

Name:	Vorname:	Matrikel-Nr.:
-------	----------	---------------

Klausur zur Vorlesung Statistik I und II, WS 2007/08

Aufgabe 1

(40 Punkte)

- (a) Um zu bestimmen, in wie weit die bayerischen Wintersportorte wettbewerbsfähig gegenüber anderen Regionen sind und wo Verbesserungspotenzial besteht, wird für jeden Ort eine statistische Übersicht über die Hotel- und Pensionsbetten, die Zahl der Lifte, die Gesamtlänge an Skipisten und Loipen sowie die Schneesicherheit und Existenz von Beschneiungsanlagen gemacht. Außerdem werden Touristen-Befragungen zur Gesamt-Attraktivität der Orte ausgewertet und nach Alpen- und Mittelgebirgs-Regionen unterschieden.

Geben Sie die statistische Gesamtheit und den Merkmalsträger an. Benennen Sie außerdem mindestens je ein nominal-, ordinal- verhältnis- und absolutskaliertes Merkmal. Geben Sie für jedes dieser Merkmale eine beispielhafte Merkmalsausprägung an.

- (b) Um den Bedarf an seniorengerechten Fahrerassistenzsystemen in den nächsten 20 Jahren abzuschätzen, prognostiziert ein Autohersteller den Anstieg der Fahrleistung der 65-85 Jährigen in diesem Zeitraum mit einer modellgestützten Prognose. Dabei sind folgende Daten bekannt:

Bevölkerungsschicht	Anteil an Gesamtbevölkerung	Führerscheinanteil
45-65 Jahre, weiblich	16%	80 %
45-65 Jahre, männlich	15%	90 %
65-85 Jahre, weiblich	11%	50 %
65-85 Jahre, männlich	9%	80 %

Die Überlebenswahrscheinlichkeit der 45-65 Jahre alten Männer für die nächsten 20 Jahre beträgt $\frac{2}{3}$, die der Frauen 87.5%. Ferner wird angenommen, dass kein über 45 Jähriger den Führerschein in diesem Zeitraum neu macht oder abgibt, dass Zu- und Abzüge vom/in das Ausland keine Rolle spielen und dass sich die jährliche Fahrleistung der Senioren mit Führerschein nicht ändert.

Um wieviel Prozent steigt die gesamte Fahrleistung aller 65-85 Jährigen in Deutschland in den nächsten 20 Jahren? Vergleichen Sie dies mit dem Anstieg der Personenzahl in dieser Altersklasse.

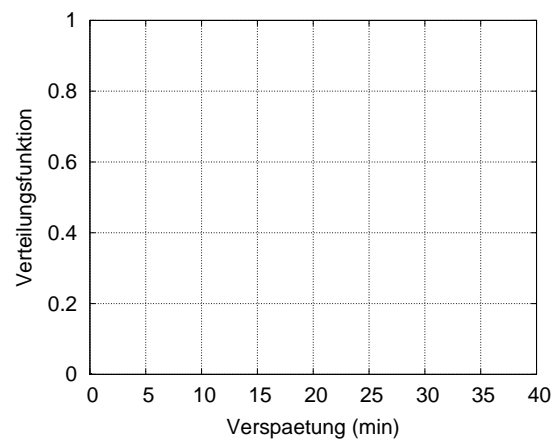
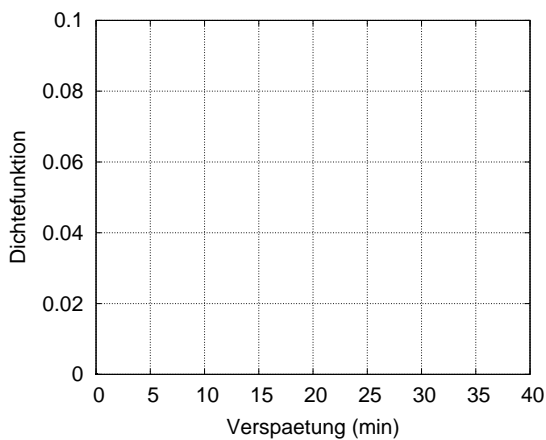
Name:	Vorname:	Matrikel-Nr.:
-------	----------	---------------

Aufgabe 2**(60 Punkte)**

Die Stiftung Warentest ermittelte im letzten Herbst u.a. die Verteilung der Verspätungen bei Fernreisezügen. Das Ergebnis lautet wie folgt:

Verspätung in Minuten	0-1	1-3	3-5	5-10	10-20	20-30	30-60
Anzahl der Züge	10699	3723	2326	2791	1861	905	956

- Wieviel Fernzüge wurden insgesamt erfasst?
- Bestimmen Sie die relativen Häufigkeiten der einzelnen Klassen.
- Bestimmen Sie die relativen Summenhäufigkeiten und die empirische Dichtefunktion für jede Klasse.
- Zeichnen Sie die Dichte- und Verteilungsfunktion in die Diagramme auf dieser Seite ein. (Die Dichte der ersten Klasse geht über den Bereich des Diagramms hinaus; diese brauchen Sie nicht einzuzichnen; dasselbe gilt in beiden Diagrammen für Zeiten oberhalb 40 Minuten).
- Bestimmen Sie das arithmetische Mittel, den Median und den Modus (ohne Feinberechnung). Ist die Schiefe dieser Verteilung positiv oder negativ?
- Bestimmen Sie Varianz, Standardabweichung und die mittlere absolute Abweichung.
- Mit welcher Wahrscheinlichkeit ist ein Fernzug pünktlich (Verspätung unter 3 Minuten)? Welcher Anteil der Fernzüge hat eine Verspätung von mehr als eine viertel Stunde? Um wieviel Minuten sind die 10% der Züge mit der längsten Verspätung mindestens verspätet?



Name:	Vorname:	Matrikel-Nr.:
-------	----------	---------------

Aufgabe 3**(20 Punkte)**

Die Entwicklung der jährlichen Zahl der Verkehrstoten und der Verkehrsleistung in Deutschland in Milliarden Personenkilometern ist wie folgt:

Jahr	Verkehrstote Westdeutschland	Verkehrstote Gesamtdeutschland	Verkehrsleistung Gesamtdeutschland (Mrd Personen-km)
1960	14 000	–	207
1970	18 200	–	430
1980	13 500	–	575
1990	8 000	11 800	715
2000	–	7 700	850
2006	–	5 100	907

- Geben Sie den Index der Zahl der Verkehrstopfer für Westdeutschland von 1960-1990 zur Basis 1960 an!
- Berechnen Sie nun einen Index für Gesamtdeutschland bis 2006. Nehmen Sie dabei vereinfachend an, dass in der Zeit bis 1990 die Wachstumsfaktoren in der ehemaligen DDR und in Westdeutschland gleich waren.
- Berechnen Sie nun den Index der personenbezogenen Risikokennziffer (Verkehrstote pro Milliarden Personenkilometer) zur Basis 1960=1.

Aufgabe 4**(20 Punkte)**

Die Stiftung Warentest hat das Ausmaß der aktuellen Zugverspätungen im Fern- und Regionalverkehr in Deutschland anhand der Analyse aller Verspätungen in 10 deutschen Städten (u.a. Dresden) im Zeitraum 23.9. bis 31.10.2007 abgeschätzt. Die in diesen Zeitraum fallenden 5 Streiktage wurden nicht berücksichtigt.

- Grenzen Sie Grundgesamtheit und Stichprobe jeweils nach den relevanten drei Kriterien ab.
- Benennen Sie die statistischen Einheiten und Gesamtheiten und das betrachtete statistische Merkmal.
- Definieren Sie unter Verwendung von Lage- und Formmaßen den Begriff der Repräsentativität einer Stichprobe. Argumentieren Sie unter Verwendung eines Lagemaßes, warum es sinnvoll war, die Streiktage auszuschließen.
- Was könnte im konkreten Beispiel noch dazu führen, dass die Stichprobe nicht repräsentativ ist?

Name:	Vorname:	Matrikel-Nr.:
-------	----------	---------------

Aufgabe 5**(40 Punkte)**

Ein Student will gleich nach der Klausur zur Freundin nach Nürnberg. Es gibt am 21.2. nach 17:00 h zwei sinnvolle Verbindungen von Dresden Hbf:

- Start 17:56, Dauer 4:21, ohne Umsteigen.
- Start 18:55, Dauer 4:39, einmal Umsteigen in Leipzig (Ankunft 20:04, Abfahrt von Leipzig 20:11).

Die Verbindung mit einmal Umsteigen ist um 18,60 € billiger als die Direktverbindung, allerdings würde man bei Verpassen des Anschlusszuges (nur 7 Minuten Zeitpolster!) für die Nacht in Leipzig gestrandet sein und erst um 8:15 am folgenden Tag in Nürnberg ankommen. Folgende vereinfachende Annahmen werden getroffen: (i) 62% aller Züge sind pünktlich, (ii) die Verspätungszeitdauer der 38% verspäteten Züge ist exponentialverteilt mit einem Erwartungswert von 10 Minuten, (iii) die Anschlusszüge sind pünktlich und warten nicht, kommen allerdings am Bahnsteig gegenüber an, so dass die Übergangszeit vernachlässigbar ist.

- Mit welcher Wahrscheinlichkeit würde man bei der zweiten Verbindung den Zug verpassen?
- Berechnen Sie die Verteilungsfunktionen der tatsächlichen Reisezeiten der beiden Verbindungen (in Minuten). Wie hoch sind jeweils die Erwartungswerte?
- Nun wird angenommen, dass sich der Anschlusszug der zweiten Verbindung auch verspäten kann. Man könnte also auch bei einer Verspätung des ersten Zuges von mehr als 7 Minuten mit Glück den Anschluss noch erreichen. Wie würde man nun bei Annahme der Unabhängigkeit der möglichen Verspätungen die Wahrscheinlichkeit berechnen, den Anschluss zu erreichen? (*Hinweis:* keine Rechnung nötig! Finden und Anpassen der korrekten, in der Formelsammlung stehenden Formel genügt!)

Name:	Vorname:	Matrikel-Nr.:
-------	----------	---------------

Aufgabe 6**(40 Punkte)**

Mittels einer Zufalls-Stichprobe vom Umfang n (gerade und ≥ 6) soll das arithmetische Mittel μ der Bruchlast X von Achsen von ICE-Zügen und ihre Varianz σ^2 bestimmt werden.

(a) Für den Mittelwert kommen folgende Schätzer zum Einsatz:

$$(1) \hat{\mu}_1 = \frac{1}{n} \sum_{i=1}^n x_i,$$

$$(2) \hat{\mu}_2 = \frac{1}{n-1} \sum_{i=1}^n x_i,$$

$$(3) \hat{\mu}_3 = \frac{2}{n} \sum_{i=1}^{n/2} x_i,$$

$$(4) \hat{\mu}_4 = \frac{1}{10} \sum_{i=1}^4 i x_i.$$

Geben Sie für jeden dieser Schätzer jeweils an, ob er

- erwartungstreu,
- konsistent,
- effizient

ist. Berechnen Sie dazu für jeden dieser Schätzer Erwartungswert und Varianz.

(b) Für die Varianz werden folgende Schätzer betrachtet:

$$(1) \hat{\sigma}_1^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2,$$

$$(2) \hat{\sigma}_2^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2,$$

$$(3) \hat{\sigma}_3^2 = \frac{2}{n-2} \sum_{i=1}^{n/2} x_{2i}.$$

Geben Sie auch hier jeweils an, ob sie erwartungstreu und/oder konsistent und/oder effizient sind (Erwartungswerte und Varianzen müssen nicht explizit berechnet werden)

Name:	Vorname:	Matrikel-Nr.:
-------	----------	---------------

Aufgabe 7**(20 Punkte)**

Biologiestudenten gelten ja im Allgemeinen "ökologischer" als Wirtschaftsstudenten. Dies soll am Beispiel Fahrzeugbesitz anhand einer Stichprobe getestet werden. Die Stichprobe ergab:

Fach	Anzahl in Stichprobe	Anzahl Kfz-Besitzer
Biologie	234	87
Wirtschaft	789	401

Bestimmen Sie mit dem Unabhängigkeitstest (Vierfeldertest), ob man bei einer Fehlerwahrscheinlichkeit von 5% die Annahme verwerfen kann, dass der Anteil der Studenten mit Kfz nicht von den beiden Fachrichtungen abhängt. Bei welcher Fehlerwahrscheinlichkeit kann man die Annahme gerade nicht mehr verwerfen?

Tabellen**Quantile $q_{\alpha}^{(m)}$ der χ^2 -Verteilung mit n Freiheitsgraden**

n	$\alpha = 0.9900$	0.9750	0.9500	0.9000	0.8000	0.5000	0.2000	0.1000	0.05000
1.0	6.635	5.024	3.841	2.706	1.642	0.4549	0.06418	0.01579	0.003932
2.0	9.210	7.378	5.991	4.605	3.219	1.386	0.4463	0.2107	0.1026
3.0	11.34	9.348	7.815	6.251	4.642	2.366	1.005	0.5843	0.3518
4.0	13.28	11.14	9.488	7.779	5.989	3.357	1.649	1.064	0.7106
5.0	15.09	12.83	11.07	9.236	7.289	4.351	2.343	1.610	1.145
6.0	16.81	14.45	12.59	10.64	8.558	5.348	3.070	2.204	1.635
7.0	18.48	16.01	14.07	12.02	9.803	6.346	3.822	2.833	2.167
8.0	20.10	17.54	15.51	13.36	11.03	7.344	4.594	3.490	2.733
9.0	21.67	19.03	16.92	14.68	12.24	8.343	5.380	4.168	3.325
10.	23.22	20.49	18.31	15.99	13.44	9.342	6.179	4.865	3.940
15.	30.59	27.49	25.0	22.31	19.31	14.34	10.31	8.547	7.261
20.	37.58	34.18	31.41	28.41	25.04	19.34	14.58	12.44	10.85
30.	50.92	46.99	43.78	40.26	36.25	29.34	23.36	20.60	18.49