TECHNISCHE UNIVERSITÄT DRESDEN

Fakultät Verkehrswissenschaften „Friedrich List“ Professur für Verkehrsökonometrie und -statistik

Statistik II für Verkehrswirtschaftler

Wintersemester 2011/12

Trainingsaufgaben

Inhaltsverzeichnis

Aufgabenübersicht

1 Zeitreihenanalyse

T.1.1	Übernachtungszahlen	a	b	c	d	e	f ......
T.1.2	Fortzüge aus Dresden	a	b				
T.1.3	Automobilbau	a	b				
T.1.4	Fahrgastzahlen	a	b	c	d	e	f ......
T.1.5	Kurort	a	b	c	d	e	f ......
T.1.6	Fahrraddiebstähle	a	b	c	d		

2 Wahrscheinlichkeitsrechnung

T.2.1	Auktion für ein Cabrio	a	b	c	d	e ......
T.2.2	Mobilitätsverhalten	a				
T.2.3	Rechtschreibfehler	a	b			
T.2.4	Zinsen	a				
T.2.5	Druckfehler	a	b	c	d	e ......
T.2.6	UEFA-Cup	a	b	c		
T.2.7	Kundenzufriedenheit	a	b	c	d	
T.2.8	Schadensmeldung	a				
T.2.9	Stau im Pendlerverkehr	a	b	c	d	
T.2.10	Steuerung	a	b	c	d	e ......
T.2.11	Kriminalität	a	b			
T.2.12	Trunkenheit am Steuer	a	b			
T.2.13	Panne	a	b	c		
T.2.14	JobTicket	a	b	c	d	
T.2.15	Autohändler	a	b	c	d	e ......

3 Wahrscheinlichkeitsverteilungen

T.3.1	Zugfahrt	a	b	c	d		
T.3.2	Verkehrsmittelwahl binomialv.	a	b	c	d	e	
T.3.3	Fußgänger	a	b	c	d		
T.3.4	Induktionsschleifen	a	b	c	d	e	f ......
T.3.5	Direktverbindung	a	b	c			
T.3.6	Schwangerschaft	a	b	c			
T.3.7	Straßenübergang mit Ampel	a	b	c			
T.3.8	Bücher und Musikstücke	a	b	c	d	e	

4 Induktive Statistik

T.4.1	Fußgängerampel	a	b	c	d	e		
T.4.2	Billigtickets	a	b	c				
T.4.3	Motorprüfstand	a	b	c	d	e		
T.4.4	Fahrassistenzsysteme	a	b					
T.4.5	Immatrikulation	a	b	c	d	e		
T.4.6	Zugachsen	a	b					
T.4.7	Ökologisch	a						
T.4.8	Wege mit dem ÖPNV	a	b	c	d			
T.4.9	Belastungstest	a	b	c	d			
T.4.10	Modal Split	a	b	c				
T.4.11	Geburtenspitzen	a	b	c	d			
T.4.12	Zugriffshäufigkeit	a	b	c				
T.4.13	Werbeaktion	a	b	c				
T.4.14	Einwohner Dresdens	a	b	c				
T.4.15	Kraftstoffverbrauch	a	b	c	d	e		
T.4.16	Weiterempfehlung	a	b	c	d			
T.4.17	Verkehrsmittelwahl KI	a	b	c	d			
T.4.18	Verkehrsmittelwahl von St.	a	b	c				
T.4.19	Fahrpreissystem	a	b	c	d	e		
T.4.20	Motorprüfstand	a	b	c	d			
T.4.21	Verspätungen	a	b	c				
T.4.22	Taxi	a	b	c	d	e	f	
T.4.23	Tag der Sachsen	a	b	c	d	e	f	g ...

1 Zeitreihenanalyse

Aufgabe T.1.1 – Übernachtungszahlen pro Saison, WS 2009/10

Die Übernachtungszahlen (in 1 000) eines Urlaubsortes in den letzten drei Jahren waren, nach Quartalen aufgegliedert, wie folgt:

	Q1	Q2	Q3	Q4
2007	240	110	170	95
2008	290	120	170	125
2009	325	150	240	130

Mit Hilfe einer additiven Zerlegung (nichtstationärer Fall) soll eine Prognose für das nächste Jahr durchgeführt werden.

(a) Ermitteln Sie die glatte Komponente.

(b) Ermitteln Sie die Saisonkomponente. Falls Sie (a) nicht gelöst haben, rechnen Sie mit folgenden Werten für die glatte Komponente:

	Q1	Q2	Q3	Q4
2007	-	-	169	168
2008	169	173	181	189
2009	201	211	-	-

(c) Geben Sie die Zufallskomponente (unbestimmte Komponente) für das Jahr 2008 an.

(d) Prognostizieren Sie die glatte Komponente für das Jahr 2010 durch Fortschreibung der letzten verfügbaren Anstiegsrate.

(e) Prognostizieren Sie die Übernachtungszahlen für die vier Quartale des Jahres 2010.

(f) Warum macht man einen Fehler, wenn man die Übernachtungszahlen direkt durch Fortschreibung der letzten verfügbaren Anstiegsrate der anfangs gegebenen Übernachtungszahlen prognostiziert? Welcher Wert wäre mit dieser Methode für das vierte Quartal 2010 zu erwarten?

Aufgabe T.1.2 – Fortzüge aus Dresden, WS 2005/06

Für den Fortzug von Dresden in die Alten Bundesländer ist folgende Zeitreihe gegeben:

	Q I	Q II	Q III	Q IV
2001	1 600	1 580	2 420	1 980
2002	1 510	1 420	1 820	1 520
2003	1 370	1 200	1 640	1 390
2004	1 240	1 060	1 590	1 440

(Q - Quartal)

(a) Begründen Sie, warum ein multiplikatives Modell dem Datenmaterial angemessen ist

(b) Führen Sie eine multiplikative Saisonbereinigung durch. Geben Sie die Saisonindexziffern an. Wie wäre nach diesem Modell die Prognose für die vier Quartale 2005?

Aufgabe T.1.3 – Absatzzahlen im Automobilbau

Eine Automobilfirma habe in den letzten 3 Jahre folgende Absatzzahlen (in 1000 PKW):

Jahr	1. Quartal	2. Quartal	3. Quartal	4. Quartal
1999	200	260	210	150
2000	220	270	220	140
2001	210	280	200	160

(a) Berechnen Sie die Saisonfigur.

(b) Steigt im 4. Quartal 2001 die saisonbereinigte Absatzkennziffer gegenüber dem Vorquartal an oder fällt sie?

Aufgabe T.1.4 – Fahrgastaufkommen, WS 2003/04

Eine S-Bahn Linie hatte in den letzten beiden Jahren die in der Tabelle angegebenen täglichen Fahrgastzahlen (in Tausend):

Jahr	Quartal 1	Quartal 2	Quartal 3	Quartal 4
2002	410	680	710	430
2003	440	720	710	460

Zur Abschätzung des zukünftigen Fahrgastaufkommens soll anhand dieser Daten eine Zeitreihenanalyse vorgenommmen werden.

(a) Berechnen Sie den gleitenden Durchschnitt der Ordnung 1 Jahr (d.h. 4 Quartale). Begründen Sie anschaulich, warum in diesem Durchschnitt die jährlichen Schwankungen weitgehend verschwinden.

(b) Führen Sie eine exponentielle Glättung der Daten mit dem Glättungsparameter $\alpha = 1/2$ durch. Verwenden Sie als Anfangswert den ersten Datenwert. Welcher Glättungszeitkonstante (in Jahren oder Quartalen) entspricht obiger Glättungsparameter?

(c) Führen Sie an den Daten eine lineare Regression durch. Welchen Wert hat die Regressionsgerade im 2. Quartal 2004?

(d) Zieht man von den Daten den durch die Regressionsgerade gegebenen Trend ab (additive Aufspaltung), erhält man folgende stationäre Zeitreihe:

Jahr	Quartal 1	Quartal 2	Quartal 3	Quartal 4
2002	-137	126	150	-137
2003	-133	140	124	-133

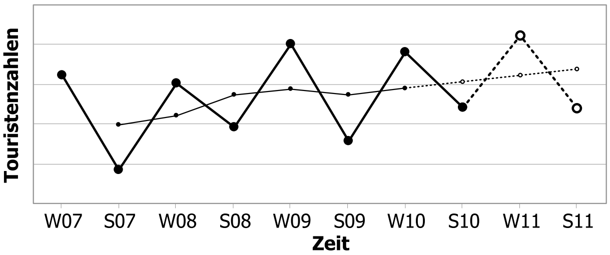
Berechnen Sie die Saisonfigur für die 4 Quartale.

(e) Geben Sie für alle 8 Zeitpunkte die Restkomponente bei einer additiven Aufspaltung an.

(f) Prognostizieren Sie die Fahrgastzahlen für die 4 Quartale des Jahres 2004 mit Hilfe der Trend- und Saisonkomponente.

Aufgabe T.1.5 – Kurort (WS 10/11)

Die Gemeindeverwaltung eines Kurortes möchte zur Kapazitätsplanung die Touristenzahlen für das bevorstehende Winter- und Sommerhalbjahr prognostizieren. Zur Prognose steht aus den vergangenen vier Jahren die Zeitreihe y_t zur Verfügung, welche angibt, wie viele Touristen (in 1000) den Kurort im Winter- bzw. Sommerhalbjahr insgesamt besuchten:



Jahr	Winter	Sommer
2007	261	142
2008	251	196
2009	301	178
2010	291	220

(a) Bestimmen Sie die glatte Komponente der Zeitreihe. Verwenden Sie ein gleitendes Mittel geeigneter Ordnung.

(b) Ermitteln Sie die Saisonfigur. Benutzen Sie das Phasendurchschnittsverfahren und geben Sie als Ergebnis die Saisonindexziffern an.

(c) Geben Sie in einem Satz an, wie der Zahlenwert der Saisonindexziffer für das Winterhalbjahr zu verstehen ist.

(d) Prognostizieren Sie die Entwicklung der glatten Komponente bis zum Sommerhalbjahr 2011 durch Fortschreiben der letzten verfügbaren Anstiegsrate.

(e) Prognostizieren Sie die zu erwartenden Touristenzahlen im Winter- und Sommerhalbjahr 2011.

(f) Geben Sie für die zurückliegende Winter- und Sommersaison 2010 an, um wie viel die Touristenzahlen über bzw. unter dem aufgrund des Modells zu erwartenden Wert gelegen haben.

Aufgabe T.1.6 – Fahrraddiebstähle (SS 2011)

Die Polizeidirektion einer Kleinstadt hatte über einen Zeitraum von drei Jahren ($i = 1, 2, 3$) jeweils in den Quartalen ($j = 1, \dots 4$) folgende Anzahl y_{ij} an Fahrraddiebstählen erfasst.

y_{ij}	$j = 1$	$j = 2$	$j = 3$	$j = 4$
$i = 1$	9	65	89	25
$i = 2$	9	41	105	25
$i = 3$	25	49	97	33

(a) Glätten Sie die Zeitreihe mit dem gleitendem Mittel der Ordnung $\tau = 4$.

(b) Berechnen Sie die Saisonfigur S_j mit dem Phasendurchschnittsverfahren.

(c) Interpretieren Sie in einem kurzen Antwortsatz den Zahlenwert $S_3 = 51$.

(d) Prognostizieren Sie die Anzahl der Fahrraddiebstähle in den vier Quartalen des Jahres $i = 4$. Gehen Sie von einer konstanten glatten Komponente $G = 50$ aus und rechnen Sie unabhängig von Ihrem Ergebnis aus (b) mit der Saisonfigur S_j von

$$S_1 = -28.5; \quad S_2 = -1.5; \quad S_3 = 51; \quad S_4 = -21. \tag{1}$$

2 Wahrscheinlichkeitsrechnung

Aufgabe T.2.1 – Auktion für ein Cabrio, SS 2009

Eine sächsische Radiostation versteigert ein Cabrio unter folgenden Bedingungen:

1. Mindestgebot ein Euro-Cent. Weiterhin ist jeder ganzzahlige Centbetrag möglich.
2. Das Cabrio bekommt derjenige Bieter mit dem *geringsten* einzelstehenden Gebot, es darf also kein weiterer Bieter genau denselben Betrag geboten haben.

(a) Es seien die Ereignisse A_n wie folgt definiert: "Zwei oder mehr Bieter haben n Cent geboten". Gewinnen kann man also nur mit Centbeträgen, für die $\overline{A_n}$ gilt. Wie hoch ist die Wahrscheinlichkeit $P(G_x)$ dafür, dass man mit einem Gebot von x Cent das Cabrio bekommt? Nehmen Sie an, dass alle Mitbieter unabhängig agieren und drücken Sie das Ergebnis als Funktion der Wahrscheinlichkeiten $P(A_n)$ aus.

(b) Gegeben sind nun die Wahrscheinlichkeiten $P(A_n) = e^{-\lambda n}$ mit $\lambda = 1/100\,000$. Mit welcher Wahrscheinlichkeit haben also mindestens zwei Bieter genau 1 000 Euro geboten?

(c) Zeigen Sie, dass man mit den bei (b) gegebenen Wahrscheinlichkeiten und der Summe

$$\sum_{n=1}^{x-1} n = \frac{x(x-1)}{2} \approx \frac{x^2}{2} \tag{2}$$

den Logarithmus der Gewinnwahrscheinlichkeit schreiben kann als

$$\ln P(G_x) \approx -\frac{\lambda x^2}{2} + \ln\left(1 - e^{-\lambda x}\right). \tag{3}$$

(d) Maximieren Sie nun die Gewinnwahrscheinlichkeit durch Variation des Gebotes x. Zeigen Sie, dass man für x die nichtlineare Bedingung

$$x\left(1 - e^{-\lambda x}\right) = 1 \tag{4}$$

bekommt. Warum ist es zur Bestimmung des optimalen Gebots gleichgültig, ob man die Wahrscheinlichkeit oder deren Logarithmus maximiert?

(e) Die Bedingung in (d) besitzt näherungsweise die Lösung $x = 1/\sqrt{\lambda}$. Wie hoch ist also das optimale Gebot und wie hoch die dazugehörige Wahrscheinlichkeit, den Zuschlag zu bekommen?

Aufgabe T.2.2 – Informationen aus dem Mobilitätsverhalten, WS 2008/09

Bei einer Erhebung zum Mobilitätsverhalten soll auch das Einkommen als relevante Einflussgröße ermittelt werden. Da es heikel ist, danach direkt zu fragen, wird das Einkommen indirekt durch Erhebung von mit dem Einkommen korrelierten, aber unproblematischeren Merkmalen abgeschätzt. Konkret wird nach der Zahl der im Haushalt verfügbaren Autos gefragt. Aus Volkszähldaten ist folgendes bekannt:

- In der Einkommensklasse 1 (unter 30 000 Euro/Jahr) besitzen 40% kein Kfz, 55% ein Kfz und nur 5% mehr als 1 Kfz. In der Einkommensklasse 2 (30 000 bis 50 000 Euro/Jahr ist die Aufteilung 20% (kein Kfz), 60% (ein Kfz) und 20% (mehrere), wähernd in der Einkommensklasse 3 (>50 000 Euro/Jahr) 5% (kein Kfz), 45% (ein Kfz) und 50% (mehrere) gilt.
- 20% der Haushalte gehören zu Einkommensklasse 1, 60% zu Klasse 2 und der Rest zu Klasse 3.

Wie hoch sind die Wahrscheinlichkeiten dafür, dass (i) Haushalte ohne Kfz, (ii) Haushalte mit einem Kfz, (iii) Haushalte mit mehreren Kfz jeweils zu den Einkommensklassen 1,2 oder 3 gehören?

Aufgabe T.2.3 – Rechtschreibfehler, SS 2008

Bei der Korrektur eines Dokuments auf Rechtschreibfehler entdeckte der erste Überprüfer 5 Fehler und der zweite 4 Fehler. Darunter waren zwei Fehler, welche von beiden entdeckt wurden. Man kann davon ausgehen, dass die beiden Personen unabhängig voneinander korrigiert haben und die Wahrscheinlichkeit für das Aufspüren eines Fehlers bei allen Fehlern dieselbe ist.

(a) Nutzen Sie die Regeln für bedingte Wahrscheinlichkeiten und Wahrscheinlichkeiten unabhängiger Ereignisse, um die Gesamtzahl der Fehler abzuschätzen. Wieviele unentdeckte Fehler sind danach zu erwarten?

(b) Könnten es auch mehr oder weniger sein? Nehmen Sie dazu als Nullhypothese H_0 an, dass sich tatsächlich 10 Fehler eingeschlichen haben und die Wahrscheinlichkeit, einen Fehler aufzuspüren, bei der ersten Person bei 1/2 und bei der zweiten bei 2/5 liegt. Berechnen Sie unter dieseren Nullhypothese die Verteilung der von der ersten und zweiten Person entdeckten Fehlerzahl sowie die Verteilung der Gesamtzahl der entdeckten Fehler.

Aufgabe T.2.4 – Zinsen, SS 2008

Eine bestimmte Bank benötigt für ein Jahr Geld und kann sich dieses nur (von anderen Banken oder Investoren) ausleihen, wenn sie 6% Zinsen bietet. Ausfallsichere Bundesanleihen bieten nur 4% Zinsen. Wie hoch ist die von den Marktteilnehmern angenommene subjektive Wahrscheinlichkeit dafür, dass diese Bank innerhalb des Jahres Pleite geht (und damit das ausgeliehene Geld samt Zinsen verloren ist) ?

Aufgabe T.2.5 – Druckfehler, SS 2007

Da man „Druckfehler“ bekanntlich selbst schlecht findet, lässt ein Student seine Bachelorarbeit vor Abgabe noch von zwei Kommilitonen korrigieren. Vom ersten Korrektor werden 75 Fehler gefunden, während der zweite Korrektor unabhängig davon 60 Fehler findet. Beim Einarbeiten der Korrekturen stellt der Student fest, dass nur 30 Fehler von beiden gefunden wurden.

(a) Die Ereignisse „ein Fehler wurde vom ersten (bzw. zweiten) Korrektor entdeckt“ seien nun mit A (bzw. B) gekennzeichnet, während C das Ereignis kennzeichnet, dass beide den Fehler finden. Wie groß sind $P(C|A)$ und $P(C|B)$?

(b) Wie groß sind die unbedingten Wahrscheinlichkeiten $P(A)$ und $P(B)$? *Hinweis:* Verwenden Sie die Produktregel für Wahrscheinlichkeiten und berücksichtigen Sie die Unabhängigkeit!

(c) Rechnen Sie nun mit $P(B) = 30/75$ weiter. Wie viele unentdeckte Fehler sind in der Arbeit noch zu erwarten?

(d) Es wird nun angenommen, dass beim Schreiben der Arbeit die Wahrscheinlichkeit eines neuen Druckfehlers nicht von den bereits zuvor gemachten Fehlern abhängt. Welche der folgenden Verteilungen kann man zur Beschreibung der *Gesamtfehlerzahl* verwenden?

    (1) Gaußverteilung,
    (2) Exponentialverteilung,
    (3) Binomialverteilung
    (4) Poissonverteilung
    (5) hypergeometrische Verteilung
    (6) Gleichverteilung

    Hinweis: Es können durchaus mehrere dieser Verteilungen anwendbar sein, teilweise mit Einschränkungen. Es kann außerdem sein, dass eine Anwendung zwar möglich, aber nicht sinnvoll ist. Geben Sie all dies an!

(e) Es seien nun pro Seite noch drei unentdeckter Fehler zu erwarten. Mit welcher Wahrscheinlichkeit ist die Kurzzusammenfassung (1,5 Seiten) fehlerfrei und mit welcher Wahrscheinlichkeit enthält sie mehr als 3 Fehler?

(f) Mit welcher Wahrscheinlichkeit gibt es in der gesamten Arbeit (50 Seiten) wenigstens eine Seite ohne Fehler?

Aufgabe T.2.6 – UEFA-Cup, WS 2006/07

Heute um 20:45 findet das UEFA-Cup Fußball-Rückspiel “Ajax Amsterdam” gegen “Werder Bremen” statt. Die Quoten eines einschlägigen Anbieters von Sportwetten sind folgende:

Ereignis	Amsterdam gewinnt	Bremen gewinnt	Unentschieden
Quote (am 15.02)	2.6	3.2	2.4

“Quote” heißt dabei, dass man beim Zutreffen des jeweiligen Ereignisses seinen Einsatz, multipliziert mit der Quote bekommt, ansonsten aber nichts. Außerdem ist bekannt, dass in der Vergangenheit 3 mal Amsterdam und einmal Bremen gewann und es keine Unentschieden gab.

(a) Berechnen Sie nach der subjektiven Definition die Wahrscheinlichkeit für die Ereignisse A (Amsterdam gewinnt), B (Bremen gewinnt) und U (Unentschieden). Berücksichtigen Sie dabei, dass die Quote umgekehrt proportional zur subjektiven Wahrscheinlichkeit ist und der Wettanbieter auch einen Gewinnanteil bekommt.

(b) Wie viel Prozent des Einsatzes verdient im Mittel der Wettanbieter, wenn die subjektiven Wahrscheinlichkeiten den tatsächlichen entsprechen? *Hinweis:* Dieser Anteil ist unabhängig davon, zu welchem Anteil die Spieler auf A, B oder 0 gesetzt haben.

(c) Berechnen Sie nun noch die statistischen Wahrscheinlichkeiten für die Ereignisse A, B und U. Warum kann man die klassische (Laplace’sche) Definition nicht anwenden?

Aufgabe T.2.7 – Kundenzufriedenheit, SS 2006

Um eine Kundenreaktion auf ihre neuen Busse, Straßenbahnen und Preise zu erlangen, startet ein städtisches ÖPNV-Unternehmen eine Internet-Umfrage zur Kundenzufriedenheit. Es wird nur unterschieden zwischen (mehr oder weniger) "zufrieden" und "nicht zufrieden". Es gab 353 positive und 402 negative Rückmeldungen. Nun ist aus vergangenen ähnlichen Umfragen bekannt, dass nur 40% aller zufriedenen Kunden, welche sich den Web-Fragebogen angeschaut haben, auch eine (positive) Rückmeldung geben, während 70% aller unzufriedenen Kunden eine negative Rückmeldung geben. Die A-Priori Wahrscheinlichkeit, dass sich Kunden den Fragebogen anschauen, hängt nicht davon ab, ob sie zufrieden sind oder nicht.

(a) Zeichnen Sie einen Wahrscheinlichkeitsbaum dieser Situation. Tragen Sie die noch unbekannte Wahrscheinlichkeit p_{zuf} dafür, dass ein Kunde zufrieden ist, zunächst als allgemeines Symbol ein.

(b) Wie groß ist der Anteilswert der positiven Rückmeldungen? Wie groß ist der Erwartungswert des Anteils zufriedener Kunden, wenn Sie Stichprobenfehler nicht berücksichtigen und das Gesetz der großen Zahlen anwenden?

(c) Wie ändert sich der Erwartungswert des Anteils zufriedener Kunden, wenn jeweils 10% der Fragebogen-Beantworter "falsch" antworten, also zufriedene Kunden eine negative Rückmeldung (und umgekehrt) geben?

(d) Geben Sie das Konfidenzintervall für den Anteil an positiven Rückmeldungen zu einer Fehlerwahrscheinlichkeit von 4% an.

Aufgabe T.2.8 – Schadensmeldung, WS 2005/06

Eine Kfz-Haftpflichtversicherung teilt ihre Kunden in vier Schadenfreiheitsklassen (SF-Klassen) ein. Zur Ermittlung der "Schadenfreiheitsrabatte" wird folgende Tabelle der im letzten Jahr eingehenden Schadensmeldungen zugrundegelegt (die "Unfallfreiheit" gilt natürlich für den Zeitraum *vor* der aktuellen Schadensmeldung):

SF-Klasse	Zahl der Mitglieder	Anteil an den Schadensmeldungen
SF 1 (unfallfrei $\leq$ 2 Jahre)	50 000	30%
SF 2 (unfallfrei 2-5 Jahre)	80 000	30%
SF 3 (unfallfrei 5-10 Jahre)	80 000	24%
SF 4 (unfallfrei > 10 Jahre)	66 667	16%

Die Mitglieder der SF 1 zahlen 100% des Referenzbeitrages, die Mitglieder der SF $k = 2$ bis 4 bekommen Rabatte, so dass der zu zahlende Versicherungsbetrag proportional zur bedingten Wahrscheinlichkeit dafür ist, dass sie im Bezugszeitraum (mindestens) eine Schadensmeldung eingereicht haben. Bestimmen Sie die Rabatte mit Hilfe des Satzes von Bayes und berücksichtigen Sie dabei das Gesetz der großen Zahlen.

Hinweis: zu berechnen sind also Quotienten der bedingten Wahrscheinlichkeiten einer Schadensmeldung eines Mitglieds, falls dieses in SF k ist, bezogen auf die Schadenshäufigkeit von Mitgliedern der SF 1.

Aufgabe T.2.9 – Stau im Pendlerverkehr, SS 2005

Ein(e) Wochenendpendler(in) kennt die Stausituation auf der Heimfahrt Freitag nachmittags von Hof nach Dresden ziemlich gut: Im Abschnitt bei Chemnitz beträgt die Wahrscheinlichkeit, in einen Stau zu geraten, 20%. Ein konkreter Stau führt jedoch nur in 80% der Fälle zu einer Staumeldung. Dennoch wird im Radio im Mittel bei jeder vierten Fahrt eine Staumeldung gegeben.

(a) Aus der Statistik der letzten Jahre wurden im betreffenden Abschnitt an den letzten 100 Freitagnachmittagen 57 Staus registriert. Die daraus berechnete Stauwahrscheinlichkeit weicht signifikant von 20% ab. Wie ist das zu erklären?

(b) Wie groß ist die Wahrscheinlichkeit, in einen Stau zu geraten, wenn (i) eine Staumeldung im Radio gegeben wird, (ii) keine Staumeldung erfolgte?

(c) Wie groß ist die Wahrscheinlichkeit von Falschmeldungen, d.h. es wird eine Stauwarnung gegeben, obwohl kein Stau vorhanden ist bzw. sich dieser bereits aufgelöst hat?

(d) Es besteht die Möglichkeit, einen eventuellen Stau auf einer Nebenstrecke zu umfahren. Allerdings führt eine Stauwarnung auf der Autobahn in 50% der Fälle dazu, dass die Nebenstrecken aufgrund des durch die Warnung induzierten Verkehrs ebenfalls verstaut sind. Ohne Stauwarnung beträgt die Stauwahrscheinlichkeit auf der Nebenstrecke nur 10%. Die auf der Autobahn und der Umfahrungsstrecke benötigten mittleren Fahrtzeiten (in Minuten) ohne und mit Stau sind durch folgende Matrix gegeben:

Strecke	ohne Stau	mit Stau
Autobahn	5	30
Umfahrung	10	40

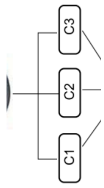
Soll der Pendler/die Pendlerin im Falle einer Stauwarnung die Umgehungsstrecke benutzen, wenn

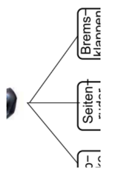
(i) ein wichtiger Termin unaufholbar verpasst werden würde, wenn man für diesen Abschnitt mehr als 20 Minuten benötigte,

(ii) der Mittelwert der erwarteten Reisezeit minimiert werden soll?

Aufgabe T.2.10 – Steuerung, WS 2004/05

"Steer By Wire" bedeutet Lenken, ohne dass eine direkte mechanische Verbindung zwischen Lenkrad und den Rädern, bzw im Flugzeug zwischen Steuerknüppel und Seitenleitwerk besteht. Fällt beim Flugzeug die elektronische Verbindung zum Seitenleitwerk aus, dann kann der Pilot das Flugzeug im Notfall noch mit den Triebwerken oder den Bremsklappen steuern. Beim Kfz dagegen ist die Verbindung zwischen Lenkrad und Rädern unersetzlich und deshalb Redundanz erforderlich.





(a) Bestimmen Sie die auf einen Flug bezogene Ausfallsicherheit von Steer-By-Wire beim Flugzeug unter den Annahmen, dass zum Steuern mindestens eines der Systeme (i) Seitenleitwerk, (ii) Triebwerke und (iii) Bremsklappen voll funktionsfähig sein muss, die drei Systeme unabhängig voneinander sind und je mit einer Wahrscheinlichkeit von 1:1000 ausfallen.

(b) Beim Vorschlag für Steer-By-Wire im Fahrzeug gemäß der Skizze wird die Lenkbewegung vom Fahrer zunächst von drei unabhängigen Systemen C1 bis C3 aus Sensor und Computer aufgenommen und an den ausfallsicheren "Voter" geschickt. Dieser wählt aus den drei anliegenden Signalen das der "Mehrheit" aus. Fällt eines der Systeme C1 bis C3 aus, wird dennoch das korrekte Signal weitergeleitet und außerdem der Fehler lokalisiert. Ein Ausfall von 2 oder 3 der Systeme führt zum Versagen. Mit welcher Wahrscheinlichkeit tritt dies pro Fahrt auf, wenn jedes der Systeme nur in einem von 10 000 Fällen versagt? *Hinweis:* Eine Berechnung auf zwei signifikante Stellen ist ausreichend.

(c) Das Signal vom Voter wird über zwei unabhängige Stellmotoren (Ausfallwahrscheinlichkeit 0.01% pro Fahrt) zu den Rädern weitergeleitet. Lenkfähigkeit ist gewährleistet, wenn mindestens einer der Stellmotoren korrekt arbeitet. Mit welcher Wahrscheinlichkeit gibt es ein Lenkversagen durch die Stellmotoren?

(d) Mit welcher Wahrscheinlichkeit versagt das Gesamtsystem "Lenkung" pro Fahrt?

(e) Mit welcher Wahrscheinlichkeit bleibt das Fahrzeug aufgrund eines "Lenkelektronikfehlers" (der Voter erhält nicht drei übereinstimmende Signale oder ein Stellmotor ist defekt) liegen? *Hinweis:* Zwei signifikante Stellen genügen.

Aufgabe T.2.11 – Kriminalität, SS 2004

Bei einer Fahrzeugkontrolle zur Ermittlung von gestohlenen Fahrzeugen und Drogenbesitz werden u.a. die Fahrzeugpapiere überprüft. Von den Fahrern nicht gestohlener Fahrzeuge haben 10% den Fahrzeugschein nicht dabei. Die Fahrer gestohlener Fahrzeuge hingegen können zu 50% gefälschte oder im Fahrzeug gefundene Papiere vorweisen. Außerdem ist bekannt, dass in 1% der Fälle ein gestohlenes Fahrzeug kontrolliert wird sowie in 10% aller gestohlenen Fahrzeuge, aber nur in 0.1% aller übrigen Fahrzeuge Drogen mitgeführt werden.

(a) Mit welcher Wahrscheinlichkeit fahren Personen, die keinen Fahrzeugschein vorzeigen können, einen gestohlenen Wagen? Wie hoch ist die Wahrscheinlichkeit bei den restlichen Personen? Wie hoch ist die unbedingte Wahrscheinlichkeit dafür, dass Papiere vorgezeigt werden können?

(b) Wie hoch sind die Wahrscheinlichkeiten, dass Personen mit bzw. ohne Fahrzeugschein Drogen transportieren? *Hinweis:* Am leichtesten geht es mit dem Wahrscheinlichkeitsbaum!

Aufgabe T.2.12 – Trunkenheit am Steuer, WS 2003/04

Aus statistischen Untersuchungen ist bekannt, dass alkoholisierte Autofahrer pro gefahrenen Kilometer ein zehnmal so hohes Unfallrisiko haben wie nichtbetrunkene Fahrer. Ferner ist von Fahrkontrollen her bekannt, dass 5% aller kontrollierten Fahrer alkoholisiert sind.

(a) Berechnen Sie mit dem Satz von Bayes die Wahrscheinlichkeit dafür, dass ein Unfallverursacher alkoholisiert ist.

Hinweis: In der Formel für den Satz von Bayes kürzen sich die nicht angegebenen absoluten Zahlenwerte für die Unfallwahrscheinlichkeiten heraus. Sie können also z.B. die bedingten Wahrscheinlichkeit dafür, dass ein alkoholisierter Fahrer einen Unfall verursacht, zu 10% annehmen.

(b) Die vor Ort verwendeten Alkoholtestgeräte sind nicht unfehlbar und zeigen nur bei 90% aller Alkoholisierten ein positives Ergebnis an, während sie bei 5% der Nichtbetrunkenen einen falschen Alarm geben. Nach einer Alkoholkontrolle mit positivem Ergebnis zweifelt ein Betroffener das Ergebnis der Messung an. Mit welcher Wahrscheinlichkeit hat er Recht, wenn die Messung

(i) im Rahmen einer gewöhnlichen Fahrerkontrolle,

(ii) nach einem von ihm verursachten Unfall

stattfand?

Hinweis: Falls Sie Teil (a) nicht gerechnet haben, verwenden Sie den Wert 34.5 % für die Wahrscheinlichkeit der Alkoholisierung eines Unfallverursachers.

Aufgabe T.2.13 – Panne, SS 2003

Eine Speditionsfirma transportiert unter anderem Maschinenteile von Deutschland in die Türkei (Wegstrecke: 4000 km). Da eine verzögerte Lieferung mit hohen Konventionalstrafen verbunden ist, ist vor jedem dieser Transporte eine Inspektion des LKW vorgesehen, die jedoch von den Fahrern aus Bequemlichkeit in 20% der Fälle nicht durchgeführt wird. Ohne Inspektion erleidet der LKW pro 1000 gefahrene km mit 3% Wahrscheinlichkeit für eine Panne, die zu einer unzulässigen Verzögerung führt, mit Inspektion nur mit 0,5%.

(a) Wie groß ist die Wahrscheinlichkeit mindestens einer Panne auf der 4000 km langen Strecke ohne und mit Inspektion?

(b) Mit welcher Wahrscheinlichkeit hat ein auf der Strecke liegengebliebener Fahrer die Inspektion nicht durchgeführt ?

Hinweis: Berechnen Sie zunächst die mittlere Pannenwahrscheinlichkeit durch entsprechende Gewichtung der in (a) berechneten Wahrscheinlichkeiten (Lösung: 3,88%) und wenden Sie dann den Satz von Bayes an!

(c) Neben Pannen gibt es mit 1% Wahrscheinlichkeit andere Gründe, die zu unzulässigen Verzögerungen führen wie z.B. Zoll oder Verkehrsstaus. Mit welcher Wahrscheinlichkeit kommen die Maschinenteile verspätet an?

Aufgabe T.2.14 – JobTicket (WS 10/11)

Mit JobTickets können Mitarbeiter größerer Unternehmen vergünstigt öffentliche Verkehrsmittel nutzen. Die Geschäftsleitung eines Dresdner Unternehmens wittert Einsparpotential an Pkw-Stellplätzen und möchte seinen Mitarbeitern dieses JobTicket anbieten. Eine Befragung unter den Mitarbeitern hat ergeben, dass 50% aller Mitarbeiter das JobTicket nutzen würden. Unter den JobTicket-Befürwortern fahren derzeit 40% dem eigenen Pkw zur Arbeit, unter den JobTicket-Verweigerern sind es 80%.

Verwenden Sie zur Beantwortung der Fragen (a) und (b) die beiden Ereignisse

$$
\begin{aligned}
A &= \text{„Mitarbeiter wird künftig das JobTicket nutzen"} \\
B &= \text{„Mitarbeiter fährt bislang mit dem Pkw zur Arbeit"}
\end{aligned}
$$

(a) Wie wahrscheinlich fährt ein zufällig ausgewählter Mitarbeiter derzeit mit dem Pkw zur Arbeit?

(b) Wie wahrscheinlich entscheidet sich ein Mitarbeiter, der heute seinen Pkw benutzt, für das JobTicket? Wenden Sie den Satz von Bayes an!

Die Geschäftsleitung geht davon aus, dass nach Einführung des JobTickets jeder Pkw-Stellplatz (unabhängig von allen anderen, d. h. es gibt keine Stellplatz-Präferenzen) nur noch mit einer Wahrscheinlichkeit von 66.7% belegt sein wird.

(c) Wie wahrscheinlich werden dann von den vier Stellplätzen, die regelmäßig die Müllabfuhr behindern, nur noch höchstens zwei belegt sein?

(d) Wie wahrscheinlich werden von den derzeit 208 Stellplätzen des Firmenparkplatzes künftig nur noch höchstens 150 benötigt? (Hinweis: Approximieren Sie die Binomialverteilung mit Hilfe des Zentralen Grenzwertsatzes durch die Normalverteilung. Lesen Sie aus der Quantilstabelle den nächstgelegenen Wert ab.)

Aufgabe T.2.15 – Autohändler (SS 2011)

Ein Autohändler stellt fest, dass die Entscheidung seiner Kunden, eher ein Familien- oder Sportmodell zu kaufen, davon abhängig ist, ob der Kunde Kinder hat. Es entscheiden sich Kunden mit mindestens einem Kind zu 80% für ein Familienmodell. Dagegen wählen 70% der Kunden, die keine Kinder haben, ein Sportmodell. Der Autohändler geht davon aus, dass 60% seiner Kunden kinderlos sind.

(a) Geben Sie bezüglich der Ereignisse

$$\begin{aligned} A &= \text{„Der Kunde hat mindestens ein Kind."} \\ B &= \text{„Der Kunde kauft ein Familienmodell."} \end{aligned}$$

die bedingten Wahrscheinlichkeiten $P(B|A)$, $P(\overline{B}|A)$, $P(B|\overline{A})$ und $P(\overline{B}|\overline{A})$ an.

(b) Ein neuer Kunde betritt das Autohaus. Berechnen Sie die Wahrscheinlichkeit dafür, dass dieser ein Familienmodell kaufen wird.

(c) Mit welcher Wahrscheinlichkeit hat dieser Kunde mindestens ein Kind, falls er sich für ein Familienmodell entscheidet? Wenden Sie den Satz von Bayes an.

(d) Mit welcher Wahrscheinlichkeit ist dieser Kunde kinderlos, falls er ein Sportmodell kauft?

(e) Der Autohändler im Nachbarort verkauft zu 55% Familienmodelle. Wie wahrscheinlich hat ein (beliebiger) Kunde dieses Verkäufers mindestens ein Kind? Legen Sie die Entscheidungswahrscheinlichkeiten aus (a) zugrunde.

3 Wahrscheinlichkeitsverteilungen

Aufgabe T.3.1 – Zugfahrt (WS 2009/10)

Für die Fahrt mit dem Zug von A-Stadt nach B-Stadt gibt es zwei Verbindungen: Einerseits einen ICE mit einmal Umsteigen in C-Stadt. Der ICE ist häufig verspätet. Die Verspätung der Ankunft in C-Stadt (in Minuten) kann (0,30)-gleichverteilt angenommen. Die Umsteigezeit in C-Stadt beträgt aber nur 20 min und der Anschlusszug ist stets pünktlich. Verpasst man ihn, gibt es eine Stunde später wieder eine (ebenfalls pünktliche) Verbindung. Alternativ kann man eine Direktverbindung von A nach B wählen, deren planmäßige Ankunftszeit aber 30 min später ist und deren Verspätung (0,20) gleichverteilt ist.

(a) Zeichnen Sie für jede der beiden Verbindungen die Verteilungsfunktionen der Ankunftszeiten T. Beziehen Sie dabei alles auf die planmäßige Ankunft der ersten Verbindung (Ankunftszeit $T = 0$).

(b) Ermitteln Sie für beide Verbindungen den Median und den Erwartungswert der Ankunftszeit.

(c) Welche Verbindung würde man sinnvollerweise wählen, wenn man

  (i) regelmäßig nach B-Stadt fährt und im Mittel möglichst viel Zeit sparen will,

  (ii) spätestens $t_c = 50$ min nach der frühestmöglichen Ankunft am Ziel sein muss?

  Berechnen Sie dazu jeweils die Wahrscheinlichkeiten $P(T \leq 50 \text{ min})$.

(d) Berechnen Sie für beide Verbindungen jeweils

  (i) die Standardabweichung,

  (ii) die Wahrscheinlichkeit $P(T < 45 \text{ min})$,

  (iii) und die Spannweite.

Aufgabe T.3.2 – Verkehrsmittelwahl binomialverteilt, SS 2009

Zehn Studenten und zehn Erwerbstätige wurden gefragt, ob sie bei gleichen Reisezeiten und gleichen Kosten einen bestimmten Weg lieber mit dem (vorhandenen) eigenen Fahrzeug oder mit dem ÖPNV zurücklegen würden. Sieben Studenten und zwei Erwerbstätige entschieden sich für den ÖPNV.

(a) Unter welchen Bedingungen kann in jeder Gruppe die Zahl der ÖPNV-Entscheidungen mit der Binomialverteilung beschrieben werden? Welcher der beiden Parameter dieser Verteilung ist dann fest?

(b) Der Parameter θ der Binomialverteilung soll nun so bestimmt werden, dass mit größtmöglicher Wahrscheinlichkeit die Zahl der getroffenen ÖPNV-Entscheidungen reproduziert wird. Wie groß sind (auf 10% gerundet) die Parameterwerte für die Studenten- und die Berufstätigengruppe? Lesen Sie das Ergebnis aus folgender Tabelle der $(10, \theta)$-Binomialverteilung ab.

θ	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80
$x = 1$	0.39	0.27	0.12	0.040	0.0098	0.0016	0.00014	4.1e-06
$x = 2$	0.19	0.30	0.23	0.12	0.044	0.011	0.0014	7.4e-05
$x = 3$	0.057	0.20	0.27	0.21	0.12	0.042	0.0090	0.00079
$x = 4$	0.011	0.088	0.20	0.25	0.21	0.11	0.037	0.0055
$x = 5$	0.0015	0.026	0.10	0.20	0.25	0.20	0.10	0.026
$x = 6$	0.00014	0.0055	0.037	0.11	0.21	0.25	0.20	0.088
$x = 7$	8.7e-06	0.00079	0.0090	0.042	0.12	0.21	0.27	0.20

(c) Leiten Sie die optimalen θ-Werte nun auch analytisch her, indem Sie die binomialverteilte Wahrscheinlichkeit, bei 10 Versuchen 2 bzw. 7 Treffer zu erhalten, nach θ ableiten und nullsetzen.

(d) Ein LKW fiel während der letzten 100 000 km zweimal wegen Defekten aus. Unter welchen Bedingungen kann man die Ausfallhäufigkeit mit einer Poissonverteilung beschreiben? Warum sind Binomial- und Gaußverteilungen auf jeden Fall hier ungeeignet?

(e) Bestimmen Sie den Parameter der Poissonverteilung, für den die Wahrscheinlichkeit zweier Ausfälle maximal ist.

Aufgabe T.3.3 – Kl. SS 2005, Fußgänger

Die Stadtverwaltung erwägt, auf einer relativ großen Straße (2 Spuren in jeder Richtung) in der Nacht die Ampeln abzuschalten. Wegen der langen Überquerungszeit zu Fuß muss jedoch sichergestellt werden, dass die Straße nach wie vor von Fußgängern überquert werden kann.

Verkehrsmessungen ergaben für die geplanten Abschaltperioden über alle zu überquerende Spuren summierte maximale Verkehrsaufkommen von 600 Fahrzeugen pro Stunde. Dabei werden die zeitlichen Abstände zwischen den vorbeifahrenden Fahrzeugen exponentialverteilt angenommen. Aus der Straßendimension ergibt sich eine anzusetzende Fußgänger-Überquerungszeit (einschließlich Sicherheitspuffer) von 12 Sekunden.

(a) Begründen Sie zunächst (ohne Rechnung!), warum bei hohem Verkehrsaufkommen auf einspuriger Strecke eine Exponentialverteilung der Abstände keine dem Sachverhalt angemessene Annahme ist, im zu untersuchenden Fall (geringe Verkehrsaufkommen, mehrere Spuren) jedoch schon.

Hinweis: Betrachten Sie den Modalwert der Exponentialverteilung und den Zusammenhang der Exponential- mit der Poissonverteilung.

(b) Mit welcher Wahrscheinlichkeit p_{frei} kann ein an die Straße kommender Fußgänger diese ohne Wartezeit sofort überqueren? Hängt die Wahrscheinlichkeit von der Zeitspanne zwischen Eintreffen des Fußgängers und Vorbeifahrt des letzten Fahrzeugs ab?

(c) Wieviel Fahrzeuge muss der Fußgänger im Mittel vorbeifahren lassen, bis er die Straße überqueren kann? Berechnen Sie dazu die Wahrscheinlichkeiten, dass eine hinreichend große Lücke erst nach $n = 0, 1, 2, \ldots$ Fahrzeugen auftritt. Diese Wahrscheinlichkeiten definieren eine Verteilung, deren Mittelwert der gesuchten mittleren Fahrzeugzahl entspricht.

Hinweis: Falls Sie (b) nicht bearbeitet haben, nehmen Sie $p_{frei} = 1/e^2$ an. Berücksichtigen Sie, dass die einzelenen Zeitlücken unabhängig sind. Verwenden Sie das Multiplikationsgesetz für Wahrscheinlichkeiten unabhängiger Ereignisse und die für $|q| < 1$ gültige Beziehung

$$\sum_{n=0}^{\infty} n q^n = \frac{q}{(1-q)^2} \tag{5}$$

(d) Wie groß ist die mittlere Wartezeit für die Füßgänger?

Aufgabe T.3.4 – Induktionsschleifen, WS 2008/09

Auf Fernstraßen und Autobahnen werden Fahrzeuge automatisch durch Induktionsschleifen detektiert, wenn diese an festen Streckenquerschnitten vorbeifahren. Dabei wird der Zeitpunkt des Überquerens (und meist auch die Geschwindigkeit) aufgezeichnet.

(a) Bei geringem Verkehrsaufkommen ist die Zeitlücke zwischen zwei Fahrzeugen angenähert exponentialverteilt. Bestimmen Sie den Parameter dieser Verteilung als Funktion des Verkehrsflusses Q (Fahrzeuge pro Zeiteinheit).

(b) Gegeben ist nun ein Fluss von 1 Fahrzeug pro Minute. Mit welcher Wahrscheinlichkeit kommen in den nächsten 5 Minuten höchstens zwei Fahrzeuge vorbei? Wie hoch ist der Median des zeitlichen Abstandes zwischen zwei Fahrzeugen?

(c) Nehmen Sie nun an, dass die Geschwindigkeit aller Fahrzeuge genau 72 km/h bzw. 20 m/s beträgt. Wie lautet die Verteilungsfunktion der räumlichen Fahrzeugabstände, wenn das Verkehrsaufkommen nach wie vor 1 Fahrzeug pro Minute beträgt?

(d) Für höhere Verkehrsaufkommen (dichter Verkehr) und Richtungsfahrbahnen mit nur einem Fahrstreifen ist die Exponentialverteilung zur Beschreibung der Abstände nicht geeignet. Ermitteln Sie dazu den aus ihr folgenden wahrscheinlichsten (zeitlichen) Abstand und erläutern Sie den Sachverhalt.

(e) Um einen zeitlichen Mindestabstand $x_{min} = 1$ s $= 1/60$ min zu berücksichtigen, wird die Verteilung modifiziert. Die Dichtefunktion der neuen Verteilung ist gegeben durch

$$f(x) = \begin{cases} 0 & x < x_{min} \\ \tilde{\lambda}\, e^{-\tilde{\lambda}(x - x_{min})} & x \geq x_{min}. \end{cases} \tag{6}$$

Ermitteln Sie $\tilde{\lambda}$ in Abhängigkeit des Verkehrsaufkommens Q so, dass der mittlere zeitliche Abstand $E(X)$ nach wie vor durch den Kehrwert $1/Q$ des Verkehrsflusses gegeben ist.

Hinweis: $\int\limits_0^\infty x e^{-\tilde{\lambda}x} \mathrm{d}x = \frac{1}{\tilde{\lambda}^2}$.

(f) Beachten Sie nun die beiden Werte $Q_1 = 1$ Kfz/min und $Q_2 = 30$ Kfz/min des Verkehrsaufkommens (geringer bzw. starker Verkehr). Für den Parameter $\tilde{\lambda}_1$ der Verteilung von Aufgabenteil (e) gilt dann $\tilde{\lambda}_1 = 1.01695$ (geringer Verkehr) bzw. $\tilde{\lambda}_2 = 60$ (starker Verkehr). Beurteilen Sie die Auswirkung der Modellierung des Mindestabstandes für beide Fälle, indem Sie jeweils

(i) den Wert des Parameters $\tilde{\lambda}$ mit dem entsprechenden Wert aus der Exponentialverteilung vergleichen,

(ii) den Median der beiden Verteilungen vergleichen.

Für welches der beiden Verkehrsaufkommen Q_1 und Q_2 fällt der Unterschied zwischen der ursprünglichen und der modifizierten Exponentialverteilung geringer aus?

Aufgabe T.3.5 – Direktverbindung, WS 2007/08

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.

(a) Mit welcher Wahrscheinlichkeit würde man bei der zweiten Verbindung den Zug verpassen?

(b) Berechnen Sie die Verteilungsfunktionen der tatsächlichen Reisezeiten der beiden Verbindungen (in Minuten). Wie hoch sind jeweils die Erwartungswerte?

(c) 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!)

Aufgabe T.3.6 – Schwangerschaft, SS 2007

Eine Schwangerschaft dauert, vom Zeitpunkt der letzten Menstruation aus gerechnet, im Mittel 280 Tage (Naegelesche Regel). Zum berechneten Termin kommen jedoch nur vier Prozent der Kinder zur Welt.

(a) Die Schwangerschaftsdauer kann sehr gut mit einer Gaußverteilung beschrieben werden. Geben Sie deren Parameter an!

(b) Kommt das Kind mehr als 3 Wochen vor dem mit der "Naegeleschen Regel" errechneten Mittelwert zur Welt, gilt dies als Frühgeburt. Wie groß ist die Wahrscheinlichkeit dafür? *Hinweis:* Nehmen Sie für diese und die folgenden Teilaufgaben eine Standardabweichung von 10 Tagen an.

(c) Das Kind ist am errechneten Geburtstermin noch nicht da und es gibt auch keine Anzeichen, dass die Geburt unmittelbar bevorsteht. Warum ist es problematisch, die Verwandten auf einen "Besichtigungstermin" in zwei Wochen zu bestellen? Berechnen Sie dazu die Wahrscheinlichkeit dafür, dass die Geburt in den nächsten zwei Wochen stattfindet! (Achtung: Sie müssen mit bedingten Wahrscheinlichkeiten rechnen!)

Aufgabe T.3.7 – Straßenübergang mit Ampel, WS 2006/07

Ein Fußgänger beachtet die Verkehrsregeln und geht nur bei "Grün" über Straßenübergänge mit Lichtsignalanlagen. Auf den Weg zur Arbeit und zurück muss er jeweils einen signalisierten Übergang überqueren. Die Ampel hat keine Grün-Anforderung, 90 s Umlaufzeit und 30 s Grünphase.

(a) Zeigen Sie, dass bei zufälliger Ankunft des Fußgängers die Verteilungsfunktion für die Wartezeit x in Sekunden durch

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 & x > 60 \\ \frac{1}{3} + \frac{x}{90} & \text{sonst} \end{cases} \tag{7}$$

gegeben ist.

(b) Der Fußgänger geht während eines Jahres 200 mal zu Fuß zur Arbeit und zurück. Berechnen Sie mit Hilfe des zentralen Grenzwertsatzes die Verteilungsfunktion für die gesamte Wartezeit. *Hinweis:* zunächst müssen Sie Erwartungswert und Varianz der Verteilung berechnen. Falls Ihnen das nicht gelingt, verwenden Sie $\mu_x = 20$ und $\sigma_x^2 = 400$.

(c) Mit welcher Wahrscheinlichkeit wartet der Fußgänger mehr als 2 h bzw. mehr als 2 1/2 Stunden im Jahr?

Aufgabe T.3.8 – Bücher und Musikstücke, SS 2006

Bei Produkten wie Büchern und Musikstücken gibt es wenige sehr gut verkäufliche Titel ("Top 100") und viele, von denen nur wenige Exemplare oder gar nichts verkauft wird.

Der von Titeln mit jährlichen Mindestumsatz 1 generierte Umsatz x (pro Titel, in Geldeinheiten) lässt sich durch folgende Verteilungsfunktion beschreiben:

$$F(x) = 1 - x^{-\alpha} \tag{8}$$

mit $\alpha = 1.1$.

(a) Mit welcher Wahrscheinlichkeit erzeugt ein Titel einen Umsatz größer 100?

(b) Wie groß ist der Umsatz, welcher von nur 0.1 % der Titel erreicht oder überschritten wird?

(c) Berechnen Sie den mittleren Umsatz $\bar{x}$ eines Titels (beachten Sie, dass wegen des Mindestumsatzes die Integration bei 1 beginnt!).

(d) Für (quasi-)kontinuierliche Variablen ist der kumulierte relative Anteil der Merkmalssumme gegeben durch

$$P(x) = \frac{1}{\bar{x}} \int_1^x x' f(x') \mathrm{d}x'. \tag{9}$$

Berechnen Sie $P(x)$. Was sagt $P(x)$ aus?

(e) Ein Buchhändler kann nur Bücher auf Lager halten, welche einen Mindestumsatz von 100 bringen, während ein Internet-Buchhändler alle Bücher mit Mindestumsatz 1 zur Verfügung stellen kann. Wie groß ist der Teil des Gesamt-Umsatzes, den der Internet-Händler mit den Büchern macht, welche pro Jahr nur 1-100 Geldeinheiten an Umsatz erbringen?

Hinweis: Falls Sie Teil (d) nicht gelöst haben, verwenden Sie $P(x) = 1 - x^{1-\alpha}$

4 Induktive Statistik

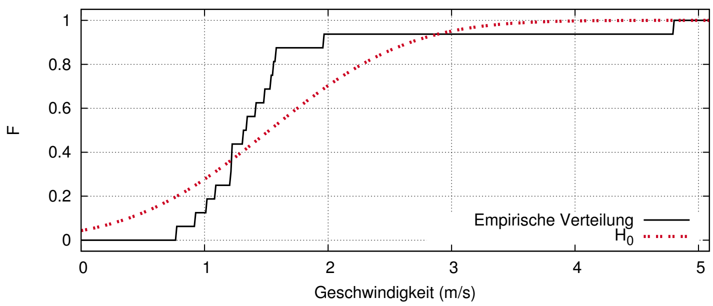
Aufgabe T.4.1 – Fußgängerampel, WS 2009/10

Um die Mindest-Grünzeit von Fußgängerampeln in Abhängigkeit der Straßenbreite festzulegen, werden zunächst die Geschwindigkeiten einer repräsentativen Stichprobe von Fußgängern ermittelt. Es ergaben sich folgende Geschwindigkeiten in Metern pro Sekunde:

1.31	1.34	1.96	0.92	4.80	1.56	1.01	1.58
1.48	1.53	1.09	1.21	0.77	1.22	1.22	1.41

(a) Geben Sie die effizienten Schätzer $\hat{\mu}$ und $\hat{\sigma}$ für Erwartungswert bzw. Standardabweichung der Geschwindigkeiten an.

(b) Testen Sie die Verteilung der Geschwindigkeiten X mit dem Kolmogorow-Smirnov-Test auf die Nullhypothese $H_0 : X \sim N(\hat{\mu}, \hat{\sigma}^2)$ bei einer Fehlerwahrscheinlichkeit von 5 %. Ermitteln Sie dabei die Realisierung der Test-Statistik graphisch unter Verwendung folgender Abbildung:



(c) In den Daten gibt es einen Ausreißer, der aber nicht auf Messfehlern beruht. Warum kann bzw. sollte man ihn trotzdem für die Ermittlung der Mindest-Freigabezeiten ignorieren?

(d) Nach Eliminerung des Ausreißers erhält man die effizienten Schätzer $\hat{\mu} = 1.31$ m/s und $\hat{\sigma} = 0.289$ m/s und H_0: $X \sim N(\hat{\mu}, \hat{\sigma}^2)$ kann nicht abgelehnt werden (dies braucht nicht gezeigt zu werden). Ermitteln Sie unter der Bedingung H_0 das 90. und das 99. Perzentil der Geschwindigkeiten.

(e) Für die Bemessung der Mindest-Grünzeiten der Fußgängerampel an einer Straße der Breite $b = 10$ m wird das 90. bzw. 99. Perzentil der Überquerungszeit $Y = b/X$ zugrundegelegt. Ist Y gaußverteilt, wenn X gaußverteilt ist? Berechnen Sie unter der Bedingung H_0 das 90. bzw. 99. Perzentil von Y.

Aufgabe T.4.2 – Billigtickets, SS 2009

Eine mögliche Verkehrslenkungsmaßnahme zur räumlichen und zeitlichen Entzerrung der Fahrgastzahlen in Zügen sind Kontingente von Billigtickets ("Dauer Spezial", "Dauer Spezial Familie", "Sparpreis-Ticket" etc), die für jeden Zug je nach erwarteter Belegung mit regulären Tickets ausgegeben werden. Es sollen nun die Kontingente für eine bestimmte Fahrt von München nach Hamburg (950 Sitzplätze) und eine von Hof nach Plauen (650 Plätze) bestimmt werden. Aus der Vergangenheit liegen für gleiche Wochentage und gleiche Zeiten folgende Belegungszahlen mit regulären Tickets vor:

München – Hamburg	675	812	634	802	777	565
Hof – Plauen	175	212	234	302	177	225

(a) Berechnen Sie für beide Verbindungen Schätzer für die mittlere Zahl der Kunden mit regulären Tickets und die dazugehörige Standardabweichung.

(b) Es sollen maximal so viele Billigtickets ausgegeben werden, dass die Wahrscheinlichkeit für eine Überfüllung des Zuges (mehr Fahrgäste als Sitzplätze) höchstens 10% beträgt. Berechnen Sie die für beide Verbindungen die Höchstzahl dieser Tickets

Hinweis: Wegen des unbekannten wahren Mittelwerts müssen Sie mit der Student-Verteilung rechnen. Ermitteln Sie unter Verwendung eines geeigneten Quantils und dem Schätzer der Standardabweichung die Zahl der regulären Tickets, die nur in 10% der Fälle überschritten wird. Ziehen Sie diese von der Sitzplatzzahl ab.

(c) Nun soll noch die Verkehrslenkungseigenschaft der Billigtickets getestet werden. Dazu stehen größere Stichproben der Gesamt-Fahrgastzahlen vor der Einführung der Billigtickets (Umfang 52) und danach (Umfang 36) zur Verfügung. Für die beiden Verbindungen ergaben sich folgende Schätzer für die Mittelwerte $\hat{\mu}_a$ bzw. $\hat{\mu}_e$ davor bzw. danach sowie die dazugehörigen Standardabweichungen:

$$\text{München – Hamburg: } \hat{\mu}_a = 890, \quad \hat{\mu}_e = 850, \quad \hat{\sigma}_a = 110, \quad \hat{\sigma}_e = 70, \tag{10}$$

$$\text{Hof – Plauen: } \hat{\mu}_a = 290, \quad \hat{\mu}_e = 350, \quad \hat{\sigma}_a = 60, \quad \hat{\sigma}_e = 90. \tag{11}$$

Ermitteln Sie mit einem geeigneten Differenztest (vgl. Vorlesungsfolien) die minimale Fehlerwahrscheinlichkeiten, bei der bei der Verbindung München-Hamburg die Nullhypothese H_0: "Fahrgastzahl ist seit Einführung der Billigtickets gesunken" und für Hof - Plauen die Nullhypothese "Fahrgastzahl ist gestiegen" abgelehnt werden kann. Ist bei einer Fehlerwahrscheinlichkeit von 1% der Anstieg bzw. die Abnahme signifikant?

Hinweis: Wegen der hinreichend großen Stichprobenumfänge können Sie hier mit der Gauß-Statistik rechnen. Verwenden Sie den nächstgelegenen Wert der Tabelle der Standardnormalverteilung

Aufgabe T.4.3 – Motorprüfstand, WS 2008/09

Auf einem Motorprüfstand wird die Leistung vier gleichartiger Motoren gemessen:

Motoren-Nr	1	2	3	4
Leistung (kW)	98.5	99.0	95.5	98.0

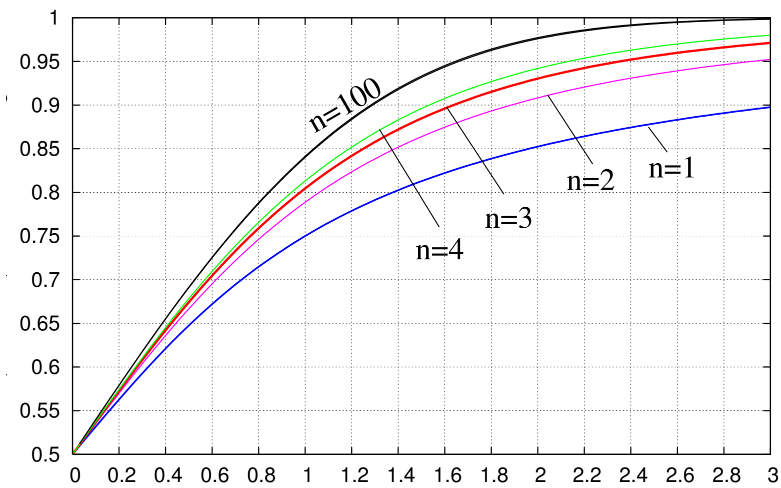
(a) Berechnen Sie zu einer Fehlerwahrscheinlichkeit von 5% das Konfidenzintervall der mittleren Leistung

(b) Der Fahrzeughersteller gibt die mittlere Leistung mit "mindestens 100 kW" an. Kann man ihn bei einer Fehlerwahrscheinlichkeit von 5% der Falschaussage bezichtigen?

(c) Das in (a) errechnete Konfidenzintervall enthält den Wert der Nullhypothese des Aufgabenteils (b), dennoch kann man in (b) die Nullhypothese verwerfen. Erläutern Sie diese scheinbar widersprüchlichen Ergebnisse.

(d) Zu welcher Fehlerwahrscheinlichkeit kann man die Nullyhpothese "Leistung mindestens 100 kW" gerade noch verwerfen? Lesen Sie das Ergebnis von der Abbildung am Ende dieser Aufgabe ab! Falls Sie (b) nicht gerechnet haben, nehmen Sie einen Stichprobenwert der Test-Statistik von -2.8 an.

(e) Wie groß ist die Wahrscheinlichkeit für einen "Fehler zweiter Art", wenn der tatsächliche Mittelwert der (gaußverteilten) Motorleistung $\mu = 98$ kW beträgt und der Test wie bei (b) durchgeführt wird? *Hinweis:* Sie müssen hier anhand der Test-Statistik bezüglich der ursprünglichen Nullhypothese $\mu_0 = 100$ kW den Wert $\bar{x}_c$ von $\bar{X}$ finden, oberhalb dem der Test fälschlicherweise angenommen wird. Tatsächlich gehorcht aber die "wahre" Statistik $\sqrt{n}(\bar{X} - \mu)/S$ und nicht $\sqrt{n}(\bar{X} - \mu_0)/S$ der Studentverteilung. Setzen Sie nun $\bar{x}_c$ in die wahre Statistik ein. Verwenden Sie wieder die Grafik



Aufgabe T.4.4 – Fahrassistenzsysteme, SS 2008

Für neue Fahrerassistenzsysteme (z.B. ein Kreuzungsassistent oder eine verbesserte Navigation) ist es wichtig, zu bestimmen, wie stark sich die Fahrtrichtung innerhalb eines bestimmten Zeitraums ändert. Diese Änderung kann prinzipiell auf dreierlei Arten gemessen werden:

1. Über den aufgezeichneten Lenkwinkel und die Geschwindigkeit,
2. über einen "Gierratensensor", welcher direkt die Drehung des Fahrzeuges misst,
3. und über die Auswertung der vom Navigationssystem gelieferten Positionen.

Für eine tatsächliche Änderung der Fahrtrichtung von 30 Grad liefern alle Methoden im Mittel das richtige Ergebnis, aber mit Unsicherheiten (einfachen Standardabweichungen) von $\sigma_1 = 2$ Grad, $\sigma_2 = 2$ Grad und $\sigma_3 = 4$ Grad. Kann man durch Kombinationen der einzelnen Messungen den Fehler verringern?

(a) Berechnen Sie zunächst die Standardabweichung der Richtungsbestimmung, falls dazu das arithmetische Mittel der drei Messungen verwendet wird.

(b) Kann man die Genauigkeit optimieren, indem man die "schlechtere" Messmethode 3 ignoriert oder weniger gewichtet? Bestimmen Sie dazu die Varianz eines gewichteten arithmetischen Mittels der drei Messungen, wobei Sie die ersten beiden Messungen gleich gewichten. Minimieren Sie die Varianz durch Variation der Gewichtungen. Mit welchem Anteil trägt die Messmethode 3 im optimalen Fall noch bei? Wie hoch ist die resultierende Varianz?

Hinweis: Die Summe der Wichtungen muss 1 betragen!

Aufgabe T.4.5 – Immatrikulation, SS 2008

Bei der Neueinschreibung an eine Universität bewerben sich die Studenten im allgemeinen an mehreren Universitäten gleichzeitig. Die Universitäte nehmen die geeigneten Kandidaten an. Allerdings erhalten viele Studenten von mehreren Universitäten Einladungen und wählen nun ihrerseits aus. Um die Studienplätze eines Studiengangs möglichst optimal zu besetzten, werden daher mehr Studenten akzeptiert als Studienplätze zur Verfügung stehen. In einem bestimmten Studiengang mit 200 Plätzen ergaben sich in den letzten Jahren bei den ersten 200 Einladungen folgende Absagen von Seiten der Studenten:

Jahr	2005	2006	2007	2008
Zahl der Absagen	83	103	63	111

(a) Geben Sie Schätzer für den Erwartungswert und die Varianz der Zahl der Absagen bzw. der Absagerquote an.

(b) Mit welcher Wahrscheinlichkeit ergibt sich höchstens 80 Absagen, wenn Sie eine Gaußverteilung der Absagenquoten mit einem Erwartungswert von 45% und einer Standardabweichung von 10% annehmen?

(c) Berechnen Sie das Konfidenzintervall der Absagenquote zur Fehlerwahrscheinlichkeit von 5%.

(d) Es gibt zwei mögliche Ursachen für die Schwankungen: (i) Schwankungen durch die endliche Zahl der Studenten, (ii) Schwankungen der wahrgenommenen Attraktivität des Studiengangs von seiten der Studenten von einem Jahr zum nächsten. Welche Schwankungsursache ist hier größer? Nehmen Sie dazu Unabhängigkeit der beiden Ursachen an und vergleichen Sie die in (a) geschätzte Gesamtvarianz mit der erwartete Varianz bei konstanter Attraktivität. In letzterem Fall gilt für jeden Studenten die gleiche Absagewahrscheinlichkeit, welche Sie wieder mit 45% annehmen können.

Hinweis: Die Zahl der absagenden Studenten ist binomialverteilt.

(e) Um den Studiengang trotz der Absagen zu füllen, werden an die Studenten 350 Einladungen für die 200 Plätze verschickt. Wie groß ist das Risiko, dass sich mehr als 200 Studenten einschreiben und damit der Studiengang überfüllt ist? Nehmen Sie nun an, dass die Schwankungen von Jahr zu Jahr überwiegen, so dass der erwartete *Anteilswert* der Absagen, nicht aber der Erwartungswert der *Anzahl* der Absagen unabhängig von der Zahl der angenommenen Studenten ist.

Aufgabe T.4.6 – Zugachsen, WS 2007/08

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:

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

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

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

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

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:

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

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

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

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

Aufgabe T.4.7 – Ökologisch, WS 2007/08

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?

Aufgabe T.4.8 – Wege mit dem ÖPNV, WS 2006/07

Die Zahl der Wege pro mobiler Person und Tag, welche mit dem ÖPNV zurückgelegt werden, ist in Dresden durch folgende Zeitreihe gegeben (Verkehrserhebung SrV):

Jahr	1991	1994	1998	2003
Mittl. ÖPNV-Wegezahl	0.69	0.71	0.63	0.62

(a) Führen Sie eine lineare Regression durch (Angabe der Zwischenergebnisse ist nötig!) und prognostizieren Sie anhand dieser die Zahl der ÖPNV-Wege pro Tag für das Jahr 2008. Verwenden Sie folgende, bereits berechnete Zwischengrößen (x =Jahr, y = Wegezahl):

$$\bar{x} = 1996.5, \quad \bar{y} = 0.6625, \quad s_x^2 = 20.25, \tag{12}$$

(b) Testen Sie, ob man bei der Fehlerwahrscheinlichkeit von 5% die Aussage "die mittlere Wegezahl ist konstant oder steigt" widerlegen und damit auf eine sinkende Wegezahl schließen kann. Testen Sie dazu in geeigneter Weise den Anstiegsparameter b der linearen Regression. Nehmen Sie dazu für den Parameter $\hat{b}$ des aus Daten gewonnenen Schätzers $\hat{y} = \hat{a} + \hat{b}x$ der linearen Regression den Wert $\hat{b} = -0.007346$ an.

(c) Geben Sie ein Konfidenzintervall bei 10% Fehlerwahrscheinlichkeit für die 2008 erwartete Zahl der ÖPNV-Wege pro Tag an. Gehen Sie davon aus, dass der Schätzer der Regressionsfunktion einer Student-Verteilung mit $n-2$ Freiheitsgraden gehorcht und eine Varianz

$$\sigma_{\hat{y}}^2 = \frac{\hat{\sigma}_R^2}{\sqrt{n}} \sqrt{1 + \frac{(x-\bar{x})^2}{s_x^2}} \tag{13}$$

mit der geschätzten Residualvarianz $\hat{\sigma}_R^2 = 0.000752$ aufweist.

(d) Gegeben ist nun die Zeitreihe mit mehr Punkten in der Vergangenheit:

Jahr	1972	1977	1982	1987	1991	1994	1998	2003
Mittl. Wegezahl	0.87	1.08	0.89	0.93	0.69	0.71	0.63	0.62

In welcher Hinsicht ist eine lineare Regression hier problematisch? Geben Sie einen plausiblen sachlichen Grund an.

Aufgabe T.4.9 – Belastungstest, SS 2006

Bei einem zerstörenden Belastungstest der Flügel eines neuen Verkehrsflugzeugs muss die Stichprobe aus Kostengründen sehr klein bleiben. Eine Stichprobe vom Umfang 4 ergab folgende Bruchlast-Werte in Einheiten der Grenzlast (maximal mögliche Belastung während des normalen Flugbetriebs):

Stichproben-Nr	Belastung (Vielfaches der Grenzlast)
1	1.45
2	1.56
3	1.48
4	1.49

(a) Testen Sie mit dem Kolmogorov-Smirnov-Test bei einer Fehlerwahrscheinlichkeit von 5%, ob die Nullhypothese einer Gaußverteilung mit Erwartungswert $\mu = \bar{x} = 1.495$ und Varianz $\sigma^2 = s_x^2 = 0.0021\overline{66}$ widerlegt werden kann.

(b) Zur Zulassung wird ein Mindestwert der mittleren Bruchlast vom 1.5-fachem der Grenzlast gefordert. Außerdem darf die Standardabweichung nicht größer als das 0.05-fache der Grenzlast sein. Testen Sie zunächst mit dem t-Test bei einer Fehlerwahrscheinlichkeit von 5%, ob die Nullhypothese “Die mittlere Bruchlast ist größer oder gleich dem 1.5-fachem der Grenzlast” widerlegt werden kann. Verwenden Sie dabei die Resultate $\bar{x} = 1.495$ und $s_x^2 = 0.0021\overline{66}$.

(c) Testen Sie nun mit dem χ^2-Varianztest die Nullhypothese “Die Standardabweichung ist kleiner als das 0.05-fache der Grenzlast”.

(d) Bei sicherheitskritischen Fragestellungen sind die bei (b) und (c) formulierten Nullhypothesen problematisch. Diskutieren Sie dies, in dem Sie prüfen, ob die Nullhypothesen “Die mittlere Bruchlast ist *kleiner* als das 1.5-fache der Grenzlast” sowie “Die Standardabweichung ist *größer* als das 0.05-fache der Grenzlast” widerlegt werden können.

Aufgabe T.4.10 – Modal Split, WS 2005/06

Einer Befragung nach der Verkehrsmittelwahl in Dresden ergab 2002 für den Arbeitsweg folgende Ergebnisse: 43% PKW Fahrer, 3% PKW Mitfahrer, 1 % Motorrad, Moped o. ä., 20% Straßenbahn, 12% Bus, 3% S-Bahn, 13% Fahrrad und nur 5% zu Fuß ohne Benutzung anderer Verkehrsmittel.

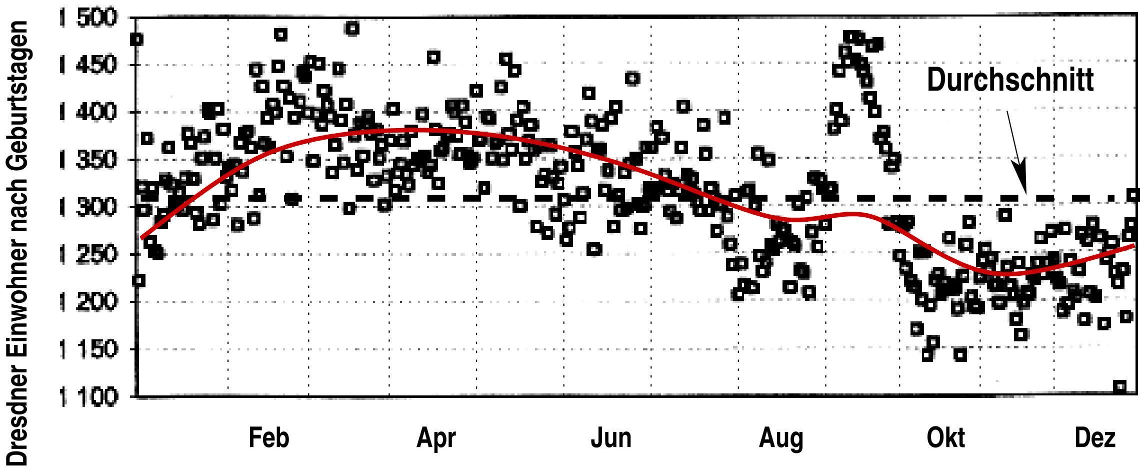
(a) Zeichnen Sie ein Tortendiagramm dieses Sachverhalts.

(b) Der Stichprobenfehler soll für alle Anteile bei einem Signifikanzniveau von 5% kleiner als 2% sein. Wie groß muss man den Stichprobenumfang mindestens wählen?

(c) Gegenüber dem Vorjahr stieg der Anteil der Fahrradbenutzer von 11% auf 13%. Beide Stichproben hatten den Umfang 2000. Sind die Unterschiede bei einer Fehlerwahrscheinlichkeit von 1% signifikant? Verwenden Sie den Differenztest für zwei unabhängige Stichproben!

Aufgabe T.4.11 – Geburtenspitzen in Dresden, WS 2005/06

Der Statistikstelle der Landeshauptstadt Dresden entnimmt man folgende Verteilung der 481 429 Einwohner Dresdens auf die einzelnen 365 Geburtstage (29. Februar nicht berücksichtigt):



(a) Die durchgezogene Kurve gibt eine Glättung der Daten an, welche man z.B. mit dem gleitenden arithmetischen Mittel gewinnen kann.

(i) Wie groß muss man den Glättungsparameter des gleitenden Mittels wählen, damit die Geburtstags-Spitze im September nahezu "weggemittelt" wird, die jahreszeitliche Verteilung "mehr Geburtstage im Frühjahr als im Herbst" aber korrekt wiedergegeben wird? (keine Rechnung, Angabe der Größenordnung genügt!)

(ii) Was muss man zusätzlich bei der Glättung beachten? (bedenken Sie, dass der Jahresbeginn am 1. Januar willkürlich ist!)

(b) Weicht die Geburtstagsverteilung bei einer Fehlerwahrscheinlichkeit von 1 % signifikant von einer Gleichverteilung ab? Berücksichtigen Sie dabei, dass in der Grafik die Fläche zwischen dem Mittelwert und der geglätteten Kurve von Mitte Januar bis Juli (und auch die entsprechende Nettofläche zwischen ungeglätteten Daten und Mittelwert) etwa sechs der von den punktierten Linien begrenzten "Kästchen" entspricht.

Hinweis: Kolmogorow-Smirnov-Anpassungstest.

(c) Kann man die Geburtstags-Spitze im September durch eine statistische Schwankung erklären, wenn man den Kolmogorow-Smirnov-Test wie oben zugrundelegt? Falls nein, hat man damit statistisch eine besonders "fruchtbaren" Empfängniszeitraum im Dezember bewiesen?

(d) In Österreich wurden bei der Volkszählung im Jahre 2001 am ersten Januar 37 700 Geburtstage erfasst, im Mittel jedoch nur 22 082. In Dresden gab es am 1. Januar ebenfalls erstaunlich viele Geburtstage (linkester Datenpunkt der Grafik). Was könnte die Ursache dieser hochsignifikanten Abweichungen sein? (jede plausible Erklärung ergibt volle Punktzahl!)

Aufgabe T.4.12 – Zugriffshäufigkeit auf einer Website, SS 2005

Von einer Webseite über Verkehrssimulationen liegen folgende wöchentlichen und täglichen Zugriffs-Statistiken vor, welche angeben, wie häufig die Seite im entsprechenden Zeitraum angeklickt wurde:

Woche	Zugriffszahl
18	1200
19	2135
20	1778
21	1334
22	2299
23	1554
24	2502
25	3199
26	2035
27	2567
28	2246

Wochentag Woche 28	Zugriffszahl
Mo	148
Di	231
Mi	264
Do	271
Fr	343
Sa	501
So	488

(a) Berechnen Sie bei einer Fehlerwahrscheinlichkeit von 5% das Konfidenzintervall für die mittlere wöchentliche Zugriffszahl der Wochen 18 bis 24.

(b) Am Ende der Woche 24 wurde im Rahmen eines Artikels von einem bekannten Journal aus ein Link auf die Webseite erstellt. Der Webmaster der Seite möchte nun wissen, ob dies die Zugriffszahlen signifikant gesteigert hat. Berechnen sie dazu das Konfidenzintervall für die Wochen 25-28 (bei einer Fehlerwahrscheinlichkeit von 5%) und vergleichen Sie die beiden Intervalle. Bei welcher Fehlerwahrscheinlichkeit kann man gerade noch eine signifikante Erhöhung feststellen?

Hinweis: Eine Interpolation der Quantilswerte aus den entsprechenden Tabellen ist nicht verlangt. Geben Sie den ersten Quantilswert an, bei der Signifikanz gegeben ist.

(c) Zu einer genaueren Analyse soll nun die tägliche Zugriffsstatistik herangezogen werden. Eine typische Verteilung der Zugriffe auf die Woche zeigt die Tagesstatistik der Woche 28. Was muss man beachten, wenn man anhand von Tagesdaten eine Änderung des Zugriffsverhaltens untersuchen will?

Hinweis: Keine Rechnung notwendig!

Aufgabe T.4.13 – Werbeaktion, WS 2004/05

Um zu untersuchen, ob die Bahn ihre Werbung regionalen Besonderheiten anpassen sollte oder nicht, wertet sie die als Stichprobe erhobenen Rückmeldungen auf die letzte Werbeaktion *"Zeit ist Geld - dank der neuen, schnelleren ICEs wird deshalb Bahnfahren trotz angehobener Preise noch günstiger"* getrennt nach Ost- und Westdeutschland aus:

Rückmeldung	genau!	eher ja	weiß nicht	eher nein	ganz und gar nicht!
Ostdeutschland	4	6	14	16	20
Westdeutschland	10	28	40	42	20

(a) Testen Sie mit dem χ^2-Homogenitätstest bei einer Fehlerwahrscheinlichkeit von $\alpha = 2.5\%$, ob die Rückmeldungen aus Ost und West mit einem homogenen Meinungsbild bezüglich dieser Werbekampagne vereinbar sind.

(b) Nehmen Sie für folgende Teilaufgabe an, dass der Wert der Testvariablen in Teilaufgabe (a) durch $Q = 10.7$ gegeben ist und die Nullhypothese eines homogenen Meinungsbildes nicht verworfen werden kann. Erläutern Sie kurz in je ein oder zwei Sätzen, ob dann folgende Aussagen begründet werden können oder nicht:

(i) Mit mindestens 97.5%-iger Wahrscheinlichkeit ist die Nullhypothese H_0 zutreffend.

(ii) Reduziert man die geforderten Zuverlässigkeit durch Erhöhen der Fehlerwahrscheinlichkeit von 2.5% auf 5%, kann man H_0 verwerfen.

(iii) Über die Wahrscheinlichkeit des Zutreffens von H_0 sind ohne weitere Angaben keine Aussagen möglich. Die Wahrscheinlichkeit dafür, dass H_0 falsch ist, kann durchaus weit über 50% liegen.

(iv) Gelingt es, einen Test zu formulieren, bei dem gegenüber dem Homogenitätstest Nullhypothese und Alternativhypothese $H_A = \overline{H_0}$ vertauscht sind und die Nullhypothese H_A dieses "Inhomogenitätstests" bei einer Fehlerwahrscheinlichkeit β zu verwerfen, so ist H_0 mit einer Wahrscheinlichkeit von mindestens $1 - \beta$ zutreffend.

Aufgabe T.4.14 – Einwohner Dresdens, SS 2004

Einer Statistik Dresdens entnimmt man für die Jahre 2001 und 2002 folgende quartalsweisen Geburtenraten und Zuzugsraten:

Jahr	2001				2002			
Quartal	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Zahl der Lebendgeburten	1100	1010	990	1020	1020	1000	1120	1020
Zahl der Zuzüge	4900	5000	6100	6400	4600	4600	5300	6200

(a) Testen Sie die Geburtenraten und die Zuzüge auf saisonale Abhängigkeiten, indem Sie für die Nullhypothese einer Gleichverteilung auf alle Zeitabschnitte einen χ^2-Test (Fehlerwahrscheinlichkeit 2.5%) durchführen

(b) Kann man bei 2.5% Fehlerwahrscheinlichkeit widerlegen, dass die Zuzüge im 4. Quartal gleichmäßig auf die Jahre 2001 und 2002 verteilt sind?

(c) Bezüglich der Zahl der Lebendgeborenen kann man die Nullhypothese im Aufgabenteil (a) nicht widerlegen. Bedeutet dies, dass die Nullhypothese in mindestens 97.5% der Fälle zutrifft? Begründen Sie Ihre Antwort! *Hinweis:* Unterscheiden Sie Fehler erster und zweiter Art und überprüfen Sie auch, ob *im konkretem Fall* die Nullhypothese erfüllt ist!

Aufgabe T.4.15 – Kraftstoffverbrauch, WS 2003/04

Bei einem Kraftstoff-Verbrauchstest werden von 5 Fahrzeugen des Typs „3-Liter Auto“ die Verbrauchswerte x (in Litern je 100 km) bestimmt:

i	x_i
1	3.05
2	3.35
3	3.20
4	3.00
5	3.00

(a) Berechnen Sie zu einer Fehlerwahrscheinlichkeit von 10% das Konfidenzintervall für den Verbrauch.

(b) Der Fahrzeughersteller gibt den mittleren Kraftstoffverbrauch des „3-Liter Autos“ mit unter 3 Litern auf 100 km an. Kann man ihn bei einer Fehlerwahrscheinlichkeit von 10% der Falschaussage bezichtigen? Geben Sie außerdem die Fehlerwahrscheinlichkeit an, bei der die Aussage gerade noch angenommen wird (es genügt, den nächsten angegebenen Wert der Tabelle zu nehmen).

(c) Das in (a) errechnete Konfidenzintervall enthält den Wert der Nullhypothese des Aufgabenteils (b), dennoch kann man in (b) die Nullhypothese verwerfen. Erläutern Sie diese scheinbar widersprüchlichen Ergebnisse.

(d) Der Tank des “3-Liter Autos” fasst 25 l Kraftstoff. Geben Sie, bei einer Fehlerwahrscheinlichkeit von 10%, das Konfidenzintervall für die Reichweite mit einer Tankfüllung an.

Hinweis: Verwenden Sie das Ergebnis der Teilaufgabe (a). Falls Sie (a) nicht gelöst haben, verwenden Sie das Verbrauchs-Konfidenzintervall [2.97 l/100km, 3.27 l/100km].

(e) Kann man bei einer Fehlerwahrscheinlichkeit von 10% die Annahme verwerfen, dass die Standardabweichung des Verbrauchs über 0.4 l/100 km ist?

Tipp: Es ist ein einseitiger Varianztest durchzuführen.

Aufgabe T.4.16 – Weiterempfehlung, WS 2003/04

Um Informationen über die Mund-zu-Mundwerbung für das Bahnfahren zu erhalten, erhebt eine Bahngesellschaft bei 100 ihrer Kunden Stichproben

(i) über die Weiterempfehlungs*absicht*,

(ii) und über *tatsächliche* Weiterempfehlungen

in Abhängigkeit von der Benutzungshäufigkeit. Folgende Tabellen zeigen die Ergebnisse:

Empfehlungsabsicht	ganz bestimmt	eher ja	vielleicht	eher nicht	ganz bestimmt nicht
Vielfahrer	16	5	4	1	14
Wenigfahrer	-	10	15	20	15

Tatsächlich weiterempfohlen	ja	nein
Vielfahrer	25	15
Wenigfahrer	25	35

(a) Testen Sie mit dem χ^2-Test bei einer Fehlerwahrscheinlichkeit von $\alpha = 1\%$, ob die Empfehlungsabsicht unabhängig von der Benutzungshäufigkeit ist.

(b) Testen Sie nun ($\alpha = 1\%$), ob die tatsächliche Häufigkeit der Weiterempfehlungen von der Benutzungshäufigkeit abhängt.

(c) Ermitteln Sie aus der Tabelle der χ^2-Quantile die zwei Fehlerwahrscheinlichkeiten, bei der im Aufgabenteil (b) die Abhängigkeit (i) gerade noch, (ii) gerade nicht mehr, gezeigt werden kann. Falls Sie Teil (b) nicht gelöst haben, nehmen Sie als Realisierung der χ^2-verteilten Testvariablen den Wert 4.17 an.

(d) Warum ist es in dieser Untersuchung wichtig, bei der Erhebung der Empfehlungs*absicht* diese zu differenzieren, also mehr als zwei Stufen zuzulassen? Begründen Sie ihre Aussage, indem Sie in obiger Tabelle die Stufen "ganz bestimmt", "eher ja" und "vielleicht" zur Aussage "ja" sowie die Stufen "eher nicht" und "ganz bestimmt nicht" zur Aussage "nein" zusammenfassen und den Test wiederholen. Können Sie das Ergebnis anschaulich begründen? (*Hinweis:* Welche Information ginge bei der Zusammenfassung verloren?)

Aufgabe T.4.17 – Verkehrsmittelwahl Konfidenzinterv., SS 2003

Um bei Fahrten innerhalb Dresdens die Aufteilung auf die verschiedenen Verkehrsmittel (Modal-Split) zu ermitteln, wird eine Zufalls-Stichprobe vom Umfang 1000 mit folgendem Ergebnis durchgeführt:

Verkehrsmittel	zu Fuß	Rad	Bahn/Bus	Kfz
Zahl der Fahrten	220	110	220	450

(a) Geben Sie Konfidenzintervalle (Irrtumswahrscheinlichkeit: 5%) für die Anteile der vier Verkehrsarten an den durchgeführten Fahrten an.

(b) Wie groß müsste man unter Annahme der obigen Anteilswerte sowie 5% Fehlerwahrscheinlichkeit den Stichprobenumfang wählen, wenn (i) der absolute Fehler des Anteilswertes in jeder der vier Verkehrsarten unter 2% bleiben soll, (ii) der relative Fehler der Anteilswerte unter 5% bleiben soll?

(c) Testen Sie bei einer Fehlerwahrscheinlichkeit von 5% die einseitige Hypothese, dass der Anteil des öffentlichen Nahverkehrs (Bahn/Bus) mindestens bei 24% liegt.

(d) Es ist nun eine Kampagne zum Umstieg von Kfz auf die anderen Verkehrsmittel geplant, deren Erfolg nach Beendigung der Kampagne mit einer weiteren Stichprobe überprüft werden soll. Schlagen Sie eine statistische Methodik (ohne Rechnung!) vor, mit der man eine "signifikante" Veränderung des Modal-Split nachweisen könnte (*Hinweis:* Sie kennen mindestens drei Möglichkeiten; jede davon ergibt volle Punktzahl!).

Aufgabe T.4.18 – Verkehrsmittelwahl von Studenten, SS 2003

Anhand von Stichproben soll überprüft werden, ob das bevorzugte Verkehrsmittel von Studenten bei den Fahrten zur bzw. von der Uni vom gewählten Studienfach abhängt. Folgende Tabelle zeigt das Ergebnis:

Fachrichtung	zu Fuß	Rad	Bahn/Bus	Kfz
Verkehrswirtschaft	18	30	30	22
Physik	8	18	10	14
BWL	34	50	50	66

(a) Zeigen Sie zunächst, dass das bevorzugte Verkehrsmittel empirisch vom Studienfach abhängt.

Hinweis: Es genügt, die Abhängigkeit an einem Beispiel zu zeigen

(b) Prüfen Sie nun auf Abhängigkeit mit dem nichtparametrischen χ^2-Test bei einer Irrtumswahrscheinlichkeit von 5%.

(c) Mit welcher Fehlerwahrscheinlichkeit kann man die Annahme einer Unabhängigkeit gerade noch widerlegen? Falls Sie Teil (b) nicht gelöst haben, nehmen Sie als Realisierung der χ^2-verteilten Testvariablen den Wert $6,19$ an.

Hinweis: Es genügt, wenn Sie aus der Tabelle das nächstliegende aufgeführte Quantil heraussuchen.

Aufgabe T.4.19 – Fahrpreissystem, WS 2002/03

Um die Akzeptanz des neuen Fahrpreissystems der Eisenbahn zu überprüfen, werden per Zufalls-Stichprobe 1000 Fragebögen verteilt und ausgewertet. Das Ergebnis ist:

Nutzerklasse	Neues System ist besser	Neues System ist schlechter	keine Meinung
Bahnbenutzer, vorwiegend Nahverkehr	20	270	10
Bahnbenutzer, vorwiegend Fernverkehr	200	60	40
Kfz-Fahrer und Sonstige	100	100	200

(a) Wieviel Prozent der Befragten finden das neue bzw. das alte Preissystem besser?

(b) Geben Sie Konfidenzintervalle ($\alpha = 5\%$) für den Anteil derjenigen an, die das neue System schlechter finden :

  (i) unter allen Befragten,

  (ii) unter den Bahnbenutzern, die vorwiegend den Nahverkehr benutzen ,

  (iii) unter den Bahnbenutzern, die vorwiegend den Fernverkehr benutzen.

(c) Kann man die Aussage der Bundesbahn: "Höchstens 40% finden das neue System schlechter" anhand des Umfrageergebnisses mit einem statistischen Test bei einer Fehlerwahrscheinlichkeit von 5% widerlegen? Wie groß ist die minimale Fehlerwahrscheinlichkeit ("Grenz-Fehlerwahrscheinlichkeit"), die eine Widerlegung dieser Aussage erlauben würde?

(d) Die Bahn will nun eine genauere Zufalls-Stichprobe durchführen, bei der der Anteil der Gegner des neuen Preissystems bei einer Fehlerwahrscheinlichkeit von 5% auf 2% genau bestimmt werden soll. Es werden Anteilswerte von etwa 40% erwartet. Wie groß muss der Stichprobenumfang mindestens sein?

(e) Können Sie sich eine Stichproben-Methodik vorstellen, welche bei gleichem Stichprobenumfang und gleicher Fehlerwahrscheinlichkeit ein genaueres Ergebnis liefert?

Aufgabe T.4.20 – Motorprüfstand, SS 2002

Tests auf dem Motorenprüfstand eines Kfz-Herstellers ergaben für eine Stichprobe von 10 Motoren folgende Leistungen und Benzinverbräuche (hochgerechnet auf Liter pro 100 km):

Motor-Nummer	Leistung (KW)	Benzinverbrauch (l pro 100 km)
1	76	7
2	83	7.2
3	80	7.1
4	78	6.5
5	78	6.6
6	72	6.8
7	73	7.4
8	77	7.1
9	85	7.0
10	78	7.1

(a) Nach Vorgabe der Qualitätskontrolle soll die mittlere Motorleistung mindestens 80 KW betragen. Der Test gilt als "bestanden", wenn die Vorgabe unter einer Irrtumswahrscheinlichkeit von 5% nicht widerlegt werden kann. Führen Sie den Test unter Annahme einer Gaußverteilung durch.

(b) Nehmen Sie nun an, dass die Motorleistungen gaußverteilt sind mit (tatsächlichem) Mittelwert $\mu = 82$ KW und Varianz $\sigma^2 = (4 \text{ KW})^2$. Mit welcher Wahrscheinlichkeit liegt die Motorleistung (i) unter 80 KW, (ii) zwischen 80 und 86 KW?

(c) Wie könnte man die Annahme einer Unabhängigkeit zwischen den Motorleistungen und den Benzinverbräuchen testen? (Angabe des Tests genügt, keine Rechnung verlangt!)

(d) Die Benzinverbräuche seien nun unabhängig von der Motorleistung sowie gaußverteilt mit Erwartungswert 7 l/(100 km) und Varianz $(0.5 \text{ l}/(100 \text{ km}))^2$. Die Verteilung der Motorleistung sei wire in Aufgabenteil (c). Mit welcher Wahrscheinlichkeit hat ein willkürlich herausgegriffener Motor

(i) Sowohl eine Leistung ≥ 80 KW als auch einem Benzinverbrauch ≤ 7 l/100km?

(ii) Zumindest entweder ≥ 80 KW oder einen Verbrauch ≤ 7 l/100km?

Aufgabe T.4.21 – Verspätungen, WS 2002/03

Um einen Überblick über die Verspätungen auf der Bahnverbindung Dresden-Berlin zu erhalten, werden die Ankunftszeiten der letzten 200 Züge dieser Verbindung festgestellt:

Zeitkategorie	Zahl der Züge
Mehr als 5 min zu früh	5
0-5 min zu früh	15
0-5 min zu spät	94
5-10 min zu spät	50
10-20 min zu spät	17
20-30 min zu spät	10
Mehr als 30 min zu spät	9

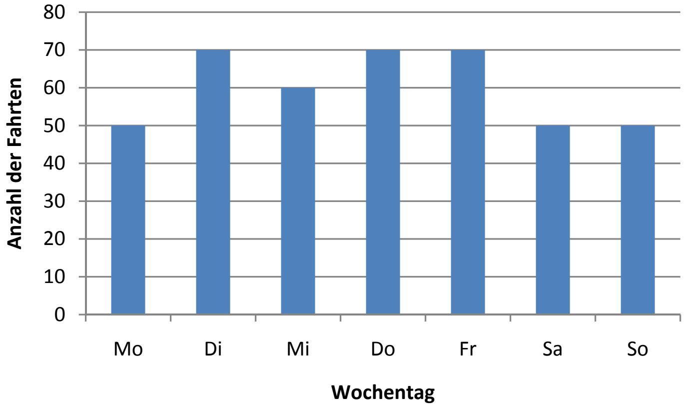
(a) Bestimmen Sie zunächst aus der Tabelle den Mittelwert $\bar{X}$ und die Stichprobenvarianz s^2. Nehmen Sie dabei an, dass die Züge der ersten Zeitkategorie im Mittel 10 min zu früh ankommen und die der letzten Kategorie im Mittel 35 min zu spät.

(b) Überprüfen Sie mit einer Irrtumswahrscheinlichkeit von 5% die obige Verteilung auf eine Gaußverteilung. Falls Sie Teil (a) nicht gerechnet haben, rechnen Sie mit der Schätzung $\mu = 6.713$ min und $\sigma^2 = 80.82$ min^2.

(c) Bestimmen und zeichnen Sie die empirische Dichtefunktion. Nehmen Sie dabei als untere Grenze der ersten Klasse 15 min zu früh und als obere Grenze der letzten Klasse 50 min zu spät an. Sieht man in der Grafik einen möglichen Grund für das Ergebnis in Teil (b)?

Aufgabe T.4.22 – Taxi (WS 10/11)

Ein kleineres Taxiunternehmen hat für die vergangene Woche die Anzahl der Fahrten, die an jedem der sieben Tage durchgeführt wurden, zusammengetragen:



Wochentag	Fahrten
Mo	50
Di	70
Mi	60
Do	70
Fr	70
Sa	50
So	50

(a) Schätzen Sie aus der gegebenen Stichprobe den Mittelwert und die Varianz der täglichen Fahrtenzahl.

(b) Für das Fortbestehen des Taxiunternehmens ist es entscheidend, dass im Langzeitmittel täglich mindestens 50 Fahrten durchgeführt werden. Würde man rechtzeitig wissen, dass dies nicht gegeben ist, könnte man den drohenden Konkurs mit Umstrukturierungs- und Einsparungsmaßnahmen jedoch noch abwenden.

Welche der folgenden drei Nullhypothesen

$$H_0 : \mu < 50 \qquad H_0 : \mu = 50 \qquad H_0 : \mu > 50$$

ist für einen entsprechenden statistischen Test heranzuziehen? Benennen Sie für Ihre Auswahl die Konsequenzen eines Fehlers 1. bzw. 2. Art und begünden Sie kurz, warum der Fehler 1. Art schwerwiegender ist als der Fehler 2. Art.

(c) Kann man die Nullhypothese $H_0 : \mu < 50$ mit einer maximalen Irrtumswahrscheinlichkeit von 5% verwerfen?

(d) Angenommen, die Nullhypothese in (c) kann verworfen werden. Was lässt sich daraus schlussfolgern? (Ein Antwortsatz genügt.)

(e) Überprüfen Sie die Nullhypothese

$$H_0 : \text{„Die Fahrten verteilen sich gleichmäßig auf die sieben Wochentage“}.$$

Verwenden Sie den χ^2-Anpassungstest mit einer maximalen Irrtumswahrscheinlichkeit von 5%. (Hilfestellung: Die vorliegende Stichprobe umfasst $n = 420$ Fahrten!)

(f) Angenommen, die Nullhypothese in (e) kann nicht abgelehnt werden. Was lässt sich daraus schlussfolgern? (Ein Antwortsatz genügt.)

Aufgabe T.4.23 – Tag der Sachsen (SS 2011)

Zum Tag der Sachsen wurden $n = 200$ von auswärts kommende Besucher dazu befragt, wie weit und mit welchem Verkehrsmittel sie angereist sind. Die absoluten Häufigkeiten h_{ij} sind in folgender Kreuztabelle aufgetragen:

		Verkehrsmittel		
		MIV	ÖV	
Entfernungsklasse	h_{ij}	$j = 1$	$j = 2$	$\sum_j$
kleiner 50 km	$i = 1$	80	33	113
zwischen 50 und 100 km	$i = 2$	37	18	55
über 100 km	$i = 3$	29	3	32
	$\sum_i$	146	54	200

(a) Wie hoch ist der ÖV-Anteil f unter allen Befragten?

(b) Geben Sie das Konfidenzintervall des ÖV-Anteils θ auf einem Signifikanzniveau von $\alpha = 5\%$ an.

(c) Obwohl das Konfidenzintervall in (b) Werte größer 33% beinhaltet, kann die Nullhypothese

$$H_0: \text{„Der ÖV-Anteil ist größer als 33\%."}$$

mit $\alpha = 5\%$ abgelehnt werden. Warum ist das kein Widerspruch?

(d) Testen Sie mit dem χ^2-Unabhängigkeitstest und einem Signifikanzniveau von $\alpha = 5\%$, ob die Verkehrsmittelwahl von der Entfernungsklasse unabhängig ist.

(e) Angenommen, die Nullhypothese in (d) kann verworfen werden. Erklären Sie in einem Antwortsatz, was sich daraus schlussfolgern lässt.

(f) Testen Sie mit dem χ^2-Anpassungstest und einem Signifikanzniveau von $\alpha = 5\%$, ob die Reiseweite (ungeachtet des gewählten Verkehrsmittels) exponentialverteilt ist. Die mittlere Reiseweite sei mit $1/\lambda = 60$ km gegeben. Nehmen Sie als Untergrenze der ersten Entfernungsklasse 0 und als Obergrenze der letzten Entfernungsklasse „unendlich" an. Hinweis: Die Verteilungsfunktion der Exponentialverteilung lautet

$$F_E(x) = 1 - e^{-\lambda x}.$$

(g) Angenommen, die Nullhypothese in (f) kann nicht abgelehnt werden. Erklären Sie in einem Antwortsatz, was sich daraus schlussfolgern lässt.