

Aufgabenteil zur Klausur zu  
Modul M1 im BA-Studiengang  
„Politikwissenschaft, Verwaltungswissenschaft, Soziologie“  
sowie zu Modul 3 im BA-Studiengang  
„Soziologie“  
sowie zum Modul M1 im neuen BA-Studiengang  
„Politikwissenschaft, Verwaltungswissenschaft, Soziologie“

Termin: 5. März 2013, 14.00 - 18.00 Uhr

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

Zweitprüfer: Dr. H.-G. Sonnenberg

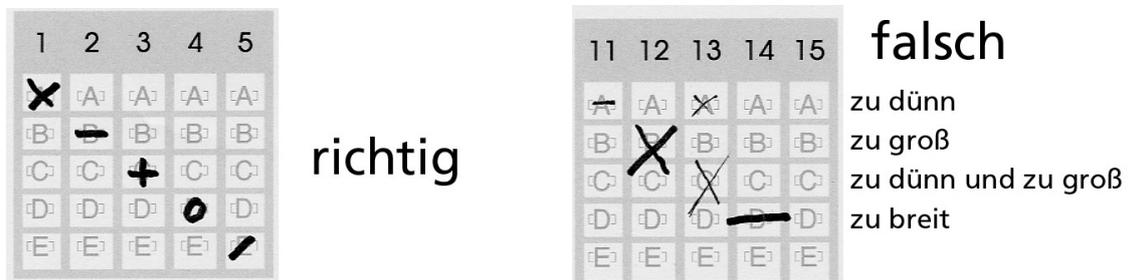
Spätere **Rückfragen** zur Klausur sind stets **an den Erstprüfer** zu richten.

**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  $\frac{\text{erreichte Rohpunkte} + \text{Sonderpunkte}}{\text{maximal erreichbare Rohpunkte}}$ . 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 bis zu 8 Wochen in Anspruch. Sehen Sie daher bis Ende April 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: $a \wedge b$	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, ist nur eine der beiden Prämissen  $P1$  und  $P2$  erfüllt.
- B) Wenn die Aussagen  $a$  und  $b$  beide falsch sind, sind sowohl  $P1$  als auch  $P2$  nicht erfüllt.
- C) Wenn eine der beiden Prämissen  $P1$  und  $P2$  erfüllt ist, ist auch  $K$  wahr, d. h. die Konklusion ist dann korrekt.
- D) Wenn die Aussage  $a$  wahr und  $b$  falsch ist, ist  $P2$  erfüllt, nicht aber  $P1$ .
- E) Wenn die Aussage  $a$  falsch und  $b$  wahr ist, 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 A und C geht es jeweils um die Beurteilung der Richtigkeit des letzten Satzes. Der aus zwei Teilaussagen bestehende letzte Satz in Aufgabenteil C gilt als richtig, wenn jede Teilaussage zutrifft. Der hier und in Aufgabenteil D verwendete Begriff „metrische Skala“ ist als Oberbegriff für „Intervallskala“, „Verhältnisskala“ und „Absolutskala“ zu verstehen. (x aus 5)

- A) Bei einer Einkommenserhebung wird u. a. der Bildungsstand von Arbeitnehmern erfasst und zwar anhand des höchsten erreichten Bildungsabschlusses (Ausprägungen des Merkmals „Bildungsstand“: ohne Schulabschluss, Hauptschule, mittlere Reife, Fachhochschulreife, Abitur, akademischer Abschluss). Das Merkmal „Bildungsstand“ ist ordinalskaliert.
- B) Das in Aufgabenteil A spezifizierte Merkmal „Bildungsstand“ ist diskret.
- C) In der Einkommensteuererklärung muss jeder Arbeitnehmer u. a. die Entfernung zwischen Wohnung und Firma angeben sowie die Zugehörigkeit zu einer Religionsgemeinschaft spezifizieren (Merkmal „Religionszugehörigkeit“ u. a. mit den Ausprägungen „römisch-katholisch“ oder „evangelisch-lutherisch“ oder „keine Zugehörigkeit zu einer Religionsgemeinschaft“). Das Merkmal „Entfernung zum Arbeitsplatz“ ist metrisch skaliert, das Merkmal „Religionszugehörigkeit“ nominalskaliert.
- D) Metrisch skalierte Merkmale können sowohl qualitativ als auch quantitativ sein.
- E) Operationen, die für ordinalskalierte Daten zulässig sind, sind auch für nominalskalierte Daten zulässig.

**Aufgabe 3 (Nominal- und Realdefinitionen)**

(5 Punkte)

Welche der folgenden Aussagen sind zutreffend?

(x aus 5)

- A) Eine Realdefinition beinhaltet eine Aussage über Eigenschaften eines Gegenstands oder Sachverhalts.
- B) Mit einer Nominaldefinition wird ein Gegenstand oder Sachverhalt (das Definiendum) durch einen anderen Gegenstand oder Sachverhalt (das Definiens) erklärt.
- C) Eine Realdefinition umfasst alle Eigenschaften des Definiendums.
- D) Sowohl Nominaldefinitionen als auch Realdefinitionen sind entweder richtig oder falsch.
- E) Eine Nominaldefinition hat keinen empirischen Informationsgehalt.

**Aufgabe 4 (Informationsgewinnung, Datenerhebung)**

(5 Punkte)

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

- A) Bei der Gestaltung von Fragebögen kann es vorkommen, dass eine Frage die Antwort auf eine nachfolgende Frage beeinflusst. Eine solche Ausstrahlungswirkung nennt man Halo-Effekt.
- B) Bei einer Befragung (per Fragebogen oder Interview) dienen Filterfragen dem Zweck, Untergruppen von Befragten zu bilden, die dann jeweils spezielle Fragen zu beantworten haben.
- C) Um theoretische Konstrukte (z. B. „Schulischer Erfolg“ oder „Lebenszufriedenheit“) messen zu können, muss man sie mit beobachtbaren Konstrukten verknüpfen, d. h. es gilt Handlungsanweisungen für die Gewinnung von Daten zu spezifizieren. Diesen Schritt bezeichnet man als Operationalisierung.
- D) Anhand des Random-Route-Verfahrens lassen sich Zufallsstichproben von Personen aus einer Zielpopulation gewinnen. Die Anwendung des Verfahrens setzt das Vorhandensein einer Namens- oder Adressdatei voraus.
- 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 (Messen / Stichprobenverfahren)**

(5 Punkte)

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

- A) Wenn eine Messung dem Gütekriterium „Reliabilität“ genügt, genügt sie auch dem Gütekriterium der „Validität“.
- B) In einer stichprobenbasierten Untersuchung sollen politische Einstellungen von Schülern an Hauptschulen in der Bundesrepublik Deutschland anhand eines standardisierten Interviews erhoben werden. Die Stichprobe wird dadurch erzeugt, dass aus allen Hauptschulen in Deutschland nach einem Zufallsverfahren 120 Schulen ausgewählt werden, in denen dann alle Schüler befragt werden. Bei dieser Vorgehensweise sind die Schüler die Auswahleinheiten.
- C) Die Klumpenauswahl ist eine zweistufige Auswahlprozedur, bei der in der ersten Stufe Teilmengen der Grundgesamtheit – sog. Klumpen – zufällig ausgewählt werden. Aus jedem Klumpen wird dann in der zweiten Verfahrensstufe eine Zufallsstichprobe gezogen.

- D) Bei einer geschichteten Stichprobenauswahl wird eine Grundgesamtheit in Teilmengen zerlegt (sog. Schichten), denen dann jeweils Zufallstichproben entnommen werden. Dabei wird jeder Schicht stets ein fester Prozentsatz von Stichprobenelementen zufällig entnommen.
- E) 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.

**Aufgabe 6 (absolute und relative Häufigkeiten)**

(5 Punkte)

Das Statistische Amt eines EU-Staates gab in einer Pressemitteilung bekannt, dass es in dem betreffenden Land Anfang 2012 insgesamt 36,0 Millionen Menschen im erwerbsfähigen Alter gab. Innerhalb dieser Population belief sich der Anteil der Frauen auf 50,5 %. Ferner wurde bekannt gegeben, dass bei den Frauen 44,5 % und bei den Männern 58,0 % als erwerbstätig gemeldet waren.

Welche der folgenden Aussagen, die sich alle auf die obige Pressemitteilung beziehen, sind richtig: (x aus 5)

- A) Die Anzahl der erwerbstätigen Frauen liegt über 8,2 Millionen.
- B) Die Anzahl der erwerbslosen Männer liegt zwischen 7,4 und 7,5 Millionen.
- C) Der Anteil der erwerbstätigen Männer an der Population der insgesamt 36,0 Millionen Menschen im erwerbsfähigen Alter beträgt mehr als 29,5 %.
- D) Die Gesamtzahl der Erwerbstätigen liegt über 18,3 Millionen.
- E) Wenn man die Anzahl der Erwerbstätigen und die der Erwerbslosen zunächst für die Personen männlichen Geschlechts und danach auch für die Personen weiblichen Geschlechts addiert, erhält man die absolute Randverteilung für das binäre Merkmal „Geschlecht“.

**Aufgabe 7 (Kenngrößen von Datensätzen)**

(5 Punkte)

Gegeben sei der folgende Datensatz für ein stetiges Merkmal  $X$ :

5,3 4,3 6,9 6,4 3,9 4,1 3,0 6,6 4,3 5,9 5,5 6,2.

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 keinen eindeutig bestimmten Modalwert.
- B) Bei obigem Datensatz ist der Median  $\tilde{x}$  kleiner als der Mittelwert  $\bar{x}$ .
- C) Wenn man den Datensatz anhand eines Boxplots visualisiert, liegt der letzte Wert der obigen ungeordneten Urliste (6,2) innerhalb der Box.
- D) Wenn man den ersten Wert (5,3) der oben wiedergegebenen ungeordneten Urliste um 0,2 senkt und den letzten Wert (6,2) um 0,2 erhöht, bleiben sowohl der Median  $\tilde{x}$  als auch der Mittelwert  $\bar{x}$  des Datensatzes unverändert.
- E) Wenn man bei der obigen ungeordneten Urliste den letzten Wert (6,2) streicht, nehmen sowohl der Median  $\tilde{x}$  als auch der Mittelwert  $\bar{x}$  ab.

*Anmerkung:*

Aufgabenteil C bezieht sich auf die Basisvariante eines Boxplots, bei der fünf Charakteristika eines Datensatzes visualisiert werden (Extremwerte und drei Quartile).

**Aufgabe 8 (Konzentrationsmessung)**

(5 Punkte)

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

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

$$x_1 = 20, x_2 = 40, x_3 = 50, x_4 = 65 \text{ und } x_5 = 80.$$

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

- A) Wenn man auf der Basis des obigen Datensatzes die Lorenzkurve zeichnet – also einen Polygonzug, der den Nullpunkt mit den Punkten  $(0, 2; v_1)$ ,  $(0, 4; v_2)$ ,  $(0, 6; v_3)$ ,  $(0, 8; v_4)$  und  $(1; 1)$  verbindet – so nimmt diese Kurve an der Stelle 0,4 einen Wert an, der zwischen 0,25 und 0,26 liegt.
- B) Der Wert, den man für den Ordinatenwert  $v_2$  errechnet, gibt an, welcher Anteil des gesamten monatlichen Taschengeldes aller 5 Schüler auf die beiden Schüler mit dem geringsten monatlichen Taschengeld entfällt.
- C) Der unnormierte Gini-Koeffizient hat bei obigem Datensatz einen Wert  $G$ , der unter 0,22 liegt.
- D) Der normierte Gini-Koeffizient hat bei obigem Datensatz einen Wert  $G^*$ , der zwischen 0,28 und 0,29 liegt.
- E) Wenn man bei dem Datensatz  $x_1 = 20, x_2 = 40, x_3 = 50, x_4 = 65$  und  $x_5 = 80$  alle Werte halbiert, bleibt der Wert des normierten und auch des unnormierten Gini-Koeffizienten unverändert.

**Aufgabe 9 (Kontingenztafeln; Randverteilungen)**

(5 Punkte)

In der nachstehenden Kontingenztafel für absolute Häufigkeiten sind Ergebnisse der Befragung einer Wählerstichprobe wiedergegeben (Daten des ZDF-Politbarometers vom 16. Oktober 2009). Die Häufigkeiten beziehen sich auf  $n = 1021$  Personen, die nach ihrer aktuellen Parteipräferenz befragt wurden. Die Tabelle zeigt die Ergebnisse differenziert nach Geschlecht und jeweils mit Angabe der Randverteilungen für die Merkmale „Parteipräferenz  $X$ “ und „Geschlecht  $Y$ “.

Beim Merkmal  $X$  wurden die Ausprägungen  $a_1 = \text{CDU /CSU}$ ,  $a_2 = \text{SPD}$ ,  $a_3 = \text{FDP}$ ,  $a_4 = \text{Linke}$ ,  $a_5 = \text{Grüne}$ ,  $a_6 = \text{Sonstige}$  unterschieden (bei  $a_1 - a_5$  jeweils mit Ausweis des Parteienlogos). Die beiden Ausprägungen des Merkmals  $Y$  sind mit  $b_1$  (= männlich) und  $b_2$  (= weiblich) codiert.

			Ausprägungen von $Y$		
			$b_1$ (männlich)	$b_2$ (weiblich)	
Ausprägungen von $X$		$a_1$	179	204	383
		$a_2$	100	117	217
		$a_3$	80	59	139
		$a_4$	67	50	117
		$a_5$	54	62	116
	Sonstige	$a_6$	21	28	49
			501	520	1021
			<b>Randverteilung von <math>Y</math></b>		

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

- A) Der Tabelle entnimmt man, dass von den 1021 Befragten der Stichprobe 520 Frauen waren.
- B) Von den befragten Frauen präferierten mehr als ca. 21,0 %, aber weniger als 22,0 % die SPD.
- C) Durch die Randhäufigkeiten der beiden Merkmale „Parteipräferenz  $X$ “ und „Geschlecht  $Y$ “ sind die 12 Häufigkeiten  $h_{ij}$  im Innern der obigen Kontingenztafel eindeutig definiert, d. h. man kann von den beiden Randhäufigkeiten von  $X$  und  $Y$  eindeutig auf die gemeinsamen Häufigkeiten  $h_{ij}$  zurückschließen.
- D) Der Frauenanteil lag bei den Befragten, die sich für die SPD entschieden hatten (Bedingung  $X = a_2$ ), unter 52,0 %.
- E) Ferner lässt sich aus der obigen Tabelle ableiten, dass von den an der Befragung beteiligten Männern (Bedingung  $Y = b_1$ ) mehr als 35,0 %, aber weniger als 36,0 % die CDU/CSU favorisierten.

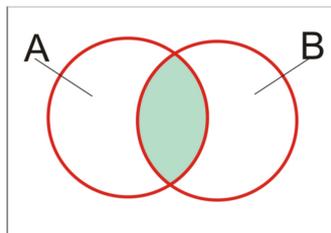
**Aufgabe 10 (Venn-Diagramme, Kombinatorik)**

(5 Punkte)

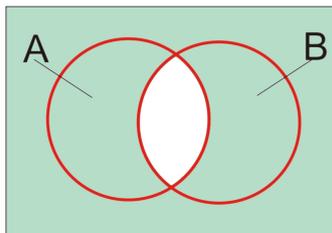
Welche der folgenden Aussagen sind richtig?

(x aus 5)

- A) Wenn man eine „faire“ Münze, also eine Münze mit gleichen Eintrittswahrscheinlichkeiten für „Kopf“ und „Zahl“, 8-mal wirft und die Anzahl  $X$  der Ausgänge mit „Zahl“ feststellt, ist die Wahrscheinlichkeit dafür, *höchstens* dreimal „Zahl“ zu erhalten, größer als 0,38.
- B) Die Wahrscheinlichkeit dafür, bei dem 8-maligem Münzwurf aus Aufgabenteil A *genau* dreimal „Zahl“ zu erhalten, ist kleiner als 0,23.
- C) Mit einem fairen Würfel werde 10-mal in Folge gewürfelt. Bezeichne  $X$  die Anzahl der Ausgänge, bei denen die Augenzahl 1 oder 6 erschien. Der Erwartungswert von  $X$  liegt zwischen 3,0 und 3,2.
- D) Das nachstehende Venn-Diagramm veranschaulicht anhand der dunkler gefärbten Fläche die Schnittmenge  $A \cap B$  zweier Mengen  $A$  und  $B$ .

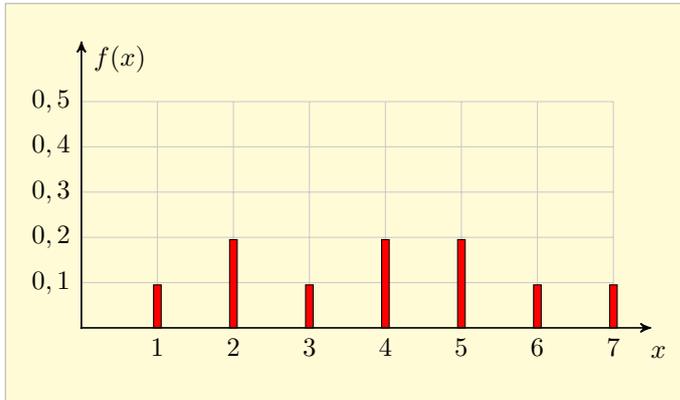


- E) Das nachstehende Venn-Diagramm veranschaulicht anhand der dunkler gefärbten Fläche die Vereinigungsmenge der Komplementärmenge von  $A$  und  $B$ , also  $\overline{A \cap B}$ .



**Aufgabe 11 (Wahrscheinlichkeits- und Verteilungsfunktion)** (5 Punkte)

Die nachstehende Abbildung zeigt die Wahrscheinlichkeitsfunktion  $f(x)$  einer diskreten Zufallsvariablen  $X$ , die sieben Ausprägungen  $x_1 = 1, x_2 = 2, \dots, x_6 = 6, x_7 = 7$  aufweist. Die Ausprägungen  $x_1, x_3, x_6$  und  $x_7$  weisen jeweils die Eintrittswahrscheinlichkeit 0,1 auf, während  $x_2, x_4$  und  $x_5$  je mit Wahrscheinlichkeit 0,2 realisiert werden.



Welche der folgenden Aussagen sind richtig?

(x aus 5)

- A) Die Verteilungsfunktion  $F(x)$  der diskreten Zufallsvariablen  $X$  nimmt für  $x = 3, 0$  den Wert 0,4 an.
- B) Die Verteilungsfunktion  $F(x)$  von  $X$  nimmt für  $x = 3, 5$  den Wert 0,4 an.
- C) Die Verteilungsfunktion  $F(x)$  von  $X$  ist nur bis  $x = 7$  definiert.
- D) Der Erwartungswert  $E(X)$  der Zufallsvariablen  $X$  hat den Wert 4,0.
- E) Wenn man die Zufallsvariable  $X$  gemäß  $Y = X - 1$  transformiert, so ist die Varianz der transformierten Variablen  $Y$  identisch mit der der ursprünglichen Variablen  $X$ .

**Aufgabe 12 (Zusammenhangsmessung)**

(5 Punkte)

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

- A) Die Kovarianz zweier Zufallsvariablen  $X$  und  $Y$  ist ein Zusammenhangsmaß, das nur Werte zwischen  $-1$  und  $+1$  annehmen kann.
- B) Der Korrelationskoeffizient  $r$  nach Bravais-Pearson misst die Stärke eines linearen oder nicht-linearen Zusammenhangs zwischen zwei Merkmalen  $X$  und  $Y$ .
- C) Wenn  $r = 0$  ist, bedeutet dies, dass die Datenpaare  $(x_1, y_1), \dots, (x_n, y_n)$  alle auf einer Geraden mit Steigung 0 liegen.
- D) Der Rangkorrelationskoeffizient  $r_{SP}$  kann nur Werte zwischen 0 und 1 annehmen.
- E) Bei einer Studie wurde bei 28 Jugendlichen jeweils das Verhalten bezüglich des Alkoholkonsums erfasst. Beim Merkmal „Alkohokonsum  $X$ “ wurden drei Ausprägungen unterschieden ( $a_1$  = durchschnittlicher Konsum von bis zu 15 ml reinen Alkohols pro Tag;  $a_2$  = über 15 ml bis einschließlich 30 ml;  $a_3$  = über 30 ml). Zusätzlich wurde noch nach Geschlecht  $Y$  unterschieden ( $b_1$  = männlich,  $b_2$  = weiblich). Die Ergebnisse wurden anhand einer Kontingenztabelle für die Häufigkeiten der 6 Kombinationen von Merkmalsausprägungen für  $X$  und  $Y$  dargestellt. Mit den Werten aus der Kontingenztabelle wurde dann der  $\chi^2$ -Koeffizient berechnet. Der für diesen Koeffizienten berechnete Wert kann nicht größer als 28 sein.

**Aufgabe 13 (Punkt- und Intervallschätzungen)**

(5 Punkte)

Gegeben seien  $n$  Stichprobenwerte, die als Ausprägungen  $x_1, x_2, \dots, x_n$  unabhängiger Zufallsvariablen  $X_1, X_2, \dots, X_n$  interpretiert werden. Die Variablen  $X_1, X_2, \dots, X_n$  seien alle als normalverteilt spezifiziert mit gleichem Erwartungswert  $\mu$  und gleicher Varianz  $\sigma^2$ . Sowohl  $\mu$  als auch  $\sigma^2$  sollen geschätzt werden.

Zur Schätzung der beiden genannten Verteilungsparameter können z. B. der Stichprobenmittelwert  $\bar{X}$  und die Stichprobenvarianz  $S^2$  herangezogen werden. Beide genannten Stichprobenfunktionen leiten sich aus den Zufallsvariablen  $X_1, X_2, \dots, X_n$  ab, wobei  $S^2$  definiert sei als Summe der quadrierten Mittelwertabweichungen  $(X_i - \bar{X})^2$ , dividiert durch  $n$ . Mit  $X_1, X_2, \dots, X_n$  sind auch  $\bar{X}$  und  $S^2$  Zufallsgrößen.

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

- A) Die Stichprobenfunktion  $\bar{X}$  liefert eine unverzerrte Schätzung für den Erwartungswert  $\mu$ .
- B) Die Varianz der Zufallsvariablen  $X_1, X_2, \dots, X_n$  ist kleiner als die Varianz der Stichprobenfunktion  $\bar{X}$ .
- C) Die Stichprobenfunktion  $S^2$  liefert eine unverzerrte Schätzung für die Varianz  $\sigma^2$ .
- D) Man kann den Erwartungswert  $\mu$  der zugrunde gelegten Normalverteilung nicht nur anhand des Stichprobenmittelwerts schätzen (Punktschätzung), sondern auch anhand eines Konfidenzintervalls (Intervallschätzung). Letzteres ist ein Intervall, das stets so groß gewählt wird, dass es den unbekanntem Parameter  $\mu$  enthält.
- E) Schätzt man  $\mu$  anhand eines Konfidenzintervalls, so hängt der Abstand der Intervallgrenzen, d. h. die Länge des Konfidenzintervalls, davon ab, wie groß der Stichprobenumfang  $n$  ist. Je größer  $n$  gewählt wird, desto kürzer wird das Konfidenzintervall.

**Aufgabe 14 (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 Aussage A gilt es 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  $x_{0,05}$  das 0,05-Quantil der Verteilung. Die Wahrscheinlichkeit dafür, dass  $X$  eine Ausprägung  $x$  mit  $x > x_{0,05}$  annimmt, beträgt 0,95.
- B) Die Wahrscheinlichkeit dafür, dass eine standardnormalverteilte oder  $t$ -verteilte Zufallsvariable einen negativen Wert annimmt, beträgt 0,5.
- C) Die  $t$ -Verteilung ist eine stetige Verteilung, deren Varianz mit zunehmender Anzahl der Freiheitsgrade zunimmt.
- D) Sind  $X_1$  und  $X_2$  zwei stetige Zufallsvariablen mit Varianz  $V(X_1)$  resp.  $V(X_2)$  und bezeichnet  $Y = X_1 + X_2$  ihre Summe (also eine aus  $X_1$  und  $X_2$  gebildete neue Zufallsvariable), so gilt stets  $V(Y) = V(X_1) + V(X_2)$ .
- E) Die Wahrscheinlichkeit  $P(X > 2,9)$  dafür, dass eine mit  $m = 5$  und  $n = 15$  Freiheitsgraden  $F$ -verteilte Zufallsvariable  $X$  eine Ausprägung annimmt, die den Wert 2,9 überschreitet, beträgt 0,05.

**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  $X_1, \dots, X_n$  aufgefasst (Normalverteilung mit unbekanntem Erwartungswert  $\mu$  und Varianz  $\sigma^2$ ). Getestet werden soll

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

und zwar zum Signifikanzniveau  $\alpha = 0,01$ . Wenn man die Varianz  $\sigma^2$  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) Wird bei obigem Test für die Prüfgröße ein Wert ermittelt, der den Wert des 0,99-Quantils der Standardnormalverteilung überschreitet, wird die Nullhypothese verworfen.
- B) Wenn man bei obigem Test für die Prüfgröße den Wert  $z = -2,3$  errechnet, wird die Nullhypothese verworfen.
- C) Die Wahrscheinlichkeit dafür, einen Fehler 1. Art zu begehen, beträgt bei obigem Test 0,01.

- 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 jeder Stelle  $\mu$  mit  $\mu \neq \mu_0$  einen Wert, der größer als 0,01 ist.

**Aufgabe 16 (Regressions- und Varianzanalyse)**

(5 Punkte)

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  $(x_1; y_1), (x_2; y_2), \dots, (x_8; y_8)$  umfassenden Datensatz wurde  $\bar{x} = 3$  und  $\bar{y} = 7$  errechnet sowie

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

- Wenn man unterstellt, dass zwischen  $x_i$  und  $y_i$  ( $i = 1, 2, \dots, 8$ ) 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. Für die Steigung  $\hat{\beta}$  der resultierenden Regressionsgerade  $\hat{y} = \hat{\alpha} + \hat{\beta}x$  gilt  $\hat{\beta} > 0,6$ .
- 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 kleiner als 0,12 ist.
- C) Die nach der KQ-Methode bestimmte Regressionsgerade nimmt in  $x = 1$  einen Wert an, der größer als 6,0 ist.
- D) Bei Anwendung der Methode der kleinsten Quadrate wird die Regressionsgerade so bestimmt, dass die Summe der quadrierten Residuen Null ist.
- E) Bei der einfaktoriellen Varianzanalyse wird die unabhängige Variable als diskret modelliert.

**Hinweis:**

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

## Numerische Aufgaben zu Block 1

### Aufgabe 41 (Rangkorrelationskoeffizient)

(3 Punkte)

Zwei Weintester eines Lebensmittel-Discounters bewerten die Qualität von 6 Weißweinen, die neu in das Sortiment aufgenommen werden sollen. Die Bewertung erfolgt anhand eines 10-stufigen Notenschemas (1 = sehr schlecht, 10 = hervorragend). Jeder Tester darf eine Notenstufe nur einmal vergeben. Die Ergebnisse der Bewertungen sind nachstehend ausgewiesen:

Wein $i$	Tester 1 Bewertung $x_i$	Tester 2 Bewertung $y_i$
1	7	7
2	8	9
3	6	3
4	4	5
5	10	8
6	9	10

Untersuchen Sie anhand des Rangkorrelationskoeffizienten  $r_{SP}$  von Spearman, ob zwischen den Bewertungen der beiden Tester ein Zusammenhang besteht.

Tragen Sie Ihr Ergebnis auf *drei Stellen* nach dem Dezimalkomma genau rechtsbündig in das Antwortfeld ein. Wenn Sie also z. B. den Wert 0,4673 errechnen, wäre 0,467 einzutragen. Verwenden Sie für das **Dezimalkomma** ein **eigenes Feld**. Übertragen Sie Ihr Ergebnis rechtzeitig vor Ende der Klausur auf den Markierungsbogen.

(numerisch)

$$r_{SP} = \boxed{\phantom{0}}\boxed{\phantom{0}}\boxed{\phantom{0}}\boxed{\phantom{0}}\boxed{\phantom{0}}\boxed{\phantom{0}}\boxed{\phantom{0}}$$

### Aufgabe 42 (Bestimmung einer Auswahlwahrscheinlichkeit)

(3 Punkte)

An einer Hochschule werden zwei Bachelorstudiengänge angeboten, die hier mit MATH und PHYS abgekürzt seien. Im Studiengang MATH sind 76 Studierende eingeschrieben (darunter 48 Männer), im Studiengang PHYS sind 58 Studierende registriert (darunter 32 Männer). Aus der Gesamtpopulation der Studierenden aus beiden Studiengängen wird eine Person zufällig ausgewählt.

Bestimmen Sie die Wahrscheinlichkeit  $P$  dafür, dass eine im Studiengang MATH eingeschriebene Frau ausgewählt wird. Geben Sie das Ergebnis auf *vier Stellen* nach dem Dezimalkomma genau an. Verwenden Sie für das **Dezimalkomma** ein **eigenes Feld**.

(numerisch)

$$P = \boxed{\phantom{0}}\boxed{\phantom{0}}\boxed{\phantom{0}}\boxed{\phantom{0}}\boxed{\phantom{0}}\boxed{\phantom{0}}\boxed{\phantom{0}}\boxed{\phantom{0}}$$

**Aufgabe 43 (Wahrscheinlichkeiten beim Roulette)**

(3 Punkte)

Beim Roulette gibt es die Ausgänge  $0, 1, \dots, 36$ . Ein Spieler nimmt hintereinander an zwei Spielen teil und setzt bei jedem der beiden Durchgänge darauf, dass das Ergebnis eine ungerade Zahl ist („Impair“, gewonnen wird hier mit jeder der Zahlen  $1, 3, \dots, 35$ ). Berechnen Sie die Wahrscheinlichkeit  $P$  dafür, dass er bei keinem der beiden Spiele gewinnt, also zweimal verliert.

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)

$$P = \begin{array}{|c|c|c|c|c|c|} \hline & & & & & \\ \hline \end{array}$$

**Aufgabe 44 (Varianz einer diskreten Zufallsvariablen)**

(3 Punkte)

Ein Unternehmen erhält eine Lieferung von 100 Tablets. Der Lieferung werden per Zufallsauswahl  $n = 7$  Tablets entnommen (Ziehen ohne Zurücklegen), bei denen dann eine Funktionsprüfung des Betriebssystems vorgenommen wird. Es sei angenommen, dass sich in der Lieferung 3 Tablets befinden, bei denen das Betriebssystem die Funktionsprüfung nicht bestehen würde. Ferner bezeichne  $X$  die Anzahl der Tablets in der Stichprobe, bei denen die Funktionsprüfung negativ verläuft, also ein fehlerhaftes Betriebssystem festgestellt wird.

Offenbar ist  $X$  eine Zufallsvariable. Berechnen Sie deren Varianz  $\sigma^2 = V(X)$ . 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)

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

**Aufgabe 45 (Korrelationsmessung)**

(3 Punkte)

Der nachstehende fiktive Datensatz ist einem Statistiklehrbuch von Hornsteiner (2012) entnommen. Er bezieht sich auf Befragungsergebnisse bei 6 Jugendlichen, bei denen die Ausprägungen der Merkmale „Höhe  $X$  des monatlichen Taschengelds (in Euro)“ und „Monatliche Ausgaben  $Y$  für Besuche von Diskotheken (in Euro)“ erhoben wurden:

$i$	$x_i$	$y_i$
1	25	15
2	35	10
3	55	25
4	100	35
5	35	15
6	50	20

Aus diesen Daten errechnet man für die Mittelwerte der beiden Merkmale die Werte  $\bar{x} = 50$  und  $\bar{y} = 20$ . Berechnen Sie auch den Korrelationskoeffizienten  $r$  nach Bravais-Pearson.

Geben Sie den Wert  $r$  auf *vier* Stellen nach dem Dezimalkomma genau an (eine Stelle vor dem Dezimalkomma). 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)

 $r =$         
**Aufgabe 46 (Normalverteilung)**

(3 Punkte)

Es sei angenommen, dass ein Merkmal  $X$  normalverteilt sei und zwar mit Erwartungswert  $\mu = 5$  und einer Standardabweichung von  $\sigma = 1$ . Wie groß ist dann die Wahrscheinlichkeit  $P = P(4 \leq X \leq 6)$  dafür, dass eine Ausprägung des Merkmals den Wert 4 nicht unter- und den Wert 6 nicht überschreitet, also in das Intervall  $[4; 6]$  fällt?

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

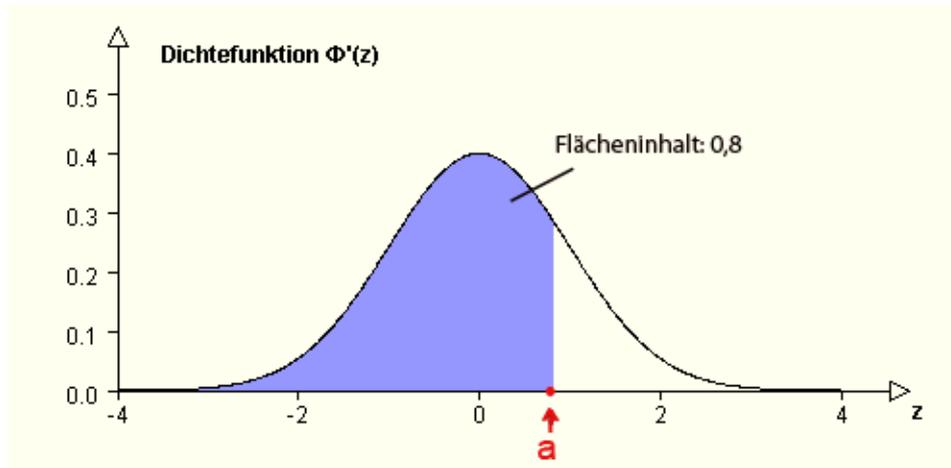
(numerisch)

 $P =$         
**Aufgabe 47 (Standardnormalverteilung)**

(3 Punkte)

Die nachstehende Grafik zeigt die Dichtefunktion  $\Phi'(z)$  einer standardnormalverteilten Zufallsvariablen. Auf der horizontalen Achse (Abszissenachse) ist ein Punkt  $z = a$  betont. Der Wert  $a$  ist dadurch charakterisiert, dass der Inhalt der markierten Fläche unterhalb

der Dichtefunktion  $\Phi'(z)$  bis zum Punkt  $z = a$  den Wert 0,8 besitzt.



Geben Sie den Wert  $a$  auf 4 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 nochmals ein Gauß-Test betrachtet, nun aber ein rechtsseitiger Test

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

zum Signifikanzniveau  $\alpha = 0,025$ . Dabei bezeichnet  $\mu$  wieder den Erwartungswert eines als normalverteilt spezifizierten Merkmals, dessen Varianz  $\sigma^2$  als bekannt vorausgesetzt wird. Als Prüfgröße des Tests wird der standardisierte Stichprobenmittelwert  $Z = (\bar{X} - \mu_0)/\sigma_{\bar{X}}$  herangezogen. Es sei  $\mu_0 = 0$  und  $\sigma = 1$ .

Wenn der Parameter  $\mu$ , auf den sich der Test bezieht, oberhalb von  $\mu_0 = 0$  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  $\mu$  den  $\mu = 0,3$  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)

$P =$