

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

Termin: 14. September 2009, 14.00 - 18.00 Uhr

Prüfer: apl. Prof. Dr. H.-J. Mittag

Aufgabenteil sowie Formelsammlung und Konzeptpapier werden *nicht* eingesammelt – eingesammelt wird nur der maschinenauswertbare Markierungsbogen. Es wird aber empfohlen, dass Sie Ihre Antworten in den Aufgabenteil eintragen und diesen zu Hause aufbewahren. So können Sie später mit der Musterlösung vergleichen.

Hinweise zur Bearbeitung der Klausur:

1. Bitte lesen Sie diese Hinweise vollständig und aufmerksam durch, bevor Sie mit der Bearbeitung beginnen. Legen Sie für die Identitätskontrolle Ihren Personalausweis und die Anmeldebestätigung neben die Klausurunterlagen. Im Klausorraum sind das Rauchen und die Benutzung von Mobiltelefonen nicht gestattet.
2. Die Klausur besteht aus 22 Aufgaben, die in 240 Minuten zu bearbeiten sind. Bitte kontrollieren Sie sofort, ob Sie ein vollständiges Klausurexemplar mit einem 22 Aufgaben umfassenden **Aufgabenteil auf weißem Papier**, eine **Formelsammlung mit Glossar und angehängtem Konzeptpapier auf farbigem Papier** sowie einen **LOTSE-Markierungsbogen** erhalten haben. Trennen Sie die farbigen Blätter zweckmäßigerweise zu Klausurbeginn vom Aufgabenteil ab.
3. Bevor Sie mit der Bearbeitung der Klausuraufgaben beginnen, füllen Sie bitte den Identifikationsteil des Markierungsbogens aus. Tragen Sie dort Ihren Namen, Ihre Matrikelnummer, Ihre Anschrift sowie das Datum ein und unterschreiben Sie.
4. Für die Bewertung der Klausur sind ausschließlich Ihre Markierungen auf dem LOTSE-Markierungsbogen ausschlaggebend. Erfahrungen haben gezeigt, daß Sie spätestens 20 Minuten vor Abgabe der Klausur mit dem Markieren beginnen sollten. Kontrollieren Sie ganz am Schluss noch einmal Ihre Markierungen, bevor Sie den Markierungsbogen abgeben. Es wird dringend empfohlen, die Markierungen auf dem Lotsebeleg mit einem Bleistift und nicht mit einem Kugelschreiber durchzuführen, weil Sie dann während der Klausur mit einem weichen Radiergummi auch noch Antworten ändern können.
5. Bei jeder Aufgabe ist die maximal erreichbare Anzahl der Punkte angegeben. Insgesamt können Sie 100 Punkte erreichen. Mit 50 Punkten haben Sie die Klausur auf jeden Fall bestanden.
6. Die Klausur besteht aus 14 Multiple-Choice-Aufgaben (je 5 Punkte) und 8 numerischen Aufgaben (zwei Aufgaben mit je 3 Punkten und sechs Aufgaben mit je 4 Punkten). Sind die *numerischen Aufgaben* richtig beantwortet, erhalten Sie die volle Punktzahl, ansonsten werden 0 Punkte vergeben. Bei den Multiple-Choice-Aufgaben sind 5 Aussagen vorgegeben, von denen x zutreffend sind mit $x \geq 1$. Zutreffende Aussagen sind zu markieren, falsche Aussagen sind nicht zu markieren. Die Bewertung der *Multiple-Choice-Aufgaben* erfolgt nach folgendem Prinzip:
 - Für jede korrekte Markierung / Nicht-Markierung werden 20 % der pro Aufgabe erzielbaren Punkte vergeben, für jede inkorrekte Markierung / Nicht-Markierung werden 20 % der bei der Aufgabe maximal erreichbaren Punktzahl abgezogen. Es werden aber pro Aufgabe mindestens 0 Punkte vergeben (Ausschluss negativer Salden).
 - Aufgaben, bei denen Sie keine Markierung vornehmen, gelten als nicht bearbeitet und werden mit 0 Punkten bewertet.

Nachstehend ist angedeutet, wie die Lösungen für die Multiple-Choice-Aufgaben aussehen könnten (die Eintragungen sind hier willkürlich vorgenommen worden):

1	2	3	4	5	6	7	8	9	10
A	A	A	A	A	A	A	A	A	A
B	B	B	B	B	B	B	B	B	B
C	C	C	C	C	C	C	C	C	C
D	D	D	D	D	D	D	D	D	D
E	E	E	E	E	E	E	E	E	E

Bewertungsbeispiel für eine Multiple-Choice-Aufgabe mit 5 zu vergebenden Punkten: Sind A und C richtig und es wurden A, C und D markiert, gibt es 3 Punkte (Für A, B, C und E gibt es jeweils 1 Punkt, also insgesamt 4 Punkte, weil die Markierung von A und C korrekt und die Nicht-Markierung von B und E ebenfalls korrekt ist. Die Antwort D ist inkorrekt markiert worden, so dass $4 - 1 = 3$ Punkte übrig bleiben).

- Für Zwischenrechnungen können Sie das der Formelsammlung angehängte Konzeptpapier verwenden. Da nur der Lotse-Markierungsbogen abzugeben ist, können Zwischenrechnungen grundsätzlich nicht in die Bewertung eingehen.
- 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.
- Täuschungen und Täuschungsversuche 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 bis zu 8 Wochen in Anspruch. Sehen Sie daher in diesem Zeitraum bitte von Nachfragen zum Klausurergebnis ab.

Viel Erfolg bei der Klausurbearbeitung!

Aufgabe 1 (Aussagenlogik)

(5 Punkte)

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

a	b	P1: $a \wedge b$	P2: $\neg b$	K: $(a \wedge b) \vee (\neg b)$
w	w			
w	f			
f	w			
f	f			

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

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

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

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

Aufgabe 2 (Definitionen, Operationalisierung)

(5 Punkte)

Welche der folgenden Aussagen, die sich auf Nominal - und Realdefinitionen beziehen, sind zutreffend? (x aus 5)

- A) Nominaldefinitionen sind Worterklärungen, die festlegen, dass ein bestimmter Begriff (das Definiendum) mit einem anderen Begriff (dem Definiens) gleichbedeutend ist.
- B) Eine Nominaldefinition ist entweder richtig oder falsch.
- C) Realdefinitionen sind Worterklärungen, die – im Gegensatz zu Nominaldefinitionen – das Definiendum durch bestimmte Eigenschaften charakterisieren.
- D) Eine Realdefinition umfasst alle Eigenschaften des Definiendums.
- E) Wenn man theoretische Konstrukte – z. B. „Intelligenz“ oder „Berufserfolg“ – empirisch erfassen will, muss man Handlungsanweisungen spezifizieren, die es erlauben, die Variablen zu messen. Man spricht in diesem Zusammenhang von Operationalisierung.

Aufgabe 3 (Messen / Stichprobenverfahren)

(5 Punkte)

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

- A) Ein Gütekriterium für Messungen ist die Reliabilität (Zuverlässigkeit). Diese charakterisiert, inwieweit ein Messinstrument bei wiederholter Messung die gleichen Messwerte liefert.
- B) Ein anderes Gütekriterium, das im Zusammenhang mit der Gewinnung von Daten bedeutsam ist, ist die Validität (Gültigkeit). Diese charakterisiert, inwieweit wirklich das gemessen wird, was gemessen werden soll.
- C) Das Quotenauswahlverfahren ist ein Verfahren der Zufallsauswahl von Stichprobenelementen, bei dem für Teilmengen einer Grundgesamtheit der Anteil der zu ziehenden Stichprobenelemente vorab festgelegt wird.
- D) Die Klumpenauswahl von Stichprobenelementen ist ein zweistufiges Verfahren, bei der man in der ersten Stufe Teilmengen einer Grundgesamtheit (sog. Klumpen) zufällig auswählt.
- E) Bei einer geschichteten Auswahl von Stichprobenelementen wird die Grundgesamtheit zunächst in Teilgesamtheiten (Schichten) zerlegt, die in sich möglichst homogen sind, und danach zieht man aus diesen Teilgesamtheiten Zufallsstichproben.

Aufgabe 4 (Datenerhebung)

(5 Punkte)

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

- A) Under- und Overcoverage sind Probleme, die nur bei systematischen Verfahren der Stichprobenziehung auftreten können.
- B) Wenn man Bevölkerungsumfragen mit offenen Online-Befragungen (freiwillige Teilnehmer) durchführt, ist mit Ergebnissen zu rechnen, die systematisch verzerrt sind, also für die zu untersuchende Grundgesamtheit nicht repräsentativ sind.
- C) Stichprobenbasierte Erhebungen haben gegenüber einer Vollerhebung den Vorteil, dass sie weniger aufwändig und kostengünstiger sind.
- D) Bei Interviews sind Antwortverzerrungen möglich, die z. B. in der Art der Frageformulierung begründet sein können.
- E) Das narrative Interview ist eine qualitative Methode der empirischen Sozialforschung.

Aufgabe 5 (Merkmalsklassifikationen)

(5 Punkte)

In etlichen Städten werden Mietspiegel erstellt, die für Mieter und Vermieter eine Übersicht für den lokalen Wohnungsmarkt liefern. In solchen Mietspiegeln werden u. a. folgende Merkmale erfasst:

- 1 Größe der Wohnung in m^2
- 2 Baujahr des Gebäudes
- 3 Miethöhe für die Wohnung einschließlich Nebenkosten in Euro
- 4 Wohnlage (fünf Kategorien von „Toplage = 1“ bis „Problemlage = 5“)
- 5 Art der Heizung (Gas, Elektro, Öl, Solar)
- 6 Anzahl der Zimmer
- 7 Art des Mietvertrags (Zeitvertrag oder unbefristeter Vertrag) .

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

- A) Die Merkmale 1 – 3 sind metrisch skaliert.
- B) Die Merkmale 4 und 5 sind ordinalskaliert.
- C) Merkmal 7 ist nominalskaliert.
- D) Bei einem nominalskalierten Merkmal kann man keine Differenzen aus den Merkmalsausprägungen bilden, wohl aber eine Rangordnung herstellen.
- E) Die Merkmale 1 und 6 sind Beispiele für diskrete Merkmale.

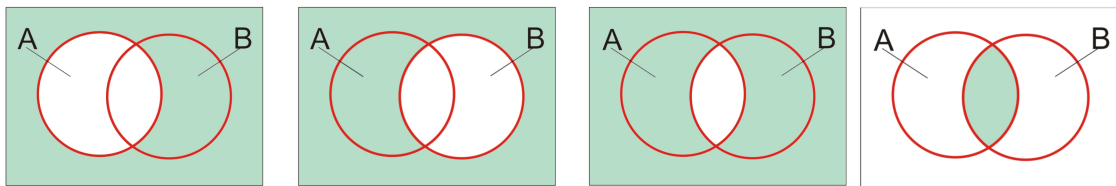
Hinweis: Der Begriff „metrische Skala“ wird als Oberbegriff für „Intervallskala“, „Verhältnisskala“ und „Absolutskala“ verwendet.

Aufgabe 6 (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 \bar{A} und \bar{B} die Komplementär-mengen 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.



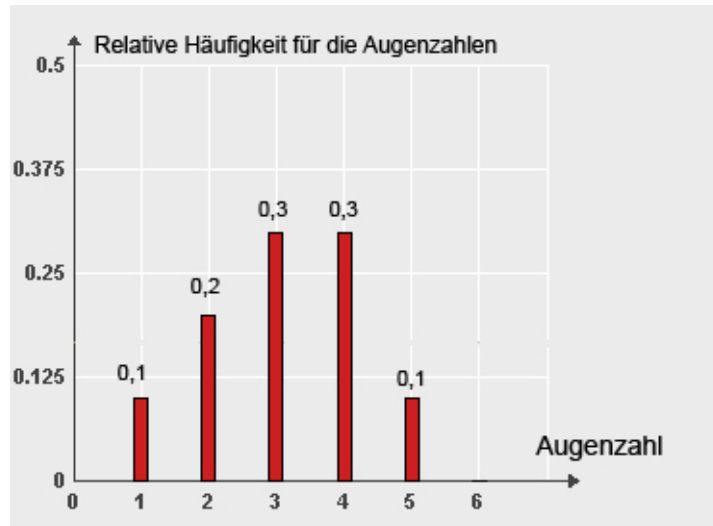
Welche der folgenden Aussagen, von denen sich einige auf die obigen Diagramme beziehen, sind richtig? (x aus 5)

- A) Das erste und zweite Venn-Diagramm veranschaulicht anhand der dunkler gefärbten Flächen die Komplementärmenge \bar{A} von A (erstes Diagramm) bzw. die Komplementärmenge \bar{B} von B (zweites Diagramm).
- B) Im dritten Venn-Diagramm ist durch die dunkler gefärbte Fläche die Schnittmenge der Komplementär-mengen von A und B dargestellt, also $\bar{A} \cap \bar{B}$.
- C) Die dunkel gefärbte Fläche im vierten Venn-Diagramm stellt die Schnittmenge von A und B dar, also $A \cap B$.
- D) Die Vereinigung der beiden Schnittmengen $A \cap B$ und $A \cap \bar{B}$ liefert A .
- E) Die Wahrscheinlichkeiten für ein Ereignis A , das sich aus zwei disjunkten Ereignissen zusammensetzt, ergibt sich als Summe der Wahrscheinlichkeiten der beiden disjunkten Ereignisse.

Aufgabe 7 (Häufigkeitsverteilungen)

(5 Punkte)

Die nachstehende Abbildung zeigt anhand eines Säulendiagramms die relative Häufigkeitsverteilung für einen 10 Werte umfassenden Datensatz, der durch ein Würfelexperiment zustande kam (10-maliges Würfeln mit einem Würfel). Die relativen Häufigkeiten sind im Säulendiagramm auch numerisch ausgewiesen.



- Welche der folgenden Aussagen sind richtig? (x aus 5)
- A) Bei den 10 Würfeln trat 4-mal eine Augenzahl auf, die größer als 3 war.
 - B) Median und Mittelwert des Datensatzes stimmen hier überein.
 - C) Die Spannweite des durch die Grafik definierten Datensatzes hat den Wert 5.
 - D) Wenn noch einmal gewürfelt wird und dabei die Augenzahl 6 erzielt wird, bleibt der Wert des Medians unverändert.
 - E) Wenn man nicht 10-mal, sondern n -mal würfeln (n groß) und die relativen Häufigkeiten für die Augenzahlen anhand eines Säulendiagramms darstellte, erhielte man eine Häufigkeitsverteilung, die sich mit zunehmendem n tendenziell der Wahrscheinlichkeitsfunktion der diskreten Gleichverteilung mit Parameter $p = \frac{1}{6}$ annähert.

Aufgabe 8 (Weitere empirische Verteilungen)

(5 Punkte)

Gegeben sei der folgende Datensatz mit 20 Werten für ein Merkmal X :

2,8 2,9 3,3 3,9 4,2 4,2 4,6 5,3 5,4 5,8
6,0 6,0 6,4 6,5 6,6 6,6 6,7 6,9 6,9 7,0.

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

- A) Der Datensatz hat einen eindeutig bestimmten Modalwert.
- B) Die Bestimmung des Modalwertes für einen Datensatz für ein Merkmal X setzt voraus, dass das Merkmal zumindest ordinalskaliert ist.
- C) Wenn der obige Datensatz um 100 Werte ergänzt würde, die alle positiv sind und unter 10,0 liegen, könnte man die Daten zu Klassen zusammenfassen (z. B. Gruppierung aller Werte von 0 bis ausschließlich 1, von 1 bis ausschließlich 2, usw.) und die relativen oder absoluten Klassenbesetzungshäufigkeiten anhand eines Histogramms darstellen.
- D) Ein Nachteil von Histogrammen ist darin zu sehen, dass der optische Eindruck, den ein Histogramm erzeugt, von der gewählten Klassenbreite abhängt.
- E) Wenn man die Daten eines größeren Datensatzes gruppiert und zwecks Erhöhung der Übersichtlichkeit nur die gruppierten Daten präsentiert, nimmt man einen Informationsverlust gegenüber dem originären Datensatz in Kauf.

Aufgabe 9 (Korrelationsmessung, lineares Regressionsmodell) (5 Punkte)

In der nachstehende Tabelle sind für zwei Merkmale X und Y Beobachtungsdaten $(x_i; y_i)$ wiedergegeben ($i = 1, 2, \dots, 8$).

i	x_i	y_i
1	1,9	3,0
2	2,7	2,5
3	3,1	4,5
4	4,0	3,5
5	3,9	4,0
6	3,4	3,0
7	2,9	4,0
8	2,1	3,5

Aus diesen Daten errechnet man $\bar{x} = 3,0$ sowie $\bar{y} = 3,5$. Für die empirischen Varianzen s_x^2 und s_y^2 erhält man bei Berücksichtigung von vier Dezimalstellen die Werte $s_x^2 = 0,5175$ resp. $s_y^2 = 0,3750$ und für die empirische Kovarianz $s_{xy} = 0,1438$. Diese fünf Werte können hier ungeprüft übernommen werden.









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

- A) Für den obigen Datensatz errechnet man für den Korrelationskoeffizienten nach Bravais-Pearson einen Wert zwischen 0,25 und 0,30 (schwach positive Korrelation).
- B) Wenn man für obige Daten $(x_i; y_i)$ das lineare Regressionsmodell $y_i = \alpha + \beta x_i + u_i$ heranzieht ($i = 1, 2, \dots, 8$), kann man die Regressionskoeffizienten nach der Methode der kleinsten Quadrate schätzen. Erhält man dabei für β den Schätzwert $\hat{\beta} = 0,281$, beinhaltet dieses Ergebnis, dass bei einer Erhöhung von X um eine Einheit mit einer Änderung des Wertes für das Merkmal Y um 0,281 Einheiten zu rechnen ist.
- C) Ist $\hat{\beta} = 0,281$ der Wert der Kleinst-Quadrat-Schätzung für β , resultiert für die Kleinst-Quadrat-Schätzung von α hier ein Wert $\hat{\alpha}$, der zwischen 2,4 und 2,5 liegt.
- D) Der nach der Kleinst-Quadrat-Methode errechnete Wert $\hat{\alpha}$ kennzeichnet den Schnittpunkt der Regressionsgeraden mit der y -Achse.
- E) Die Güte der Anpassung der mit der Kleinst-Quadrat-Methode erhaltenen Regressionsgeraden an den Datensatz lässt sich anhand des Bestimmtheitsmaßes bewerten. Wenn das Bestimmtheitsmaß z. B. den Wert 0,25 annimmt, bedeutet dies, dass 25 % der Gesamtvariation des Datensatzes durch das verwendete Regressionsmodell erklärt wird.

Aufgabe 10 (Kontingenztafeln; Randverteilungen)

(5 Punkte)

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

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

Welche der folgenden Aussagen sind richtig?

(x aus 5)

- A) Der Tabelle entnimmt man, dass von den 931 Befragten 14 Männer und 23 Frauen die FDP präferierten.
- B) Aus der Tabelle ergibt sich ferner, dass von den 931 Befragten 486 Frauen waren.
- C) Wenn man die bedingten relativen Häufigkeiten für das Merkmal „Parteipräferenz X “ unter der Bedingung $Y = b_2$ (Geschlecht Y hat die Ausprägung „weiblich“) berechnen will, benötigt man die Randverteilung des Merkmals X .
- D) Aus den Werten in der obigen Tabelle (Kontingenztafel mit Randverteilungen) lässt sich u. a. ableiten, dass von den an der Befragung beteiligten Personen, die weiblichen Geschlechts waren (Bedingung $Y = b_2$), ca. 32,9 % die SPD präferierten.
- E) Wenn man die in der obigen Tabelle wiedergegebenen absoluten Häufigkeiten und Randhäufigkeiten durch n dividiert, hier also durch 931, resultieren relative Häufigkeiten bzw. relative Randhäufigkeiten.

Aufgabe 11 (Kombinatorik / diskrete Verteilungen)

(5 Punkte)

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

- A) Beim Würfeln mit zwei Würfeln liegt die Wahrscheinlichkeit dafür, eine Augensumme zu erhalten, die größer als 10 ist, zwischen 0,05 und 0,07.
- B) Wenn man eine „faire“ Münze, also eine Münze mit gleichen Eintrittswahrscheinlichkeiten für „Kopf“ und „Zahl“, 4-mal wirft und die Anzahl X der Ausgänge mit „Zahl“ feststellt, ist die Wahrscheinlichkeit dafür *höchstens* zweimal „Zahl“ zu erhalten, größer als 0,65.
- C) Die Wahrscheinlichkeit dafür, bei dem 4-maligem Münzwurf aus Aufgabenteil B *genau* dreimal „Zahl“ zu erhalten, ist 0,25.
- D) Aus einer Lostrommel, in der jedes fünfte Los einen Gewinn repräsentiert, werden nacheinander 3 Lose *mit* Zurücklegen gezogen. Die Wahrscheinlichkeit dafür, nach den 3 Ziehungen keinen Gewinn gezogen zu haben, ist nicht größer als 0,5.
- E) Aus einer Lostrommel mit 200 Losen, von denen 40 Gewinne sind, werden nacheinander 5 Lose *ohne* Zurücklegen gezogen. Der Erwartungswert $\mu = E(X)$ für die Anzahl X der nach den 5 Ziehungen erhaltenen Gewinne hat den Wert 1.

Aufgabe 12 (Punkt- und Intervallschätzungen)

(5 Punkte)

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

- A) Gegeben seien n Stichprobendaten x_1, x_1, \dots, x_1 , die als Ausprägungen unabhängiger Zufallsvariablen X_1, X_2, \dots, X_n interpretiert werden. Der aus den Daten gebildete Stichprobenmittelwert liefert eine unverzerrte (Punkt-) Schätzung für den Erwartungswert μ eines normalverteilten Merkmals.
- B) Die Qualität einer Schätzfunktion hängt allein von der Größe der Verzerrung ab.
- C) Man kann den Erwartungswert μ eines normalverteilten Merkmals auch durch Angabe eines Konfidenzintervalls schätzen. Letzteres ist ein Intervall, das μ mit einer Wahrscheinlichkeit von $1 - \alpha$ enthält, wobei α eine vorzugebende Irrtumswahrscheinlichkeit ist.
- D) Die Grenzen des in Aufgabenteil C genannten Konfidenzintervalls für μ sind bei gegebenem α und gegebenem n fest, hängen also nicht von den Ausprägungen der Stichprobenvariablen X_1, X_2, \dots, X_n ab.
- E) Die Länge des Konfidenzintervalls für μ zum Konfidenzniveau $1 - \alpha$ nimmt ab, wenn der Stichprobenumfang n reduziert wird.

Aufgabe 13 (Stetige Verteilungen)

(5 Punkte)

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

- A) Es sei X eine stetige Zufallsvariable mit der Verteilungsfunktion $F(x)$. 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 0,95.
- B) Es sei X eine χ^2 -verteilte Zufallsvariable. Wenn man die Dichtefunktion $f(x)$ grafisch darstellt und auf der x -Achse das 0,95-Quantil $x_{0,95}$ der Verteilung markiert, so hat der bis zum Punkt $x_{0,95}$ gerechnete Flächeninhalt unter der Dichtekurve den Wert 0,95 (Flächeninhalt zwischen Dichtekurve und x -Achse bis zum 0,95-Quantil).
- C) Wenn X eine t -verteilte Zufallsvariable ist, dann liegen das $x_{0,05}$ -Quantil und das $x_{0,95}$ -Quantil der Verteilung gleich weit vom Nullpunkt entfernt, d. h. ihre Werte unterscheiden sich nur durch das Vorzeichen.
- D) Die t -Verteilung und die Standardnormalverteilung sind stetige Verteilungen mit Dichtefunktionen, die bezüglich des Nullpunkts symmetrisch sind.
- E) Die t -Verteilung lässt sich durch die Standardnormalverteilung approximieren, wobei die Güte der Approximation mit zunehmender Anzahl der Freiheitsgrade der t -Verteilung zunimmt.

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, \dots, X_n aufgefasst (Normalverteilung mit unbekanntem Erwartungswert μ und Varianz σ^2). Getestet werden soll

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

und zwar zum Signifikanzniveau $\alpha = 0,01$. Wenn man die Varianz σ^2 als bekannt voraussetzt, kann man den standardisierten Stichprobenmittelwert $Z = (\bar{X} - \mu_0)/\sigma_{\bar{X}}$ als Prüfgröße für den Test heranziehen (Gauß-Test).

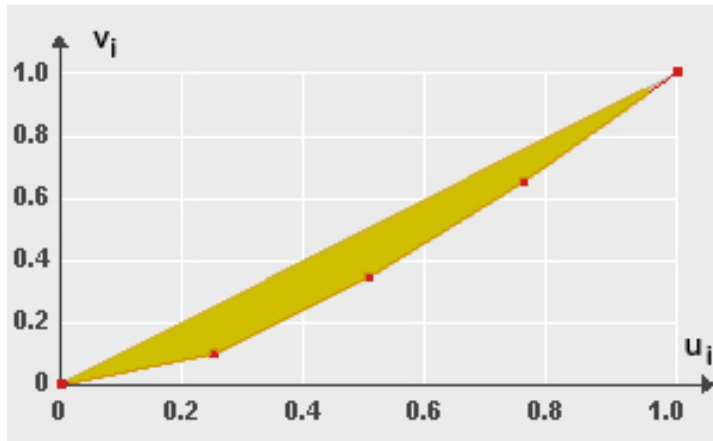
Welche der folgenden Aussagen sind richtig? (x aus 5)

- A) Die Nullhypothese wird bei obigem Test verworfen, wenn die Prüfgröße den Wert 1,96 überschreitet.
- B) Wenn die Prüfgröße im Annahmebereich liegt, kann die Nullhypothese als statistisch „bewiesen“ angesehen werden, in dem Sinne, dass ihre Gültigkeit mit einer maximalen Irrtumswahrscheinlichkeit von α als gesichert angenommen werden kann.
- C) Die fälschliche Verwerfung der Nullhypothese H_0 wird als Fehler 1. Art bezeichnet. Die Wahrscheinlichkeit für den Eintritt eines Fehlers 1. Art beträgt bei dem hier betrachteten Test mit $\alpha = 0,01$ höchstens 0,01.
- D) Wenn man bei obigem Test die Varianz σ^2 resp. die Standardabweichung σ nicht als bekannt voraussetzen kann und eine Schätzung $\hat{\sigma}$ heranzieht (Schätzung von σ durch die korrigierte Stichprobenstandardabweichung), ist die resultierende Prüfgröße t -verteilt mit $n - 1$ Freiheitsgraden. Die Nullhypothese wird nun (bei unveränderter Wahl von $\alpha = 0,01$) verworfen, wenn die Prüfgröße größer als das 0,99-Quantil der t -Verteilung mit $n - 1$ Freiheitsgraden ist.
- E) Man kann für die Testentscheidung (Verwerfung oder Beibehaltung der Nullhypothese) auch den p -Wert heranziehen. In diesem Falle basiert die Testentscheidung auf einem Vergleich des Signifikanzniveaus α mit dem p -Wert.

Aufgabe 41 (Konzentrationsmessung)

(4 Punkte)

Es seien $x_1 = 40$, $x_2 = 100$, $x_3 = 100$ und $x_4 = 160$ die Umsätze von vier konkurrierenden Baumärkten in 2008 (Umsätze jeweils in Millionen Euro). Die nachstehende Abbildung zeigt die auf der Basis dieser Umsatzdaten errechnete Lorenzkurve (Polygonzug). Die Stützpunkte (u_i, v_i) der Lorenzkurve sind in der Tabelle neben der Grafik schon z. T. wiedergegeben – nur die Werte für v_1 , v_2 und v_3 sind noch aus den Daten zu errechnen.



i	u_i	v_i
0	0	0
1	0,25	v_1
2	0,50	v_2
3	0,75	v_3
4	1	1

Bestimmen Sie zunächst die drei fehlenden Ordinatenwerte v_1 , v_2 und v_3 und berechnen Sie dann den Wert des (stets zwischen 0 und 1 liegenden) normierten Gini-Koeffizienten G^* . (Der Inhalt der in der Grafik betonten Fläche entspricht der Hälfte des unnormierten Gini-Koeffizienten G .)

Geben Sie Ihre Lösung mit *drei* Nachkommastellen rechtsbündig in das Antwortfeld ein. Verwenden Sie für das **Dezimalkomma** unbedingt ein **eigenes Feld**. Vergessen Sie nicht, Ihre Antwort rechtzeitig vor Ende der Klausur auf den Markierungsbogen zu übertragen.

(numerisch)

$G^* =$

Aufgabe 42 (Kombinatorik)

(4 Punkte)

Aus einer Gruppe von 9 Personen P_1, P_2, \dots, P_9 werden 3 Personen per Los ausgewählt. Wieviele Möglichkeiten der Zufallsauswahl gibt es hier, wenn die Reihenfolge, in der die 3 Personen bestimmt werden, hier unerheblich sei?

Tragen Sie Ihr Ergebnis rechtsbündig in das Antwortfeld ein. Übertragen Sie Ihr Ergebnis rechtzeitig vor Ende der Klausur auf den Markierungsbogen.

(numerisch)

--	--	--	--	--	--

Aufgabe 43 (χ^2 -Koeffizient)

(4 Punkte)

Das folgende Beispiel ist adaptiert aus TOUTENBURG / SCHOMAKER / WISSMANN (2009, Abschnitt 4.2.5):

Bei einer medizinischen Studie wurde für $n = 72$ Personen erfasst, ob die Beteiligten regelmäßig einen deutlich erhöhten Alkoholkonsum hatten (Überschreitung eines gewissen Schwellenwerts, bezogen auf reinen Alkohol) und ob sie Leberfunktionsstörungen aufwiesen. Es sei X das Merkmal „Alkoholkonsum“ mit den Ausprägungen a_1 (oberhalb des Schwellenwerts) und a_2 (nicht oberhalb des Schwellenwerts) und Y das Merkmal „Leberstatus“ mit den Ausprägungen b_1 (Funktionsstörungen vorhanden) und b_2 (keine Funktionsstörungen nachweisbar).

	b_1	b_2
a_1	12	20
a_2	2	38

Berechnen Sie den χ^2 -Koeffizienten auf der Basis der obigen (2 x 2)-Kontingenztafel. Runden Sie das Ergebnis auf drei Stellen nach dem Dezimalkomma und tragen Sie Ihr Ergebnis rechtsbündig in das Antwortfeld ein. Das **Dezimalkomma** belegt wieder ein **eigenes Feld**. Vergessen Sie nicht, Ihre Antwort rechtzeitig vor dem Ende der Klausur auf den Markierungsbogen zu übertragen.

(numerisch)

$$\chi^2 = \begin{array}{|c|c|c|c|c|c|} \hline & & & & & \\ \hline \end{array}$$

Hinweis: Ergänzen Sie die obige Vierfeldertafel zunächst um die Randverteilungen.

Aufgabe 44 (Rangkorrelationskoeffizient)

(4 Punkte)

Zwei Banken beurteilen unabhängig voneinander und auf der Basis unterschiedlicher interner Bewertungsrichtlinien für vier mittelständische Unternehmen das mit der Vergabe eines Kredits verbundene Risiko. Bei beiden Banken wird die Risikobewertung anhand einer 10-stufigen Ratingskala vorgenommen, wobei die Punktzahl 10 die beste Bewertung repräsentiert. Die Ergebnisse der Bewertungen sind nachstehend ausgewiesen.

Unternehmen i	Bank A Bewertung x_i	Bank B Bewertung y_i
1	3	5
2	7	8
3	9	10
4	8	7

Untersuchen Sie anhand des Rangkorrelationskoeffizienten r_{SP} von Spearman, ob zwischen den Bewertungen der beiden Banken ein Zusammenhang besteht. Tragen Sie Ihr Ergebnis auf eine Stelle 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 45 (Boxplots)**

(3 Punkte)

Gegeben sei erneut der folgende Datensatz mit 20 Werten für ein Merkmal X :

2,8 2,9 3,3 3,9 4,2 4,2 4,6 5,3 5,4 5,8
6,0 6,0 6,4 6,5 6,6 6,6 6,7 6,9 6,9 7,0.

Diesen Datensatz kann man z. B. anhand eines Boxplots visualisieren. Dieser besteht aus einer Box, die dann noch mit den Extremwerten des Datensatzes durch Striche zu verbinden ist. Berechnen Sie die Länge der Box (Boxplot ohne die beiden Striche).

Tragen Sie Ihr Ergebnis auf eine Stelle 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)

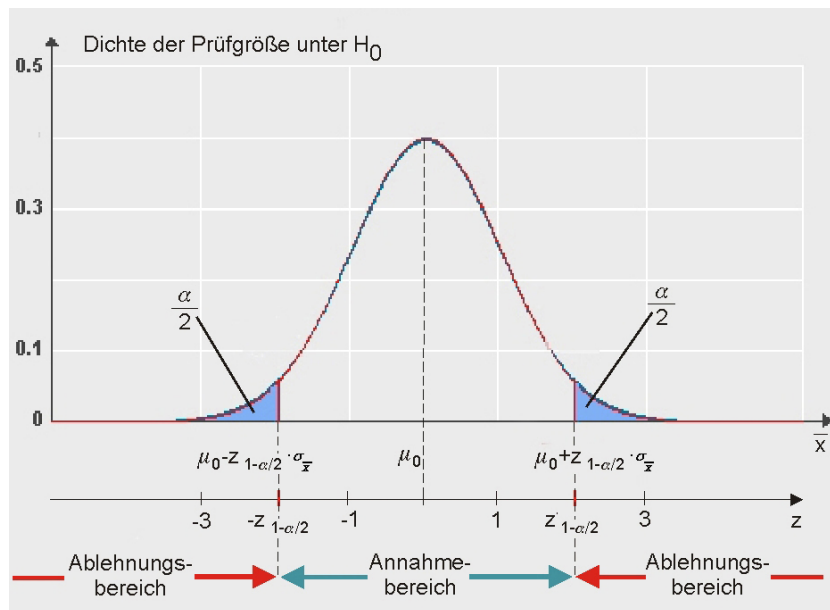
Aufgabe 46 (Gauß-Test)

(3 Punkte)

Wenn man für den Erwartungswert μ eines normalverteilten Merkmals X , dessen Varianz σ^2 als bekannt vorausgesetzt werde, die Hypothesen

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

testet, kann man als Prüfgröße entweder den Stichprobenmittelwert \bar{X} direkt verwenden oder aber zweckmäßigerweise diesen zunächst standardisieren, um dann mit einer standardnormalverteilten Prüfgröße Z arbeiten zu können. Testet man zum Signifikanzniveau α , liegt die Prüfgröße bei zutreffender Nullhypothese H_0 mit Wahrscheinlichkeit $1 - \alpha$ in dem in der nachstehenden Abbildung wiedergegebenen Annahmebereich.



Wie groß darf bei Verwendung von $\alpha = 0,1$ der Betrag $|z|$ des aus den Daten errechneten Werts der standardisierten Prüfgröße Z höchstens sein, damit die Nullhypothese nicht verworfen wird? Geben Sie ihre Antwort auf vier Dezimalstellen nach dem Komma genau an. Verwenden Sie für das **Dezimalkomma** auch hier wieder ein **eigenes Feld**.

(numerisch)

--	--	--	--	--	--

Aufgabe 47 (Wahrscheinlichkeiten bei Normalverteilung)

(4 Punkte)

Das Geburtsgewicht X von Neugeborenen in einer Region sei in guter Näherung durch eine Normalverteilung mit Erwartungswert $\mu = 3050$ und Standardabweichung $\sigma = 125$ zu modellieren. Wie groß ist unter dieser Annahme die Wahrscheinlichkeit dafür, dass ein Neugeborenes ein Geburtsgewicht zwischen 3000 und 3200 Gramm aufweist?

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. „0,6934“ errechnen, tragen Sie in die letzten sechs Felder „0,6934“ ein. Vergessen Sie nicht, Ihre Antwort rechtzeitig vor dem Ende der Klausur auf den Markierungsbogen zu übertragen.

(numerisch)

$$P = \boxed{}\boxed{}\boxed{}\boxed{}\boxed{}\boxed{}$$

Aufgabe 48 (Kleinst-Quadrat-Schätzung)

(4 Punkte)

In einem Ökonometrielehrbuch von VON AUER (2007) findet man ein kleines Illustrationsbeispiel zur Kleinst-Quadrat-Schätzung. Das Beispiel bezieht sich auf $n = 3$ Restaurantbesucher, für die die Merkmale „Rechnungsbetrag X in Euro“ und „gezahltes Trinkgeld Y in Euro“ erfasst wurden. Es sei angenommen, dass $(25; 2)$, $(34; 4)$ und $(31; 3)$ die drei beobachteten Datenpaare (x_i, y_i) sind ($i = 1, 2, 3$) und dass die Höhe des Trinkgelds eine nur durch zufällige Störeinflüsse u_i überlagerte lineare Funktion des Rechnungsbetrags ist. Letzteres beinhaltet, dass man davon ausgeht, dass das einfache lineare Regressionsmodell

$$y_i = \alpha + \beta x_i + u_i$$

hier anwendbar ist. Berechnen Sie auf der Basis dieses Datensatzes nach der Methode der kleinsten Quadrate eine Schätzung $\hat{\beta}$ für den Regressionskoeffizienten β .

Tragen Sie Ihr Ergebnis rechtsbündig und auf drei Stellen nach dem Dezimalkomma genau in das Antwortfeld ein. Sie benötigen also fünf Felder, wenn etwa „1,642“ Ihre Lösung wäre, denn das **Dezimalkomma** muss auch hier wieder ein **eigenes Feld** belegen. Vergessen Sie nicht, Ihre Antwort rechtzeitig vor dem Ende der Klausur auf den Markierungsbogen zu übertragen.

(numerisch)

$$\hat{\beta} = \boxed{}\boxed{}\boxed{}\boxed{}\boxed{}$$