FernUniversität in Hagen

LOTSE

Aufgabenteil zur Klausur zu

Modul M1 im BA-Studiengang
„Politikwissenschaft, Verwaltungswissenschaft, Soziologie“

sowie zu Modul 3 im BA-Studiengang
„Soziologie“

Termin: **3. September 2013, 14.00 - 18.00 Uhr**

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

Zweitprüfer: **Dr. H.-G. Sonnenberg**

Spätere **Rückfragen** zur Klausur sind stets **an den Erstprüfer** zu richten.

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:

- Bitte lesen Sie diese Hinweise aufmerksam durch, bevor Sie mit der Bearbeitung beginnen. Legen Sie 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**!

- Die Klausur umfasst Multiple-Choice-Aufgaben (Antwort-Auswahl-Verfahren) und einige Aufgaben, bei denen die Antwort eine Zahl ist. Die Klausurdauer beträgt 240 Minuten.

- 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.

- 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 Ihre Eintragungen bewertet wurden. Sie können so Ihre Antworten auch später mit der Musterlösung vergleichen.

- 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.

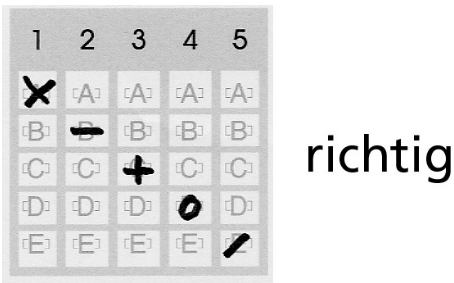
- Für Zwischenrechnungen können Sie das der Formelsammlung angehängte Konzeptpapier verwenden. Zwischenrechnungen gehen nicht in die Bewertung ein, weil nur der LOTSE-Bogen eingesammelt und verarbeitet wird.

- 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.

- Sind die *Aufgaben mit numerischer Antwort* richtig beantwortet, erhalten Sie die volle Punktzahl, ansonsten werden i. d. R. 0 Punkte vergeben.

- Bei den *Multiple-Choice-Aufgaben* sind fünf Aussagen vorgegeben, die – im Extremfall – alle zutreffend oder aber auch alle nicht zutreffend sein können. Zutreffende Aussagen sind auf dem Markierungsbogen mit einem Strich, einem Kreuz oder einem Kreis zu kennzeichnen, falsche Aussagen sind nicht zu markieren. Für jede richtige Antwort (Markierung einer korrekten Aussage oder Nicht-Markierung einer unzutreffenden Aussage) wird 1 Punkt vergeben.

- Die von Ihnen bei den Multiple-Choice-Aufgaben und den Aufgaben mit numerischer Antwort insgesamt erreichten Punkte (= Rohpunkte), werden am Ende in voll erreichte (d. h. nicht erst nach Rundung nach oben erreichte) ganzzahlige Prozentwerte umgerechnet. Das Ergebnis, im Computerergebnisbogen später als *Prozentpunkte* angesprochen, entspricht der größten ganzen Zahl, die nicht größer ist als das 100-fache des Quotienten $\frac{\text{erreichte Rohpunkte}}{\text{maximal erreichbare Rohpunkte}}$.
- Unter Umständen werden die Rohpunkte noch ergänzt durch *Sonderpunkte* zur Berücksichtigung von Kohortenspezifika. Das so errechnete Ergebnis, im Computerergebnisbogen später als *Klausurpunkte* bezeichnet, entspricht der größten ganzen Zahl, die nicht größer ist als das 100-fache des Quotienten $\frac{\text{erreichte Rohpunkte+Sonderpunkte}}{\text{maximal erreichbare Rohpunkte}}$. Wenn keine Sonderpunkte vergeben werden, sind Prozentpunkte und Klausurpunkte identisch.
- Das angewendete Notenschema, das wieder nach Abschluss der Klausurauswertung veröffentlicht wird, bezieht sich auf Klausurpunkte.
- 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).



- Täuschungen, Täuschungsversuche und andere Verstöße gegen die Prüfungsdisziplin können zum Ausschluss von der Klausur und zur Bewertung mit „nicht ausreichend“ (5,0) führen.
- 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 6 – 8 Wochen in Anspruch. Sehen Sie daher bis Mitte Oktober von Nachfragen zum Klausurergebnis ab.

Viel Erfolg bei der Klausurbearbeitung!

Multiple-Choice-Aufgaben zu Block 1

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 Prämissen $P1$ und $P2$ beide erfüllt sind, ist auch K wahr, d. h. die Konklusion ist dann 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? Bei den Aufgabenteilen B und D geht es jeweils um die Beurteilung der Richtigkeit des letzten Satzes. Eine aus mehreren Teilaussagen bestehende Aussage gilt als richtig, wenn jede Teilaussage zutrifft. Der in den Aufgabenteilen D und E verwendete Begriff „metrische Skala“ ist als Oberbegriff für „Intervallskala“, „Verhältnisskala“ und „Absolutskala“ zu verstehen. (x aus 5)

A) Alle 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 Bildungssabschlusses (Ausprägungen des Merkmals “Bildungsstand“: ohne Abschluss, Hauptschule, mittlere Reife, Fachhochsschulreife, Abitur, akademischer Abschluss). Das Merkmal „Bildungsstand“ ist ordinalskaliert.

C) Das Körpergewicht und die Körpergröße von Personen sind stetige Merkmale.

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 3 (Definitionen, Operationalisierung) (5 Punkte)

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

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

B) Eine Nominaldefinition hat stets einen empirischen Informationsgehalt.

C) Eine Realdefinition umfasst alle Eigenschaften des Definiendums.

D) Eine Realdefinition kann sowohl richtig als auch falsch sein.

E) Wenn man theoretische Konstrukte (z. B. die nicht direkt beobachtbare 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 4 (Messen, Datenerhebung, Stichprobenverfahren) (5 Punkte)

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

A) Ein Gütekriterium für Messungen ist die „Reliabilität“. Diese charakterisiert, inwieweit wirklich das gemessen wird, was gemessen werden soll.

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

C) Bei der Klumpenauswahl repräsentieren die Klumpen die Erhebungseinheiten.

D) Bei einer Untersuchung in der Psychologie zum Spracherwerb bei Kleinkindern werden aus einer Grundgesamtheit von Kindern zwei Gruppen gebildet – eine Versuchsgruppe und eine Kontrollgruppe. Anschließend werden in einer Gruppe – der Versuchsgruppe – Einflussgrößen planmäßig verändert. Wenn die Zuordnung der Kinder zur Versuchs- und zur Kontrollgruppe nicht auf der Basis einer Zufallsauswahl erfolgt, spricht man von einem Quasi-Experiment.

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 (Kenngrößen von Datensätzen) (5 Punkte)

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

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

C) Wenn man bei obigem Datensatz den kleinsten Wert $(3,9)$ der Urliste um $1,8$ 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 eingangs aufgeführten Datensatz den kleinsten Wert streicht, werden sowohl der Mittelwert $\overline{x}$ als auch der Median $\widetilde{x}$ des Datensatzes größer.

E) Wenn man den letzten Wert $(6,5)$ des eingangs aufgeführten Datensatz streicht, bleibt die Spannweite des Datensatzes unverändert.

Aufgabe 6 (Charakterisierung eines Datensatzes) (5 Punkte)

Das Landesamt für Statistik in Baden-Württemberg hat einen aktuellen Datensatz, der für alle Einwohner des Landes die Ausprägung des Merkmals „Konfessionszugehörigkeit“ ausweist. Hierüber soll in der Presse berichtet werden. Für die Presse sind natürlich nicht die Urwerte von Interesse. Sie will die in dem Datensatz enthaltene Information verdichten und zwar entweder anhand von Zahlen (Informationsverdichtung anhand numerischer Werte) oder anhand einer Grafik (Visualisierung der im Datensatz enthaltenen Information).

Welche der folgenden Kenngrößen, Statistiken bzw. Grafiken, die zur Charakterisierung von Datensätzen üblich sind, kommen bei diesem speziellen Datensatz in Betracht?

Sinnvoll einsetzbar sind hier (x aus 5)

A) die absoluten oder relativen Häufigkeiten der einzelnen Merkmalsausprägungen.

B) die Kenngröße „Modalwert“ (oder – wenn es mehrere gibt – die Modalwerte).

C) die Kenngröße „Median“.

D) ein Kreisdiagramm.

E) ein Boxplot.

Anmerkung:

In Betracht kommen nur Kenngrößen, Statistiken bzw. Grafiken, die ausschließlich Rechenoperationen voraussetzen, die bei diesem speziellen Datensatz auch erklärt sind.

Aufgabe 7 (Konzentrationsmessung) (5 Punkte)

Bei einer Aktiengesellschaft verteilt sich der Aktienbesitz im Gesamtwert von 60 Millionen Euro auf nur 4 Aktionäre. Aktionär 1 besitzt mit 10 % des Aktienpakets den kleinsten, Aktionär 2 mit 40 % den größten Anteil. Die beiden Aktionäre 3 und 4 sind mit je 25 % beteiligt. Der Aktienbesitz ist also innerhalb der kleinen Aktionärsgruppe nicht gleichmäßig verteilt.

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 $(0,25; v_1)$, $(0,5; v_2)$, $(0,75; v_3)$ und $(1; 1)$ verbindet – so nimmt diese Kurve an der Stelle $0,5$ einen Wert an, der zwischen $0,34$ und $0,36$ liegt.

B) Der Wert, den man für den Ordinatenwert v_3 errechnet, gibt an, welcher Anteil des gesamten Aktienpakets auf diejenigen 3 Aktionäre entfällt, die die größten Anteile am gesamten Aktienpaket halten.

C) Allgemein gilt, dass der unnormierte Gini-Koeffizient für einen Datensatz des Umfangs $n = 4$ den Wert $0,75$ nicht überschreiten kann.

D) Der normierte Gini-Koeffizient hat bei obigem Datensatz einen Wert G^*, der zwischen $0,31$ und $0,33$ liegt.

E) Wenn Aktionär 1 seinen Anteil an Aktionär 2 verkaufte, Aktionär 1 also aus der kleinen Aktionärsgruppe ausschiede, würde der Wert des dann resultierenden normierten Gini-Koeffizienten G^* unterhalb von $0,26$ liegen.

Aufgabe 8 (Randverteilungen, bedingte Wahrscheinlichkeiten) (5 Punkte)

Ein Automobilhersteller will vor der Markteinführung eines neuen Modells zunächst zwei unterschiedliche Designvarianten D_1 und D_2 potenziellen Kunden vorstellen und beauftragt ein Marktforschungsinstitut damit, die Kundenpräferenzen bezüglich der beiden Varianten zu ermitteln. Das Marktforschungsinstitut befragt insgesamt 147 Personen, von denen sich 84 für die Produktvariante D_1 entschieden. Von den befragten 147 Personen waren 81 Männer. Es präferierten 45 Männer die Designvariante D_2.

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

A) Von den weiblichen Befragten favorisierten mehr als 29 % die Produktvariante D_2.

B) Wählt man aus der Population aller befragten Personen eine Person zufällig aus, so liegt die Wahrscheinlichkeit, dass diese die Variante D_2 bevorzugt, über $0,44$.

C) Wählt man aus der Population aller befragten Personen eine Person zufällig aus, so liegt die Wahrscheinlichkeit, dass eine Frau mit Präferenz für D_2 ausgewählt wird, unterhalb von $0,13$ – in Prozentwerten ausgedrückt also unterhalb von 13 %.

D) Von den befragten Personen mit Präferenz für die Designvariante D_2 waren weniger als 30 % weiblich.

E) Wenn man bei den befragten Frauen und bei den befragten Männern jeweils den Anteil der Personen mit Präferenz für die Designvariante D_1 ermittelt, stellt man fest, dass der Anteil bei den Frauen mehr als doppelt so groß ist wie der Anteil bei den Männern.

Aufgabe 9 (diskrete Zufallsvariablen / Venn-Diagramme) (5 Punkte)

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

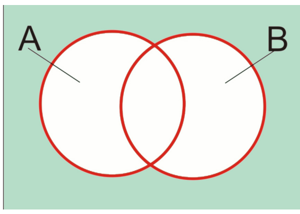
A) Wenn man eine faire Münze, also eine Münze mit gleichen Eintrittswahrscheinlichkeiten für „Kopf“ und „Zahl“, 6-mal wirft und die Anzahl X der Ausgänge mit „Zahl“ feststellt, liegt die Wahrscheinlichkeit dafür, *mindestens* drei Mal „Zahl“ zu erhalten, zwischen $0,62$ und $0,64$.

B) Die Wahrscheinlichkeit dafür, bei dem 6-maligem Münzwurf aus Aufgabenteil A *genau* drei Mal „Zahl“ zu erhalten, ist größer als $0,30$.

C) Beim Würfeln mit zwei fairen Würfeln, liegt die Wahrscheinlichkeit dafür, eine Augensumme zu erhalten, die nicht kleiner als 11 ist, unterhalb von $0,1$.

D) Der Erwartungswert für die Augensumme beim Würfeln mit zwei fairen Würfeln beträgt 7.

E) Nachstehend ist ein Venn-Diagramm wiedergegeben. Dies ist hier ein Rechteck, in dem zwei Ereignisse oder Mengen A und B als Kreise dargestellt 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.



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

Aufgabe 10 (Wahrscheinlichkeits- und Verteilungsfunktion) (5 Punkte)

Es sei eine diskrete Zufallsvariable X mit 10 Ausprägungen gegeben, nämlich den Ausprägungen $x_1 = 0, x_2 = 1, \ldots, x_{10} = 9$. Alle 10 Ausprägungen besitzen dieselbe Eintrittswahrscheinlichkeit.

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

A) Die Wahrscheinlichkeitsfunktion $f(x)$ der diskreten Zufallsvariablen X nimmt für $x = 2$ den Wert $0,1$ an.

B) Die Verteilungsfunktion $F(x)$ von X nimmt für $x = 2$ den Wert $0,2$ an.

C) Die Verteilungsfunktion $F(x)$ von X nimmt für $x = 9$ den Wert 1 an.

D) Für den Erwartungswert $\mu = E(X)$ der Zufallsvariablen X gilt $\mu = 4$.

E) Die Varianz $V(Y)$ der durch $Y = 0,5 \cdot X$ definierten Zufallsvariablen Y ist halb so groß wie die Varianz von X.

Aufgabe 11 (Zusammenhangsmessung, unabhängige Zufallsvariablen) (5 Punkte)

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

A) Hat man eine Vierfeldertafel, die sich auf n Beobachtungen für zwei binäre Merkmale bezieht, so kann man den Wert des χ^2-Koeffizienten errechnen. Der resultierende Wert für dieses Zusammenhangsmaß kann nicht größer als n sein.

B) Wenn man auf der Basis eines Datensatzes $(x_1; y_1)$, $(x_2; y_2)$, ... , $(x_n; y_n)$ für zwei Merkmale X und Y für den Korrelationskoeffizienten r nach Bravais-Pearson den Wert 0 errechnet, beinhaltet dies, dass zwischen den beiden Merkmalen kein Zusammenhang vorliegt.

C) Der Rangkorrelationskoeffizient r_{SP} kann keine negativen Werte annehmen.

D) Bei einer Bank wird die Risikobewertung bei der Vergabe größerer Kredite von zwei unabhängig voneinander tätigen Sachbearbeitern A und B vorgenommen. Die Bewertung erfolgt jeweils anhand einer 10-stufigen Ratingskala, wobei die Punktzahl 10 die beste Bewertung repräsentiert. Die Ergebnisse der Bewertungen für vier Anträge auf Bewilligung solcher Kredite sind nachstehend ausgewiesen.

Kreditantrag i	Sachbearbeiter A Bewertung x_i	Sachbearbeiter B Bewertung y_i
1	4	5
2	7	9
3	9	8
4	8	6

Mit diesen Bewertungen resultiert für den Rangkorrelationskoeffizienten r_{SP} von Spearman ein Wert, der kleiner als $0,5$ ist.

E) Sind X und Y zwei unabhängige Zufallsvariablen mit den Varianzen $V(X)$ resp. $V(Y)$ und bezeichnet $U = X + Y$ die Summe der beiden Zufallsvariablen, so gilt für die Varianz $V(U)$ der neuen Zufallsvariablen die Gleichung $V(U) = V(X) + V(Y)$.

Aufgabe 12 (Stetige Verteilungen) (5 Punkte)

Welche der folgenden Aussagen sind richtig? Beachten Sie, dass eine aus mehreren Teilen bestehende Aussage nur dann als richtig gilt, wenn sie in allen Teilen zutrifft. Bei den Aussagen A und E gilt es jeweils den Wahrheitsgehalt des letzten Satzes zu beurteilen. (x aus 5)

A) Es sei X eine stetige Zufallsvariable mit der Verteilungsfunktion $F(x)$ und es bezeichne $x_{0,95}$ das $0,95$-Quantil der Verteilung. Die Wahrscheinlichkeit dafür, dass X eine Ausprägung x mit $x > x_{0,95}$ annimmt, beträgt $0,05$.

B) Die Dichtefunktion einer standardnormalverteilten Zufallsvariable nimmt an der Stelle 0 den Wert $0,5$ an.

C) Die Wahrscheinlichkeit $P(Z = 0)$ dafür, dass eine standardnormalverteilte Zufallsvariable Z den Wert 0 annimmt, beträgt 0.

D) Die t-Verteilung ist eine stetige Verteilung, deren Dichtefunktion symmetrisch bezüglich des Nullpunkts ist und mit abnehmender Anzahl der Freiheitsgrade flacher verläuft.

E) Es sei X eine χ^2-verteilte Zufallsvariable. Wenn man die Dichtefunktion $f(x)$ grafisch darstellt und auf der x-Achse das $0,05$-Quantil $x_{0,05}$ und das $0,95$-Quantil $x_{0,95}$ der Verteilung markiert, so hat der vom Punkt $x_{0,05}$ bis zum Punkt $x_{0,95}$ gerechnete Flächeninhalt unter der Dichtekurve den Wert $0,90$ (Flächeninhalt zwischen Dichtekurve und x-Achse).

Aufgabe 13 (Vermischte Aussagen zur schließenden Statistik) (5 Punkte)

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

A) In Deutschland wird das Lottospiel „6 aus 49“ angeboten (Lotto ohne Zusatzzahl). Dabei werden 6 Kugeln aus einer Trommel mit 49 Kugeln gezogen. Jeder Lottospieler kann danach die Anzahl X der „Richtigen“ durch Vergleich der gezogenen Zahlen mit den Zahlen auf dem abgegebenen Lottoschein ermitteln. Wenn man das Lottospiel „6 aus 45“ spielte (mit nur 45 Kugeln), hätte der Erwartungswert für die Zufallsvariable Anzahl X der „Richtigen“ einen Wert oberhalb von $0,82$.

B) Wenn eine Zufallsvariable t-verteilt ist, so ist die Wahrscheinlichkeit dafür, dass sie einen positiven Wert annimmt, genauso groß wie die Wahrscheinlichkeit, dass sie einen negativen Wert annimmt.

C) Getestet werden sollen zu einem vorgegebenen Signifikanzniveau die Hypothesen

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

die sich auf den Erwartungswert μ eines normalverteilten Merkmals beziehen. Anhand des aus Stichprobendaten errechneten Wertes einer geeigneten Prüfgröße kommt man zu einer Testentscheidung, die entweder richtig ist (kein Fehler) oder falsch ist (Auftreten eines Fehlers 1. Art oder 2. Art). Für jeden beliebig gewählten Wert von μ gilt, dass bei diesem μ nur einer der beiden genannten Fehler auftreten kann.

D) Der zweiseitige Test aus Aufgabenteil C, der sich auf den Erwartungswert μ eines normalverteilten Merkmals bezieht, kann nur bei bekannter Varianz der Normalverteilung als Gauß-Test durchgeführt werden. Falls die Varianz der Normalverteilung nur in Form einer Schätzung vorliegt, ist der Test als t-Test (t-verteilte Prüfgröße) durchzuführen. Die kritischen Grenzen, die den Annahmebereich vom Ablehnungsbereich trennen, liegen beim t-Test weiter vom Nullpunkt entfernt als beim Gauß-Test.

E) Der Erwartungswert einer t-verteilten Zufallsvariablen stimmt stets mit dem Erwartungswert der Standardnormalverteilung überein.

Aufgabe 14 (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 $x_1, x_2, \ldots, x_n$ schätzen. Letztere lassen sich als Realisationen unabhängiger Zufallsvariablen $X_1, X_2, \ldots, X_n$ interpretieren (auch Stichprobenvariablen genannt). Aus den n Stichprobenvariablen lässt sich der Stichprobenmittelwert $\overline{X}$ bilden.

Welche der folgenden Aussagen sind richtig? Bei den Aussagen C und 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) Wenn man die quadrierten Abweichungen $(X_1 - \overline{X})^2, (X_2 - \overline{X})^2, \ldots, (X_n - \overline{X})^2$ aufsummiert und die resultierende Summe durch n dividiert, also den Mittelwert aus den quadrierten Abweichungen bildet, hat man eine unverzerrte Schätzung für die Varianz σ^2 des Merkmals X.

C) 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.

D) Der Abstand der Intervallgrenzen eines Konfidenzintervalls für μ, d. h. die Länge des Konfidenzintervalls, nimmt ab, wenn die Irrtumswahrscheinlichkeit α reduziert wird.

E) Schätzt man μ anhand eines Konfidenzintervalls, so hängt der Abstand der Intervallgrenzen davon ab, wie groß der Stichprobenumfang n ist. Je größer n gewählt wird, desto kürzer wird das Konfidenzintervall.

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

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

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) Wird bei obigem Test für die Prüfgröße Z der Wert $z = -1,58$ ermittelt, ist die Nullhypothese abzulehnen.

B) Die Wahrscheinlichkeit dafür, einen Fehler 1. Art zu begehen, hat im Falle $\mu = \mu_0$ den Wert $0,05$.

C) Wenn $\mu > \mu_0$ gilt, besitzt die Wahrscheinlichkeit dafür, einen Fehler 1. Art zu begehen, einen Wert unterhalb von $0,05$.

D) Wenn man den Stichprobenumfang n erhöht, wird für alle Werte μ, an denen ein Fehler 2. Art auftreten kann, die Eintrittswahrscheinlichkeit für den Fehler 2. Art reduziert.

E) Wenn für den unbekannten Parameter μ die Ungleichung $\mu < \mu_0$ gilt, beinhaltet die Ablehnung von H_0 den Eintritt eines Fehlers 2. Art.

Aufgabe 16 (Korrelationsmessung und Regressionsanalyse) (5 Punkte)

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

A) Für einen zehn Beobachtungspaare $(x_1; y_1)$, $(x_2; y_2)$, ... ,$(x_{10}; y_{10})$ umfassenden Datensatz wurde $\overline{x} = 3$ und $\overline{y} = 5$ errechnet sowie

$$\sum_{i=1}^{10}(x_i - \overline{x})^2 = 4,5; \qquad \sum_{i=1}^{10}(y_i - \overline{y})^2 = 5,8; \qquad \sum_{i=1}^{10}(x_i - \overline{x})(y_i - \overline{y}) = 2,2.$$

Wenn man unterstellt, dass zwischen x_i und y_i $(i = 1, 2, ..., 10)$ 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 $[3,2; 3,4]$.

B) Wenn man das lineare Regressionsmodell $y_i = \alpha + \beta x_i + u_i$ für obigen Datensatz $(x_i; y_i)$ heranzieht $(i = 1, 2, ..., 10)$, kann man die Regressionskoeffizienten nach der Methode der kleinsten Quadrate schätzen. Erhielte man dabei für β z. B. den Schätzwert $\widehat{\beta} = 0,525$ (fiktiver Wert), beinhaltete dieses Ergebnis, dass bei einer Erhöhung von X um eine Einheit bei Zugrundelegung des Modells mit einem Anstieg des Wertes für das Merkmal Y um $0,525$ Einheiten zu rechnen wäre.

C) 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 größer als $0,20$ ist.

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

E) Für eine nach der Methode der kleinsten Quadrate bestimmte Regressionsgerade gilt, dass die Summe der quadrierten Residuen nur dann Null ist, wenn alle Datenpunkte auf der Geraden liegen.

Anmerkung:

Wenn man die in Aufgabenteil A angegebenen Summen jeweils durch den Umfang $n = 10$ des Datensatzes dividiert, erhält man die empirischen Varianzen s_x^2 und s_y^2 bzw. die empirische Kovarianz s_{xy}.

Hinweis:

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

Numerische Aufgaben zu Block 1

Aufgabe 41 (Phi-Koeffizient) (3 Punkte)

Bei einer Klausur mit $n = 48$ Teilnehmern wurden für die Merkmale „Klausurerfolg X“ mit den Ausprägungen a_1 und a_2 (bestanden / nicht bestanden) und „Geschlecht Y“ mit den Ausprägungen b_1 und b_2 (männlich / weiblich) die absoluten Häufigkeiten erfasst. Die Ergebnisse sind in der nachstehenden Kontingenztafel zusammengefasst:

	b_1	b_2
a_1	18	15
a_2	6	9

Berechnen Sie den Phi-Koeffizienten auf der Basis der obigen (2 x 2)-Kontingenztabelle. Runden Sie das Ergebnis auf *drei Stellen nach dem Dezimalkomma* und tragen Sie Ihr Ergebnis 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) $\Phi =$ | | | | | | |

Aufgabe 42 (Kombinatorik) (3 Punkte)

Im Rahmen eines Betriebsjubiläums bei einem kleinen IT-Unternehmen werden unter den 15 Mitarbeitern drei Preise unterschiedlicher Wertigkeit verlost (1. Preis: 4-tägige Reise nach Dresden mit Übernachtungen in einem 4-Sterne-Hotel mit Halbpension und Übernahme der Anreisekosten; 2. Preis: 4-tägige Reise nach Dresden mit Übernachtungen im gleichen 4-Sterne-Hotel mit Frühstück, ohne Übernahme der Anreisekosten; 3. Preis: ein Hotelgutschein für zwei Übernachtungen im gleichen 4-Sterne-Hotel mit Frühstück, ohne Übernahme der Anreisekosten). Die Verlosung ist so organisiert, dass in eine Schachtel 15 Lose gegeben werden, die mit $1, 2, .., 15$ nummeriert sind. Jeder Mitarbeiter ist durch genau eine der Nummern repräsentiert. Aus der Schachtel werden dann nacheinander 3 Lose gezogen – zuerst der 3. Preis, dann der 2. Preis und am Ende der Hauptgewinn. 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 15 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 43 (Wahrscheinlichkeit bei einem Gesellschaftsspiel) (3 Punkte)

Bei dem bekannten Spiel *Mensch ärgere Dich nicht* darf jeder Spieler zu Beginn drei Mal Würfeln. Sobald dabei eine Sechs gewürfelt wird, darf eine Spielfigur starten, also auf den Rundparcours gesetzt werden. Spieler A beginnt. Wie groß ist die Wahrscheinlichkeit P dafür, dass bei Spieler A auch nach dem dritten Wurf mit dem Würfel noch keine Sechs gefallen ist? Setzen Sie voraus, dass der verwendete Würfel fair ist, also gleiche Eintrittswahrscheinlichkeiten für die einzelnen Augenzahlen aufweist.

Tragen Sie Ihr Ergebnis rechtsbündig und auf *vier* 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) $P =$

Anmerkung:

Die obige Aufgabe war in modifizierter Fassung Teil eines umfassenden Mathematiktests, der zahlreichen Testpersonen vorgelegt und danach in der Ausgabe 33/2013 vom 29. Mai 2013 der Wochenzeitschrift *Die Zeit* veröffentlicht wurde (dort als Aufgabe 25).

Aufgabe 44 (Wahrscheinlichkeit bei einem Gewinnspiel) (3 Punkte)

Aus einer Gruppe von 6 Personen, die aus 2 Männern und 4 Frauen besteht, werden im Rahmen eines Gewinnspiels zwei Gewinner ermittelt. Dazu wird jeder Person eine der Zahlen 1, 2, ..., 6 zugeordnet, die jeweilige Zahl auf einem Zettel notiert und die Zettel in identischen Briefumschlägen abgelegt. Nach Durchmischen der Umschläge werden nacheinander und ohne Zurücklegen zwei Umschläge zufällig ausgewählt. Die in den gezogenen Umschlägen enthaltenen Zahlen definieren dann die Gewinner.

Wie groß ist die Wahrscheinlichkeit dafür, dass das Gewinnerpaar aus einer Frau und einem Mann besteht? Tragen Sie Ihr Ergebnis, also einen Wert aus dem Intervall $[0; 1]$, auf *drei* 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 45 (Verteilungsfunktion der Normalverteilung) (3 Punkte)

Es sei eine Zufallsvariable X betrachtet, die normalverteilt ist mit Erwartungswert $\mu = 0,5$ und Varianz $\sigma^2 = 1$. Für welchen Wert x nimmt die Verteilungsfunktion $F(x)$ dieser Normalverteilung den Wert $0,9$ an?

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. für x den Wert 1,7541 errechnen, tragen Sie in die letzten sechs Felder 1,7541 ein. Vergessen Sie nicht, Ihre Antwort rechtzeitig vor dem Ende der Klausur auf den Markierungsbogen zu übertragen.

(numerisch) $x =$ | | | | | | | |

Aufgabe 46 (Normalverteilung) (3 Punkte)

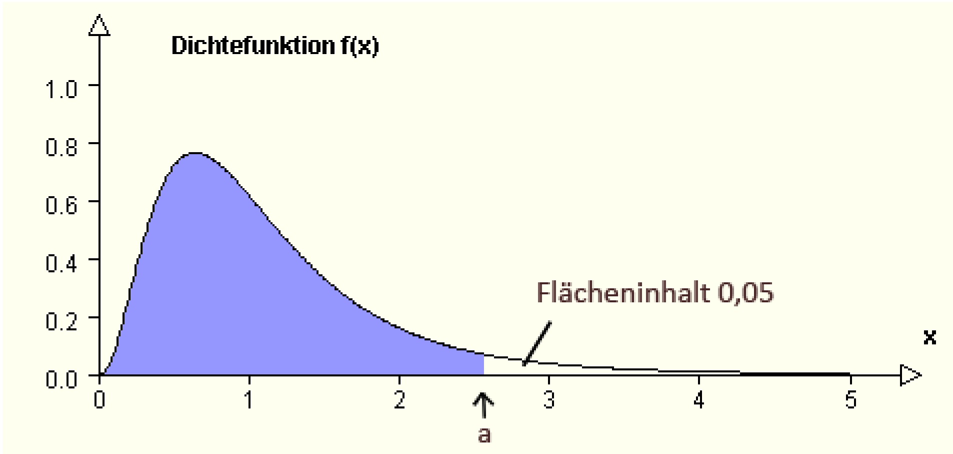
Es sei angenommen, dass sich in einer größeren Population von Personen ein Merkmal X anhand einer Normalverteilung mit Erwartungswert $\mu = 100$ und Standardabweichung $\sigma = 15$ modellieren lässt.

Wie groß ist bei Gültigkeit dieser Annahme die Wahrscheinlichkeit dafür, dass die Ausprägung des Merkmals für eine zufällig ausgewählte Person mindestens 94, aber nicht mehr als 115 beträgt? 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. $0,3864$ errechnen, tragen Sie in die letzten sechs Felder $0,3864$ ein. Vergessen Sie nicht, Ihre Antwort rechtzeitig vor dem Ende der Klausur auf den Markierungsbogen zu übertragen.

(numerisch) $P =$ | | | | | | | |

Aufgabe 47 (F-Verteilung) (3 Punkte)

Die nachstehende Grafik zeigt die Dichtefunktion $f(x)$ einer mit $m = 7$ und $n = 18$ Freiheitsgraden F-verteilten Zufallsvariablen. Auf der horizontalen Achse (Abszissenachse) ist ein Punkt $x = a$ betont. In der Abbildung ist die von der Abszissenachse und der Dichtekurve eingeschlossene Fläche anhand einer bei $x = a$ ansetzenden vertikalen Linie in zwei Teile zerlegt. Die kleinere Teilfläche – in der Grafik ist dies die nicht-markierte Fläche rechts vom Punkt $x = a$ – hat hier den Flächeninhalt $0,05$.



Geben Sie den durch diese Zerlegung definierten Wert a auf 2 Stellen nach dem Dezimalkomma genau an. Tragen Sie Ihr Ergebnis 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) $a =$ | | | | | |

Aufgabe 48 (Gauß-Test) (3 Punkte)

Es sei ein einseitiger Gauß-Test betrachtet, der sich auf die Hypothesen

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

bezieht und mit dem Signifikanzniveau $\alpha = 0,025$ arbeitet. Dabei bezeichnet μ den Erwartungswert eines als normalverteilt spezifizierten Merkmals, dessen Varianz σ^2 als bekannt vorausgesetzt wird. Als Prüfgröße des Tests wird der standardisierte Stichprobenmittelwert $Z = (\overline{X} - \mu_0)/\sigma_{\overline{X}}$ herangezogen (Gauß-Test). Es sei $\mu_0 = 1$ und $\sigma = 0,2$.

Wenn der Parameter μ, auf den sich der Test bezieht, unterhalb von $\mu_0 = 1$ liegt, sollte der Test möglichst zu dem Ergebnis führen, dass die Nullhypothese verworfen wird. Wie groß ist bei dem beschriebenen Test die Wahrscheinlichkeit einer Verwerfung von H_0, wenn der Parameter μ den Wert $\mu = 0,98$ hat und der Test mit einer Stichprobe des Umfangs $n = 9$ arbeitet?

Geben Sie die Wahrscheinlichkeit auf *vier* Stellen nach dem Dezimalkomma genau an. Tragen Sie Ihr Ergebnis 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) $P =$ | | | | | | |