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: 6. März 2012, 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. Während der Klausur sind das Rauchen und die Benutzung von Mobiltelefonen und anderen Geräten, die eine Verbindung zum Internet herstellen können, strikt untersagt. **Bitte schalten Sie Ihr Mobiltelefon 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.

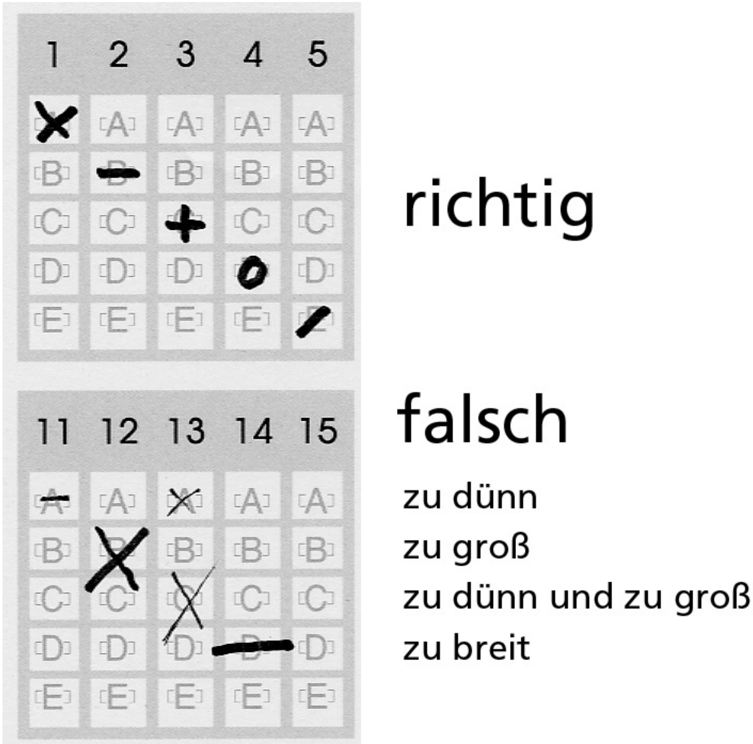
4. Für die Bewertung der Klausur sind ausschließlich Ihre Markierungen auf dem LOTSE-Markierungsbogen ausschlaggebend. Sie können nach Auswertung der Klausur online über das LOTSE-Korrektursystem einsehen, was Sie auf dem Markierungsbogen eingetragen haben und wie das System Ihre Eintragungen bewertet 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 i. d. R. 0 Punkte 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 dem 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: $a \vee \neg b$	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 $P1$ und $P2$ wahr sind, ist auch K wahr, d. h. die Konklusion ist korrekt.

C) Wenn die Aussage a falsch und b wahr ist, ist $P2$ erfüllt, nicht aber $P1$.

D) Wenn die Aussage a wahr und b falsch ist, ist $P1$ erfüllt, nicht aber $P2$.

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

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 (Merkmalsklassifikationen) (5 Punkte)

Welche der folgenden Aussagen sind richtig? Der Begriff „metrische Skala“ ist als Oberbegriff für „Intervallskala“ und „Verhältnisskala“ (letztere einschließlich des Sonderfalls „Absolutskala“) zu verstehen. (x aus 5)

A) Bei einer Nominal- und auch bei einer Ordinalskala ist die Bildung von Differenzen nicht zulässig.

B) Operationen, die für ordinalskalierte Daten zulässig sind, sind ebenso für metrisch skalierte Daten zulässig.

C) Das Merkmal „Art der Heizung“ (mit den Ausprägungen „Gas“, „Elektro“, „Öl“ und „Solar“) ist ein ordinalskaliertes Merkmal.

D) Das Merkmal „Art der Heizung“ (mit den o. g. Ausprägungen) ist ein diskretes Merkmal.

E) Die bei einem Volksmarathonlauf erreichten Plätze sind Ausprägungen eines metrisch skalierten Merkmals.

Aufgabe 3 (Messung, Informationsgewinnung) (5 Punkte)

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

A) In der empirischen Sozialforschung versteht man unter „Operationalisierung“ eine Handlungsanweisung, die es ermöglicht, nicht direkt beobachtbare theoretische Konstrukte einer Messung zugänglich zu machen.

B) Undercoverage ist ein Selektionsfehler, der bei stichprobenbasierten Datenerhebungen auftreten kann. Er entsteht, wenn nicht alle Elemente der Population, aus der eine Stichprobe gezogen werden soll, tatsächlich bei der Stichprobenziehung berücksichtigt werden.

C) Ein Gütekriterium für Messungen ist die Validität. Diese charakterisiert, inwieweit wirklich das gemessen wird, was gemessen werden soll.

D) Aus der Reliabilität einer Messung folgt stets auch deren Validität.

E) Das Random-Route-Verfahren ist ein bei der Befragung von Haushalten durch Interviewer angewendetes mehrstufiges Auswahlverfahren. Bei diesem wird der Startpunkt des Laufwegs des Interviewers zufällig ausgewählt, alle weiteren Verfahrensschritte sind dem Interviewer aber vorgegeben.

Aufgabe 4 (Experimente, Stichprobenverfahren) (5 Punkte)

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

A) Ein Ansatz zur empirischen Überprüfung von Forschungshypothesen ist die Durchführung von Experimenten. Bei Experimenten mit Menschen, z. B. in der Medizin oder der Psychologie, werden i. a. zwei Gruppen von Teilnehmern gebildet (manchmal auch mehr als zwei Gruppen). In beiden genannten Gruppen werden dann Einflussgrößen planmäßig verändert.

B) Die Klumpenauswahl ist eine zweistufige Auswahlprozedur, bei der in der ersten Stufe Teilmengen der Grundgesamtheit (sog. Klumpen oder Cluster) zufällig ausgewählt werden.

C) Die Klumpenauswahl ist eine zweistufige Auswahlprozedur, bei der in der zweiten Verfahrensstufe aus den ausgewählten Teilmengen der Grundgesamtheit Zufallsstichproben gezogen 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. $n_3 = 66$ gezogen. Das damit praktizierte Auswahlverfahren repräsentiert eine geschichtete Stichprobenauswahl mit proportionaler Schichtung.

E) Bei interviewbasierten Datenerhebungen in der Markt- und Meinungsforschung wird oft das Quotenauswahlverfahren bei der Bestimmung der zu interviewenden Personen herangezogen. Bei diesem Verfahren wird dem Interviewer genau vorgegeben, welche Personen zu interviewen sind.

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

Gegeben sei der folgende Datensatz für ein Merkmal X:

$4,9 \quad 6,2 \quad 4,2 \quad 4,9 \quad 3,9 \quad 4,9 \quad 4,2 \quad 6,0.$

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) Der Median $\widetilde{x}$ des obigen Datensatzes ist kleiner als dessen Mittelwert $\overline{x}$.

C) Wenn man bei obigem Datensatz den letzten Wert $(6,0)$ der Urliste um $0,4$ erhöht, hat dies zur Folge, dass sowohl der Mittelwert $\overline{x}$ als auch der Median $\widetilde{x}$ des Datensatzes größer werden.

D) Wenn man bei dem obigem Datensatz den ersten Wert $(4,9)$ streicht, bleibt der Median $\widetilde{x}$ unverändert, nicht aber der Mittelwert $\overline{x}$.

E) Mit der in Aufgabenteil C spezifizierten Veränderung des letzten Wertes der eingangs aufgeführten Urliste (Erhöhung dieses Wertes von $6,0$ auf $6,4$) verändert sich die Spannweite des Datensatzes.

Aufgabe 6 (Visualisierung empirischer Verteilungen) (5 Punkte)

Welche der folgenden Aussagen sind richtig? In Aufgabenteil A geht es um die Beurteilung des Wahrheitsgehalts des letzten Satzes. (x aus 5)

A) Die Lottozahlen, die in Deutschland im Jahr 2011 beim Lotto „6 aus 49“ gezogen wurden, definieren einen Datensatz, der Zahlen im Bereich von 1 bis 49 umfasst. Die durch den Datensatz definierte empirische Verteilung lässt sich anhand eines Stabdiagramms für absolute Häufigkeiten oder relative Häufigkeiten visualisieren. Beim Übergang von der einen zur anderen Darstellungsform ändert sich nur die Skalierung der Achse, an der die Häufigkeiten abgetragen werden (Ordinatenachse).

B) Die empirische Verteilung des Datensatzes „Gezogene Lottozahlen im Jahr 2011“ lässt sich ohne Informationsverlust auch anhand eines Boxplots veranschaulichen.

C) Bei dem Boxplot für den genannten Datensatz liegt der Median nicht notwendigerweise in der Mitte der Box.

D) Bei einem Boxplot sind Anfang und Ende der Box durch das untere resp. das obere Quartil des Datensatzes bestimmt.

E) Wenn man die Lottozahlen in 7-er Gruppen einteilt (Gruppen $1-7$, $8-14$, ... , $43-49$) und die Besetzungshäufigkeiten für die Gruppen visualisiert, resultiert ein Histogramm.

Aufgabe 7 (Konzentrationsmessung) (5 Punkte)

In der EU-27 teilen sich 5 Unternehmen, die hier mit U1, U2, ..., U5 bezeichnet seien, ein bestimmtes Marktsegment im Bereich der beruflichen Weiterbildung. Das Umsatzvolumen aller 5 Unternehmen im Geschäftsjahr 2011 belief sich auf insgesamt 40 Milliarden Euro. Davon entfallen auf U1 nur 10 %, auf U2 hingegen 30 %, während U3, U4 und U5 je 20 % des Gesamtumsatzes realisieren.

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 $0,28$ und $0,29$ liegt.

B) Der Wert, den man für den Ordinatenwert v_2 errechnet, gibt an, welcher Anteil des gesamten Aktienpakets auf diejenigen 2 Unternehmen entfällt, die die kleinsten Anteile am gesamten Umsatzvolumen in Höhe von 40 Milliarden Euro halten.

C) Allgemein gilt, dass der unnormierte Gini-Koeffizient G für einen Datensatz des Umfangs $n = 5$ den Wert $0,8$ nicht überschreiten kann.

D) Der normierte Gini-Koeffizient hat bei obigem Datensatz einen Wert G^*, der zwischen $0,19$ und $0,21$ liegt.

E) Wenn das Unternehmen U5 aus dem Markt ausschiede und der Umsatzanteil von U5 sich gleichmäßig auf die vier im Markt verbleibenden Unternehmen verteilte (Umsatzzuwachs von je 5 % für U1, ... , U4), würde der Wert des dann resultierenden normierten Gini-Koeffizienten G^* oberhalb von $0,21$ liegen.

Aufgabe 8 (Zusammenhangsmessung) (5 Punkte)

Welche der folgenden Aussagen sind richtig? Die Aussagen A - D beziehen sich auf empirische Zusammenhangsmaße, sind also aus Beobachtungsdaten errechenbar. Aussage E bezieht sich hingegen auf ein theoretisches Zusammenhangsmaß, also auf ein Zusammenhangsmaß für Zufallsvariablen. In Aufgabenteil D geht es um die Beurteilung des Wahrheitsgehalts des letzten Satzes. (x aus 5)

A) Der Korrelationskoeffizient r nach Bravais-Pearson misst die Stärke eines linearen oder nicht-linearen Zusammenhangs zwischen zwei Merkmalen X und Y.

B) Wenn $r = 1$ ist, bedeutet dies, dass die Datenpaare $(x_1, y_1), ..., (x_n, y_n)$ alle auf einer steigenden Geraden liegen.

C) Der Rangkorrelationskoeffizient r_{SP} lässt sich auch auf nominalskalierte Merkmale anwenden.

D) Bei einer Studie wurden für alle Mitglieder einer Gruppe von Senioren jeweils zwei Merkmale X und Y erfasst. Das Merkmal X weise die Ausprägungen a_1 und a_2, das Merkmal Y die Ausprägungen b_1 und b_2 auf. Die beobachteten Häufigkeiten für die einzelnen Ausprägungen sind in der nachstehenden Tabelle zusammengefasst:

	b_1	b_2
a_1	20	25
a_2	40	15

Mit diesen Werten erechnet sich für den χ^2-Koeffizienten ein Wert, der unterhalb von $8,5$ liegt.

E) Die Kovarianz zweier Zufallsvariablen X und Y ist ein Zusammenhangsmaß, das nur Werte zwischen -1 und $+1$ annehmen kann.

Aufgabe 9 (Randverteilungen, bedingte Wahrscheinlichkeiten) (5 Punkte)

An einer Hochschule wurde für Studierende zweier Bachelorstudiengänge, die hier mit POL und SOZ abgekürzt seien, eine gemeinsame Methodenklausur angeboten. An der Klausur nahmen insgesamt 236 Studierende teil, darunter 184 aus dem Studiengang POL. Unter den 236 Klausurteilnehmern waren 120 Frauen. Von den weiblichen Klausurnehmern waren 24 im Studiengang SOZ eingeschrieben.

Welche der folgenden Aussagen sind richtig? Gehen Sie davon aus, dass kein Klausurteilnehmer in beiden Studiengängen gleichzeitig eingeschrieben ist. (x aus 5)

A) Weniger als ein Viertel aus der Gruppe aller männlichen Klausurteilnehmer war im Studiengang SOZ eingeschrieben.

B) Wählt man aus aus der Gesamtpopulation aller Klausurteilnehmer eine Person zufällig aus, so liegt die Wahrscheinlichkeit, dass diese SOZ studiert, über $0,23$.

C) Wählt man aus der Gesamtpopulation aller Klausurteilnehmer eine Person zufällig aus, so liegt die Wahrscheinlichkeit dafür, dass eine im Studiengang POL eingeschriebene Frau ausgewählt wird, zwischen $0,39$ und $0,41$.

D) Wählt man aus der Teilpopulation der Studierenden mit Studienfach POL eine Person zufällig aus, so liegt die Wahrscheinlichkeit dafür, dass die Zufallsauswahl auf einen Mann fällt, unter $0,46$.

E) Die Quote der weiblichen Klausurteilnehmer lag im Fach POL niedriger als beim Fach SOZ.

Aufgabe 10 (Wahrscheinlichkeiten beim Roulettespiel) (5 Punkte)

Beim Roulettespiel ist der Ausgang durch eine diskrete Zufallsvariable X mit 37 Ausprägungen gegeben, nämlich den Ausprägungen $x_1 = 0, x_2 = 1, \ldots, x_{37} = 36$. Alle 37 Ausprägungen besitzen dieselbe Eintrittswahrscheinlichkeit.

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

A) Die Verteilungsfunktion $F(x)$ von X nimmt für $x = 0$ den Wert 0 an.

B) Für den Wert $F(4)$, den die Verteilungsfunktion $F(x)$ für $x = 4$ annimmt, gilt $F(4) > 0,12$.

C) Die Verteilungsfunktion $F(x)$ nimmt an der Stelle $x = 36$ den Wert 1 an.

D) Die Wahrscheinlichkeit dafür, dass der Ausgang beim einmaligen Roulettespiel eine zweistellige Zahl ist, liegt zwischen $0,72$ und $0,74$.

E) Die Wahrscheinlichkeit dafür, beim zweimaligen Roulettespiel in beiden Spielen die Zahl 10 zu realisieren, ist genau das Doppelte der Wahrscheinlichkeit dafür, beim einmaligen Roulettespiel eine 10 zu erzielen.

Aufgabe 11 (Zufallsvariablen) (5 Punkte)

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

A) Die Summe zweier unabhängiger und standardnormalverteilter Zufallsvariablen ist nicht standardnormalverteilt.

B) Das $0,9$-Quantil einer mit 4 Freiheitsgraden t-verteilten Zufallsvariablen ist größer als das $0,9$-Quantil einer standardnormalverteilten Zufallsvariablen.

C) Es sei X eine mit $m = 4$ und $n = 15$ Freiheitsgraden F-verteilte Zufallsvariable. Das $0,95$-Quantil dieser F-Verteilung ist kleiner als $2,9$.

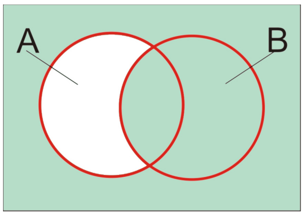
D) Es bezeichne $F(x)$ die Verteilungsfunktion einer normalverteilten Zufallsvariablen. Die Wahrscheinlichkeit $P(X \leq 1)$ dafür, dass X eine Ausprägung annimmt, die nicht größer als 1 ist, ist dann durch $F(1)$ gegeben, also durch den Wert der Verteilungsfunktion an der Stelle $x = 1$.

E) Eine mit 8 Freiheitsgraden χ^2-verteilte Zufallsvariable nimmt mit Wahrscheinlichkeit $0,1$ eine Ausprägung an, die oberhalb von $13,362$ liegt.

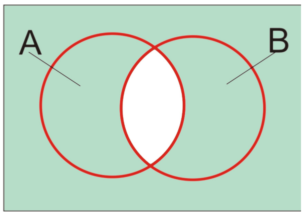
Aufgabe 12 (Venn-Diagramme / Kombinatorik / diskrete Verteilungen) (5 Punkte)

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

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



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



C) Mit einem fairen Würfel, also einem Würfel mit gleichen Eintrittswahrscheinlichkeiten für alle Augenzahlen, werde 8-mal in Folge gewürfelt. Bezeichne X die Anzahl der Ausgänge mit einer Augenzahl, die kleiner als 3 ist. Der Erwartungswert von X liegt zwischen $2,6$ und $2,7$.

D) Beim Würfeln mit zwei fairen Würfeln, liegt die Wahrscheinlichkeit dafür, eine Augensumme zu erhalten, die nicht größer als 10 ist, zwischen $0,85$ und $0,90$.

E) Auf einem Stadtteilfest wird ein Glücksrad mit 5 Feldern (Kreissektoren) gleicher Größe eingesetzt. Interessierte Besucher dürfen das Rad dreimal drehen (drei Spiele) und könnnen bei jedem Spiel einen Preis erhalten. Ob ein Preis zugesprochen wird, hängt davon ab, welcher Sektor nach einem Spiel jeweils am Ende oben steht. Nur bei 2 der 5 Sektoren ist ein Preis zu bekommen. Die Wahrscheinlichkeit dafür, nach 3 Spielen keinen Gewinn zu erhalten, ist kleiner als $0,23$.

Aufgabe 13 (Schätzen von Modellparametern) (5 Punkte)

Gegeben seien n Stichprobenwerte, die als Ausprägungen x_1, x_2, ... ,x_n unabhängiger $N(\mu;\sigma^2)$-verteilter Zufallsvariablen X_1, X_2, ... ,X_n interpretiert werden. Aus den Zufallsvariablen X_1, X_2, ... ,X_n leiten sich der Stichprobenmittelwert $\overline{X}$ und die empirische Varianz S^2 ab – letztere definiert als Summe der quadrierten Mittelwertabweichungen $(X_i - \overline{X})^2$, dividiert durch n.

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

A) Wenn man den Erwartungswert μ anhand des Stichprobenmittelwerts $\overline{X}$ schätzt, hat man mit $\overline{X}$ eine Schätzfunktion, deren Varianz auf ein Viertel des ursprünglichen Werts zurückgeht, wenn man den Stichprobenumfang verdoppelt.

B) Die Stichprobenfunktion S^2 liefert eine verzerrte Schätzung für die Varianz σ^2.

C) Der mittlere quadratische Fehler (MSE) und die Varianz der Schätzfunktion $\overline{X}$ stimmen überein.

D) Man kann den Erwartungswert μ eines normalverteilten Merkmals auch durch Angabe eines Konfidenzintervalls schätzen. Ein solches Intervall ist so konstruiert, dass es μ einschließt.

E) Die Grenzen von Konfidenzintervallen sind zufallsabhängig.

Aufgabe 14 (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 $\overline{y} = 5$ errechnet sowie

$$\sum_{i=1}^{8}(x_i - \overline{x})^2 = 4,5; \qquad \sum_{i=1}^{8}(y_i - \overline{y})^2 = 5,8; \qquad \sum_{i=1}^{8}(x_i - \overline{x})(y_i - \overline{y}) = 2,4.$$

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 $[3,3; 3,5]$.

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 $[0,23; 0,25]$ liegt.

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

D) Würde man für das Bestimmtheitsmaß den Wert $0,48$ errechnen, würde dies beinhalten, dass 48 % der Gesamtvariation des Datensatzes durch das verwendete Regressionsmodell erklärt ist.

E) Bei der einfaktoriellen Varianzanalyse werden sowohl die unabhängige als auch die abhängige Variable als diskret modelliert.

Hinweis:

Wenn man die in Aufgabenteil A angegebenen Summen jeweils durch den Umfang $n = 8$ des Datensatzes dividiert, erhält man die empirischen Varianzen s_x^2 und s_y^2 bzw. die empirische Kovarianz s_{xy}.

Aufgabe 15 (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

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

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) Wird bei obigem Test für die Prüfgröße ein Wert ermittelt, der den Wert des $0,95$-Quantils der Standardnormalverteilung überschreitet, wird die Nullhypothese verworfen.

B) Wird bei obigem Test für die Prüfgröße ein Wert ermittelt, der kleiner als $-1,96$ ist, wird die Nullhypothese verworfen.

C) Die Wahrscheinlichkeit dafür, einen Fehler 1. Art zu begehen, beträgt bei obigem Test $0,05$.

D) Wenn $\mu > \mu_0$ gilt, ist ein Fehler 2. Art möglich, nicht aber ein Fehler 1. Art.

E) Die Gütefunktion $G(\mu)$ des Tests hat an der Stelle $\mu = \mu_0$ den Wert $0,05$.

Hinweis:

Es folgen einige numerische Aufgaben, die nur aus technischen Gründen die Aufgabennummern 41 - 48 tragen.

Numerische Aufgaben

Bei den folgenden Aufgaben sind die Antworten jeweils rechtsbündig mit der angegebenen Genauigkeit in die Lösungsfelder einzutragen. Bei negativen Ergebnissen ist das Vorzeichen (Minuszeichen) in das Feld vor der ersten Ziffer einzutragen.

Aufgabe 41 (Rangkorrelationskoeffizient) (3 Punkte)

Es sei angenommen, dass - in Konkurrenz zu den etablierten Ratingagenturen - zwei neue europäische Ratingagenturen A und B ihre Arbeit aufgenommen haben und u. a. das Ausfallrisiko für Staatsanleihen bewerten. Die beiden neuen Ratingagenturen beurteilen unabhängig voneinander das kurzfristige Ausfallrisiko von Staatsanleihen für vier Länder der Eurozone. Die Risikobewertung wird anhand einer 12-stufigen Ratingskala vorgenommen. Die Stufen seien hier mit 1, .. ,12 codiert, wobei die Punktzahl 12 die schlechteste Bewertung darstelle (höchstes Ausfallrisiko) und 1 die beste Bewertung. Die Ergebnisse der Bewertungen sind nachstehend ausgewiesen.

Land i	Agentur A Bewertung x_i	Agentur B Bewertung y_i
1	3	4
2	7	9
3	11	10
4	9	8

Untersuchen Sie anhand des Rangkorrelationskoeffizienten r_{SP} von Spearman, ob zwischen den Bewertungen der beiden Ratingagenturen ein Zusammenhang besteht. Tragen Sie Ihr Ergebnis auf *zwei* Stellen nach dem Dezimalkomma genau rechtsbündig in das Antwortfeld ein. Verwenden Sie für das **Dezimalkomma** ein **eigenes Feld**. Übertragen Sie Ihr Ergebnis rechtzeitig vor Ende der Klausur auf den Markierungsbogen.

(numerisch) $r_{SP} =$ | | | | | | |

Aufgabe 42 (mehrfacher Münzwurf) (3 Punkte)

Mit einer „fairen" Münze, also einer Münze mit gleichen Eintrittswahrscheinlichkeiten für „Zahl" und „Kopf", wird 6-mal nacheinander geworfen. Wie groß ist die Wahrscheinlichkeit mindestens 4-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 43 (Kombinatorik) (3 Punkte)

Im Rahmen eines Betriebsjubiläums bei einem kleinen Unternehmen werden unter den 20 Mitarbeitern drei Preise unterschiedlicher Wertigkeit verlost (1. Preis: 4-tägige Reise nach Hamburg mit Übernachtungen in einem 4-Sterne-Hotel, Halbpension und Musicalbesuch; 2. Preis / 3. Preis: ein Hotelgutschein für zwei Übernachtungen / eine Übernachtung im gleichen 4-Sterne-Hotel). Die Verlosung ist so organisiert, dass in eine Schachtel 20 Lose gegeben werden, die mit $1, 2, .., 20$ nummeriert sind. Jeder Mitarbeiter ist durch genau eine der Nummern repräsentiert. Aus der Schachtel werden dann nacheinander 3 Lose gezogen – zuerst der Hauptgewinn, dann der 2. Preis und am Ende der 3. Preis. Um auszuschließen, dass jemand mehr als einen Preis gewinnt, wird eine gezogene Nummer vor dem Ziehen der nächsten Nummer nicht zurückgelegt.

Wieviele Möglichkeiten gibt es, die drei Preise innerhalb der 20 Personen umfassenden Belegschaft bei Anwendung dieses Losverfahrens zu verteilen? Tragen Sie Ihr (ganzzahliges) Ergebnis rechtsbündig in das Antwortfeld ein. Übertragen Sie Ihr Ergebnis rechtzeitig vor Ende der Klausur auf den Markierungsbogen.

(numerisch)

Aufgabe 44 (Lotto in Litauen) (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 Litauen wird „6 aus 30“ gespielt, also nur 30 Kugeln verwendet, von denen wie in Deutschland 6 gezogen werden. Bezeichne X die Anzahl der Richtigen bei der litauischen 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 45 (Wahrscheinlichkeiten bei normalverteiltem Merkmal) (3 Punkte)

Für Personen einer bestimmten Grundgesamtheit soll die Ausprägung der Variablen „Intelligenz X“ bestimmt werden. Dabei wird angenommen, dass X sich als eine Zufallsvariable X modellieren lässt, die normalverteilt ist mit Erwartungswert $\mu = 100$ und Varianz $\sigma^2 = 225$, also $\sigma = 15$.

Bestimmen Sie die Wahrscheinlichkeit $P(91 \leq X \leq 109)$ dafür, dass für die Variable X bei einer zufällig aus der betrachteten Grundgesamtheit ausgewählten Person ein Wert gemessen wird, der im Intervall $[91; 109]$ liegt. 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. $0,654$ errechnen, tragen Sie in die letzten sechs Felder $0,6540$ ein. Vergessen Sie nicht, Ihre Antwort rechtzeitig vor dem Ende der Klausur auf den Markierungsbogen zu übertragen.

(numerisch)

Aufgabe 46 (Wahrscheinlichkeiten bei Normalverteilung) (4 Punkte)

Bestimmen Sie einen Wert, der von einer standardnormalverteilten Zufallsvariablen mit Wahrscheinlichkeit $0,2$ nicht überschritten wird. Geben Sie das Ergebnis auf *vier* Stellen nach dem Dezimalkomma genau an. Verwenden Sie für das **Dezimalkomma** wieder ein **eigenes Feld**.

(numerisch)

Aufgabe 47 (lineares Regressionsmodell / KQ-Schätzung) (3 Punkte)

Auf der Basis eines 20 Wertepaare umfassenden Datensatzes $(x_1, y_1), (x_2, y_2), ..., (x_{20}, y_{20})$, für dessen Schwerpunkt $(\overline{x}, \overline{y})$ sich $\overline{x} = 5,24$ und $\overline{y} = 2,28$ errechnet, wurde eine Regressionsgerade nach der Methode der kleinsten Quadrate (KQ-Methode) geschätzt.

Welchen Wert nimmt die KQ-Regressionsgerade $\widehat{y} = \widehat{\alpha} + \widehat{\beta}x$ an der Stelle $x = 1,5$ an? Gehen Sie davon aus, dass die KQ-Methode für die Steigung der Regressionsgeraden den Wert $\widehat{\beta} = -0,25$ lieferte. Tragen Sie das Ergebnis auf *drei 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)

Aufgabe 48 (KQ-Schätzung / Residuen) (3 Punkte)

Auf der Basis eines Datensatzes $(x_1, y_1), (x_2, y_2), ..., (x_{10}, y_{10})$ seien für die Koeffizienten α und β des linearen Regressionsmodells nach der Methode der kleinsten Quadrate die Schätzungen $\widehat{\beta} = 0,58$ und $\widehat{\alpha} = 0,35$ bestimmt worden. Der Datenpunkt (x_1, y_1) sei durch die Koordinaten $x_1 = 2,40$ und $y_1 = 1,96$ definiert.

Berechnen Sie das Residuum $\widehat{u}_1$, also die Abweichung zwischen dem Ordinatenwert y_1 des Datenpunktes (x_1, y_1) und dem Wert der Regressionsgeraden an der Stelle x_1. Tragen Sie das Ergebnis auf *drei 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)
