

Wiederholung der Klausur STATISTIK

Name, Vorname: _____

Matrikel-Nr. _____

Die Klausur enthält zwei Typen von Aufgaben:

Teil A besteht aus Fragen mit mehreren vorgegebenen Antwortvorschlägen, von denen mindestens eine Antwort richtig ist und von denen mehrere Antworten richtig sein können. Kreuzen Sie alle richtigen Antworten in den Kästchen unterhalb der Aufgabe an. Sind alle Kreuze richtig, erhalten Sie für die Aufgabe 3 Punkte. Jede Abweichung ergibt 1,5 Punkte Abzug. Es werden keine negativen Punktezahlen vergeben, Sie erhalten also für jede Aufgabe mindestens 0 Punkte. Wenn Sie keine Antwort ankreuzen, gilt die Aufgabe als nicht bearbeitet und Sie erhalten 0 Punkte.

Teil B enthält ausführlich zu lösende Aufgaben. Nur mit der Darstellung der einzelnen Rechenschritte kann die volle Punktzahl erreicht werden.

Zulässige Hilfsmittel: Nicht programmierbarer Taschenrechner, Lehrbuch von Schira, eine handschriftlich von Ihnen selbst beschriebene Seite im DIN A4 Format ("Spickzettel", kann auf beiden Seiten beschrieben sein).

Teil A umfasst 8 Aufgaben und Teil B umfasst 5 Aufgaben. Bitte überprüfen Sie die Vollständigkeit Ihres Exemplars.

Die maximal zu erreichende Punktzahl ist 60, davon können maximal 24 Punkte in Teil A und maximal 36 Punkte in Teil B erreicht werden.

Studierende im Diplomstudiengang VWL bestehen mit mindestens 24 erreichten Punkten die Klausur.

Für Studierende im BSc VWL und alle sonstigen Studierenden werden die in der Klausur und der Zwischenklausur erzielten Punkte addiert. Diese Studierenden bestehen die Prüfung mit mindestens 32 erreichten Punkten.

Auswertung - Teil A

Aufgabe	1	2	3	4	5	6	7	8
Erreichte Punktzahl								

Auswertung - Teil B

Aufgabe	1	2	3	4	5
Erreichte Punktzahl					

Erreichte Gesamtpunktzahl

Teil A (24 Punkte)

1. Bestimmen Sie das Minimum der Funktion $f(z) = \sum_{i=1}^4 |x_i - z|$ für folgende Daten:

i	1	2	3	4
x_i	5	1	-1	1

- A) Das Minimum beträgt $z = 1$.
 B) Das Minimum beträgt $z = 0$.
 C) Das Minimum entspricht dem arithmetischen Mittel der $x_1, \dots, x_4$.
 D) Das Minimum entspricht dem Median der $x_1, \dots, x_4$.

A	B	C	D

2. Die Preise (p) und Mengen (q) für zwei Güter $i = 1, 2$ haben sich von 2000 bis 2008 wie folgt entwickelt:

i	2000		2008	
	q	p	q	p
1	10	1	5	2
2	10	2	20	2

- A) Unter der Annahme, dass 2008 die gleiche Menge wie im Jahr 2000 abgesetzt wurde, nimmt der Preisindex für 2008 zur Basis 2000 den Wert 133 an.
 B) Der Paasche-Preisindex für 2008 zur Basis 2000 beträgt 108.
 C) Der Fisher-Preisindex für 2008 zur Basis 2000 beträgt 121,5.
 D) Der Laspeyres-Preisindex zeigt einen schwächeren Preisanstieg als der Paasche-Index an, da das teurer gewordene Gut 2008 weniger stark nachgefragt wurde.

A	B	C	D

3. Sie beobachten Löhne W für Männer und Frauen in Euro und berechnen $y = \ln(w)$ für jede Beobachtung. Sie erhalten folgende Angaben:

Männerstichprobe

$$\bar{y}_m = 3,2, y_m[0, 25] = 2,7, y_m[0, 5] = 3,1, y_m[0, 75] = 3,7$$

Frauenstichprobe

$$\bar{y}_f = 2,7, y_f[0, 25] = 2,3, y_f[0, 5] = 2,63, y_f[0, 75] = 3,0$$

Die Ausdrücke der eckigen Klammern bezeichnen die entsprechenden empirischen Quantile. Ihre Stichprobe besteht zu 60% aus Männern.

- A) Das geometrische Mittel der Löhne W der Männer beträgt €22,20 (gerundet).
- B) Das geometrische Mittel der Löhne W der Frauen beträgt €14,88 (gerundet).
- C) Das geometrische Mittel der Löhne W aller Arbeitnehmer beträgt €20,09 (gerundet).
- D) Der Median der Löhne W der Frauen beträgt €14,11 (gerundet).

A	B	C	D

4. Ihnen liegen die gleichen Angaben wie in Aufgabe 3 vor.

- A) Das untere Quartil der Löhne W der Männer beträgt €14,88 (gerundet).
- B) Gemessen am Median verdienen Männer 47 Logarithmenprozente mehr als Frauen.
- C) Der Interquartilsabstand für Männer impliziert, dass das obere Quartil der Löhne in Euro um den Faktor 2,52 (gerundet) höher als das untere Quartil ist.
- D) Der relative Unterschied zwischen dem Median und dem unteren Quartil ist für Frauen größer als für Männer.

A	B	C	D

5. Die Investitionen in einem Wirtschaftszweig betrugen 1,45 Mrd. Euro im Jahr 2006 und 0,95 Mrd. Euro im Jahr 2008.
- A) Der Wachstumsfaktor für das Wachstum zwischen 2006 und 2008 beträgt 0,6402.
- B) Die Investitionen schrumpften zwischen 2006 und 2008 um 32,1%
- C) Die jährliche Schrumpfungsrates über den Zeitraum von 2006 bis 2008 in Logarithmen beträgt $-21,14\%$.
- D) Die Antworten A) bis C) sind falsch.

A	B	C	D

6. Die Zufallsvariablen X_1 und X_2 haben die Varianzen $V(X_1) = 2$ und $V(X_2) = 9$. Die Kovarianz sei $Cov(X_1, X_2) = -3$.
- A) $V(X_1 + X_2) < V(X_1) + V(X_2)$
- B) $V(X_1 - X_2) > V(X_1) + V(X_2)$
- C) $Cov(2X_1, X_2) = -12$
- D) Die Antworten A) bis C) sind falsch.

A	B	C	D

7. Folgende Wahrscheinlichkeitsfunktion ist gegeben:

x_i	-1	0	1	2
$f(x_i)$	0,1	0,4	0,4	0,1

- A) $F(0) = 0,5$
- B) $\text{Median}(X) = 0,5$
- C) $E(X) = 0,5$
- D) $V(X) = 0,65$

A	B	C	D

8. Eine Textilfabrik betreibt 80 Webstühle. In einem Jahr fallen im Durchschnitt acht Webstühle aus und müssen repariert werden. Die Zufallsvariable Dauer bezeichne die Dauer zwischen zwei Reparaturzeitpunkten. Unterstellen Sie, dass Dauer exponential verteilt sei und dass die 80 Webstühle statistisch unabhängig arbeiten.

- A) Im Durchschnitt liegen 10 Jahre zwischen zwei Reparaturzeitpunkten, also ist $E(\text{Dauer}) = 10$ Jahre.
- B) $V(\text{Dauer}) = 100 \text{ Jahre}^2$
- C) $P(\text{Dauer} > 1 \text{ Jahr}) = 0,975$
- D) Die Antworten A) bis C) sind falsch.

A	B	C	D

Teil B (36 Punkte)

1. Für das kleine Land Trizonia liegen folgende Angaben zur Anzahl der Einkommensbezieher n_j , zur Gesamtsumme der Einkommen SY_j und zum Einkommensteueraufkommen ST_j nach drei Einkommensintervallen vor:

Einkommen von über ... bis ... Tausend TM	Anzahl der Einkommens- bezieher (n_j)	Summe der Einkommen in Tausend TM (SY_j)	Einkommensteuer- aufkommen in Tausend TM (ST_j)
0 – 5	10	25	0
5 – 20	80	720	72
20 – 100	10	255	48

Das Einkommen und das Einkommensteueraufkommen werden in TM (Trizoniamark) gemessen. Die Einkommensteuer ist progressiv.

Die Angaben in der dritten und vierten Spalte sind die **Merkmalssummen** über alle Steuerpflichtigen in dem betreffenden Einkommensintervall.

- a) Berechnen Sie die relativen Häufigkeiten h_j , die relativen Häufigkeitsdichten $\bar{h}_j$ und die Verteilungsfunktionen $H(\cdot)$ für die drei Variablen ‘Anzahl der Einkommensbezieher’, ‘Einkommen’ und ‘Einkommensteueraufkommen’ nach den drei Einkommensintervallen. Tragen Sie Ihre Ergebnisse in eine Tabelle ein.

(4,5 Punkte)

- b) Zeichnen Sie die Histogramme (Häufigkeitsdichte) für die drei Merkmale ‘Anzahl der Einkommensbezieher’, ‘Einkommen’ und ‘Einkommensteueraufkommen’. Was stellen die drei unterschiedlichen Histogramme inhaltlich dar? Was sind die ökonomischen Gründe für die Unterschiede zwischen den drei Histogrammen?

(4,5 Punkte)

- c) Verwenden Sie die Ergebnisse in a), um die Lorenzkurven für das Einkommen und für das Einkommensteueraufkommen zu bestimmen. Zeichnen Sie die beiden Lorenzkurven. Welche der beiden Lorenzkurven zeigt höhere Ungleichheit an? Wie hängt das Ergebnis mit der Tatsache zusammen, dass die Einkommensteuer progressiv ist?

(3 Punkte)

(12 Punkte)

2. Ihnen liegen in Excel **saisonbereinigte** Quartalsdaten über das reale BIP und den privaten Konsum im Zeitraum 1991 bis 2008 vor. Was versteht man unter dem realen BIP? Wie würden Sie den konjunkturellen Zusammenhang zwischen BIP und privatem Konsum empirisch untersuchen? Erläutern Sie kurz, wie Sie diese empirische Analyse konkret mit Excel umsetzen. Analysieren Sie hierzu die Jahreswachstumsraten relativ zum Vorjahresquartal.

Hinweis: Eine graphische Darstellung und die Diskussion von Kennzahlen können hilfreich sein.

(10 Punkte)

3. An einem Pferderennen nehmen 8 Pferde teil. Wie groß ist jeweils die Wahrscheinlichkeit, dass ein Pferderennen-Wetter ohne Kenntnis der Stärken und Schwächen der einzelnen Pferde die folgenden Ereignisse richtig tippt:

a) siegendes Pferd,

(1 Punkt)

b) die ersten beiden Pferde in der Reihenfolge ihres Einlaufens,

(1 Punkt)

c) die ersten beiden Pferde ohne Berücksichtigung der Reihenfolge ihres Einlaufens?

(1,5 Punkte)

d) Wie ändern sich die Wahrscheinlichkeiten für a) bis c), wenn der Wetter mit Sicherheit weiß, dass ein bestimmtes Pferd Platz 2 einnehmen wird?

(1,5 Punkte)

(5 Punkte)

4. a) Betrachten Sie eine Stichprobe der Variablen X mit n Beobachtungen. Die Stichprobe besteht aus zwei disjunkten Teilgruppen X_1 und X_2 mit jeweils n_1 und n_2 Elementen. Beweisen Sie die Varianzzerlegung.

$$s_{ges}^2 = \sum_{j=1}^2 h_j s_j^2 + \sum_{j=1}^2 h_j (\bar{x}_j - \bar{x})^2$$

wobei $h_j = \frac{n_j}{n}$, s_j^2 die empirische Varianz in Gruppe j und $\bar{x}_j$ das arithmetische Mittel in Gruppe j .

s_{ges}^2 ist die empirische Varianz in der Gesamtstichprobe und $\bar{x}$ das arithmetische Mittel in der Gesamtstichprobe.

(3 Punkte)

- b) Erläutern und interpretieren Sie die Varianzzerlegung am Beispiel der Stichprobe für den Bargeldbestand von n Personen. Die beiden Teilstichproben sind die Frauen und die Männer in der Gesamtstichprobe.

(3 Punkte)

(6 Punkte)

5. Die stetige Zufallsvariable X sei im Bereich $0 < x < 2$ gleichverteilt.

- a) Was ist der Erwartungswert von X ?

(1 Punkt)

- b) Was ist die Varianz von X ?

(1 Punkt)

- c) Berechnen Sie den Erwartungswert der Zufallsvariablen $g(X) = e^X$, d.h. $E(g(X))$.

Unterscheidet er sich vom Wert $g(E(X))$ und wenn ja, warum?

(1 Punkt)

(3 Punkte)