(0 \leq X \leq 2)$ dafür, dass X zwischen 0 und 2 liegt, einen Wert, der größer als $0,4$ ist.

C) Bezeichnet $[-a; a]$ ein Intervall, in das die Ausprägung einer standardnormalverteilten Zufallsvariablen Z mit Wahrscheinlichkeit $0,9$ fällt, so ist die Wahrscheinlichkeit, dass eine mit 4 Freiheitsgraden t-verteilte Zufallsvariable X in das Intervall $[-a; a]$ fällt, größer als $0,9$.

D) Die Wahrscheinlichkeit dafür, dass eine standardnormalverteilte oder t-verteilte Zufallsvariable einen positiven Wert annimmt, beträgt $0,5$.

E) Die Wahrscheinlichkeit dafür, dass eine mit $m = 10$ und $n = 15$ Freiheitsgraden F-verteilte Zufallsvariable X eine Ausprägung annimmt, die oberhalb von $2,7$ liegt, ist kleiner als $0,05$.

Aufgabe 13 (Punkt- und Intervallschätzungen) (5 Punkte)

Bei einem statistischen Experiment mit n unabhängigen Wiederholungen wird jedesmal die Ausprägung einer Variablen X festgestellt (z. B. die Augenzahl beim n-fachen Wurf eines Würfels). Man will den Erwartungswert $\mu = E(X)$ und die Varianz $\sigma^2 = V(X)$ von X unter Heranziehung der beobachteten Werte x_1, x_2, ... , x_n schätzen. Letztere lassen sich als Realisationen unabhängiger Zufallsvariablen X_1, X_2, ... ,X_n interpretieren (auch Stichprobenvariablen genannt). Aus den n Stichprobenvariablen lässt sich der Stichprobenmittelwert $\overline{X}$ bilden.

Welche der folgenden Aussagen sind richtig? Bei Aussage E geht es um die Beurteilung des Wahrheitsgehalts des zweiten Satzes. (x aus 5)

A) Der Stichprobenmittelwert $\overline{X}$ repräsentiert eine unverzerrte Schätzung für den Erwartungswert μ.

B) Falls die obige Aussage A zutrifft, gilt auch, dass der mittlere quadratische Fehler des Stichprobenmittelwerts $\overline{X}$ und die Varianz von $\overline{X}$ übereinstimmen.

C) Wenn man die Summe der quadrierten Abweichungen $(X_1 - \overline{X})^2$, $(X_2 - \overline{X})^2$, ... , $(X_n - \overline{X})^2$ bildet und diese durch n dividiert, hat man eine unverzerrte Schätzung für die Varianz σ^2 des Merkmals X.

D) Die Varianz von $\overline{X}$ geht auf ein Viertel des Ausgangswertes zurück, wenn man n verdoppelt.

E) Man kann den Erwartungswert μ auch durch Angabe eines Konfidenzintervalls schätzen. Letzteres ist ein Intervall, das stets so groß gewählt wird, dass es den unbekannten Parameter μ enthält.

Aufgabe 14 (Testen, Fehler beim Testen) (5 Punkte)

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

$$H_0: \ \mu \geq \mu_0 \qquad \text{gegen} \qquad H_1: \ \mu < \mu_0$$

und zwar zum Signifikanzniveau $\alpha = 0,05$. Wenn man die Varianz σ^2 als bekannt voraussetzt, kann man den standardisierten Stichprobenmittelwert $Z = (\overline{X} - \mu_0)/\sigma_{\overline{X}}$ als Prüfgröße für den Test heranziehen (Gauß-Test).

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

A) Die Nullhypothese wird bei obigem Test verworfen, wenn die Prüfgröße den Wert des $0,95$-Quantils der Standardnormalverteilung unterschreitet.

B) Die Wahrscheinlichkeit dafür, die Nullhypothese zu verwerfen, wenn $\mu = \mu_0$ gilt, beträgt hier $0,05$.

C) Die Wahrscheinlichkeit dafür, die Nullhypothese zu verwerfen, wenn $\mu > \mu_0$ gilt, liegt unterhalb von $0,05$.

D) Die Wahrscheinlichkeit für den Eintritt eines Fehlers 2. Art wird immer kleiner, je stärker μ den Wert μ_0 unterschreitet.

E) Die Wahrscheinlichkeit dafür, die Nullhypothese zu verwerfen, wenn $\mu = \mu_0$ gilt, ist identisch mit dem Wert der Gütefunktion des Tests an der Stelle $\mu = \mu_0$.

Aufgabe 15 (Korrelationsmessung, Regressions- und Varianzanalyse) (5 P.)

Welche der folgenden Aussagen sind richtig? Bei den Aussagen A und B geht es darum, den Wahrheitsgehalt des jeweils letzten Satzes zu bewerten. (x aus 5)

A) Für einen acht Beobachtungspaare $(x_1; y_1)$, $(x_2; y_2)$, ... ,$(x_8; y_8)$ umfassenden Datensatz wurde $\overline{x} = 3$ und $\overline{y} = 7$ errechnet sowie

$$\sum_{i=1}^{8}(x_i - \overline{x})^2 = 4,1; \qquad \sum_{i=1}^{8}(y_i - \overline{y})^2 = 12; \qquad \sum_{i=1}^{8}(x_i - \overline{x})(y_i - \overline{y}) = 2,3.$$

Wenn man unterstellt, dass zwischen x_i und y_i $(i = 1, 2, ..., 8)$ ein linearer Zusammenhang besteht, kann man diesen durch das Regressionsmodell $y_i = \alpha + \beta x_i + u_i$ beschreiben und die Regressionskoeffizienten nach der Methode der kleinsten Quadrate schätzen. Die resultierende Regressionsgerade $\widehat{y} = \widehat{\alpha} + \widehat{\beta}x$ schneidet die y-Achse innerhalb des Intervalls $[5,0; 5,1]$.

B) Die Güte der Anpassung der mit der Kleinst-Quadrat-Methode erhaltenen Regressionsgeraden an den Datensatz lässt sich anhand des Bestimmtheitsmaßes bewerten. Für dieses errechnet man mit den Angaben aus Aufgabenteil A einen Wert, der im Intervall $[0,10; 0,12]$ liegt.

C) Der Korrelationskoeffizient für den Datensatz aus Aufgabenteil A ist positiv.

D) Bei Anwendung der Methode der kleinsten Quadrate wird die Regressionsgerade so bestimmt, dass die Summe der quadrierten Residuen Null ist.

E) Bei der einfaktoriellen Varianzanalyse wird die abhängige Variable als diskret modelliert.

Numerische Aufgaben

Aufgabe 41 (Kombinatorik) (3 Punkte)

Mit einer „fairen“ Münze, also einer Münze mit gleichen Eintrittswahrscheinlichkeiten für „Zahl“ und „Kopf“, wird 8-mal nacheinander geworfen. Wie groß ist die Wahrscheinlichkeit mindestens 5-mal „Zahl“ zu erhalten? Tragen Sie Ihr Ergebnis, also ein Wert aus dem Intervall $[0; 1]$, auf *vier Stellen nach dem Dezimalkomma* genau rechtsbündig in das Antwortfeld ein. Das **Dezimalkomma** belegt ein **eigenes Feld**. Vergessen Sie nicht, Ihre Antwort rechtzeitig vor dem Ende der Klausur auf den Markierungsbogen zu übertragen.

(numerisch)

Aufgabe 42 (Kombinatorik) (3 Punkte)

Den Mitarbeitern einer Firma wird als Intranet-Paßwort eine nur aus Großbuchstaben bestehende Buchstabenfolge zugewiesen, wobei Buchstaben auch mehrfach auftreten dürfen. Wieviele Mitarbeiter könnte man maximal anhand solcher Buchstabenfolgen unterscheiden, wenn für jeden Mitarbeiter genau 4 Großbuchstaben aus der Teilmenge $\{A, B, C, D, E, F\}$ des Alphabets verwendet werden, also z. B. $BAFB$, $AFBB$ oder $AECD$?

Tragen Sie die von Ihnen errechnete Anzahl, also ein ganzzahliges Ergebnis, rechtsbündig in das Antwortfeld ein. Übertragen Sie Ihr Ergebnis rechtzeitig vor Ende der Klausur auf den Markierungsbogen.

(numerisch)

Aufgabe 43 (Lotto in Schweden) (3 Punkte)

Beim deutschen Lotto „6 aus 49“ werden 6 Kugeln aus einer Trommel gezogen, die 49 fortlaufend nummerierte Kugeln enthält (Ziehen ohne Zurücklegen). In Schweden wird „7 aus 35“ gespielt, also nur 35 Kugeln verwendet, von denen dann 7 gezogen werden. Bezeichne X die Anzahl der Richtigen bei der schwedischen Lottovariante. Berechnen Sie den Erwartungswert $\mu = E(X)$ der Zufallsvariablen X.

Tragen Sie Ihr Ergebnis rechtsbündig und auf *drei* Nachkommastellen genau in das Antwortfeld ein. Das **Dezimalkomma** belegt ein **eigenes Feld**. Übertragen Sie Ihr Ergebnis rechtzeitig vor Ende der Klausur auf den Markierungsbogen.

(numerisch) $\mu =$

Aufgabe 44 (Gewinnwahrscheinlichkeit beim „Glücksrad“) (4 Punkte)

Bei einem Stadtteilfest ist eine als „Glücksrad“ angesprochene drehbare Scheibe installiert, die in fünf gleich große Teile mit den Farben Schwarz, Rot, Weiß, Gelb und Grün unterteilt ist. Jeder Besucher des Festes darf das Rad 3x nacheinander drehen und bei jedem Drehversuch wird die Farbe festgestellt, die am Ende oben steht. Wenn bei allen drei Versuchen „Grün“ resultiert, gibt es ein Freiabonnement für das Stadttheater für die nächste Saison.

Wie groß ist die Wahrscheinlichkeit, dass ein an diesem Spiel teilnehmender Festbesucher nach seinen drei Versuchen ein Freiabonnement erhält? Tragen Sie Ihr Ergebnis, also ein Wert aus dem Intervall $[0; 1]$, wieder rechtsbündig und auf *vier* Nachkommastellen genau in das Antwortfeld ein. Das **Dezimalkomma** belegt auch hier wieder ein **eigenes Feld**. Übertragen Sie Ihr Ergebnis rechtzeitig vor Ende der Klausur auf den Markierungsbogen.

(numerisch) | | | | | | | |

Aufgabe 45 (Verteilungsfunktion einer Normalverteilung) (2 Punkte)

Eine Zufallsvariable X sei normalverteilt mit Erwartungswert $\mu = 5$ und Varianz $\sigma^2 = 1$. Die Verteilung von X kann man sowohl durch die Dichtefunktion $f(x)$ als auch durch die Verteilungsfunktion $F(x)$ charakterisieren. Welchen Wert nimmt $F(x)$ an der Stelle $x = 5$ an?

Geben Sie das Ergebnis auf *vier Stellen nach dem Dezimalkomma* genau an. Verwenden Sie für das **Dezimalkomma** ein **eigenes Feld**.

(numerisch) $F(5) =$ | | | | | | | |

Aufgabe 46 (Wahrscheinlichkeiten bei Normalverteilung) (3 Punkte)

Es sei erneut eine Zufallsvariable X betrachtet, die normalverteilt ist mit Erwartungswert $\mu = 5$ und Varianz $\sigma^2 = 1$. Bestimmen Sie einen Wert a, für den $P(X \leq a) = 0,1$ gilt, also einen Wert a, der mit Wahrscheinlichkeit $0,1$ nicht überschritten wird.

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. 0,2471 errechnen, tragen Sie in die letzten sechs Felder 0,2471 ein. Vergessen Sie nicht, Ihre Antwort rechtzeitig vor dem Ende der Klausur auf den Markierungsbogen zu übertragen.

(numerisch) $a =$ | | | | | | | |

Aufgabe 47 (lineares Regressionsmodell / KQ-Schätzung) (4 Punkte)

Für fünf Besucher eines Restaurants wurden der am Ende zu zahlende Rechnungsbetrag X und das gezahlte Trinkgeld Y erfasst. Dies führte zu folgenden Beobachtungsdaten $(x_1; y_1)$, $(x_2; y_2)$, ... ,$(x_5; y_5)$ für die beiden Merkmale X und Y:

$$(12; 1); \quad (15; 0); \quad (16; 1); \quad (19; 1); \quad (23; 2).$$

Wenn man davon ausgeht, dass zwischen dem Rechnungsbetrag und dem gezahlten Trinkgeld ein linearer Zusammenhang $y = \alpha + \beta x$ besteht, kann man auf der Basis des obigen kleinen Datensatzes unter Heranziehung der KQ-Schätzmethode eine Regressionsgerade $\widehat{y} = \widehat{\alpha} + \widehat{\beta} x$ bestimmen.

Ermitteln Sie die Steigung $\widehat{\beta}$ dieser Regressionsgeraden und tragen Sie Ihr Ergebnis auf *vier* Stellen nach dem Dezimalkomma genau rechtsbündig in das Antwortfeld ein. Verwenden Sie für das **Dezimalkomma** ein **eigenes Feld**. Vergessen Sie nicht, Ihre Antwort rechtzeitig vor dem Ende der Klausur auf den Markierungsbogen zu übertragen.

(numerisch) $\widehat{\beta} =$ | | | | | | | |

Aufgabe 48 (Gauß-Test) (3 Punkte)

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

$$H_0: \ \mu \leq \mu_0 \qquad \text{gegen} \qquad H_1: \ \mu > \mu_0$$

und zwar zum Signifikanzniveau $\alpha = 0,01$. Wenn man die Varianz σ^2 als bekannt voraussetzt, kann man den standardisierten Stichprobenmittelwert $Z = (\overline{X} - \mu_0)/\sigma_{\overline{X}}$ als Prüfgröße für den Test heranziehen (Gauß-Test). Man wird die Nullhypothese verwerfen, wenn der aus den Daten zu errechnende Wert der Prüfgröße Z einen bestimmten kritischen Wert überschreitet. Bestimmen Sie diesen kritischen Wert.

Tragen Sie Ihr Ergebnis auf *vier* Stellen nach dem Dezimalkomma genau rechtsbündig in das Antwortfeld ein. Verwenden Sie für das **Dezimalkomma** ein **eigenes Feld**. Vergessen Sie nicht, Ihre Antwort rechtzeitig vor dem Ende der Klausur auf den Markierungsbogen zu übertragen.

(numerisch) | | | | | | |