

# Wahrscheinlichkeitsrechnung und **Statistik** für Informatik

Grossmann / Neuwirth / Cenker

28. Oktober 2003

<b>Name</b>	
<b>Matrikelnummer</b>	
<b>Studienkennzahl</b>	

Bsp	1	2	3	4	5	6	Ges	Note
Punkte								

Liebe Kandidatin!  
Lieber Kandidat!

Sie finden auf den folgenden Seiten 6 Beispiele (10 Punkte pro Beispiel). Bitte beachten Sie, dass nur diese (zusammengehefteten) Blätter abgegeben werden können. Zusatzblätter werden nicht angenommen!

Tragen Sie oben Ihren Namen, Ihre Matrikelnummer und die Studienkennzahl der Studienrichtung ein, für die diese Prüfung angerechnet werden soll. Sind diese Daten falsch bzw. bei der Prüfungsanmeldung falsch eingegeben worden, kann kein Zeugnis ausgestellt werden.

Viel Glück!

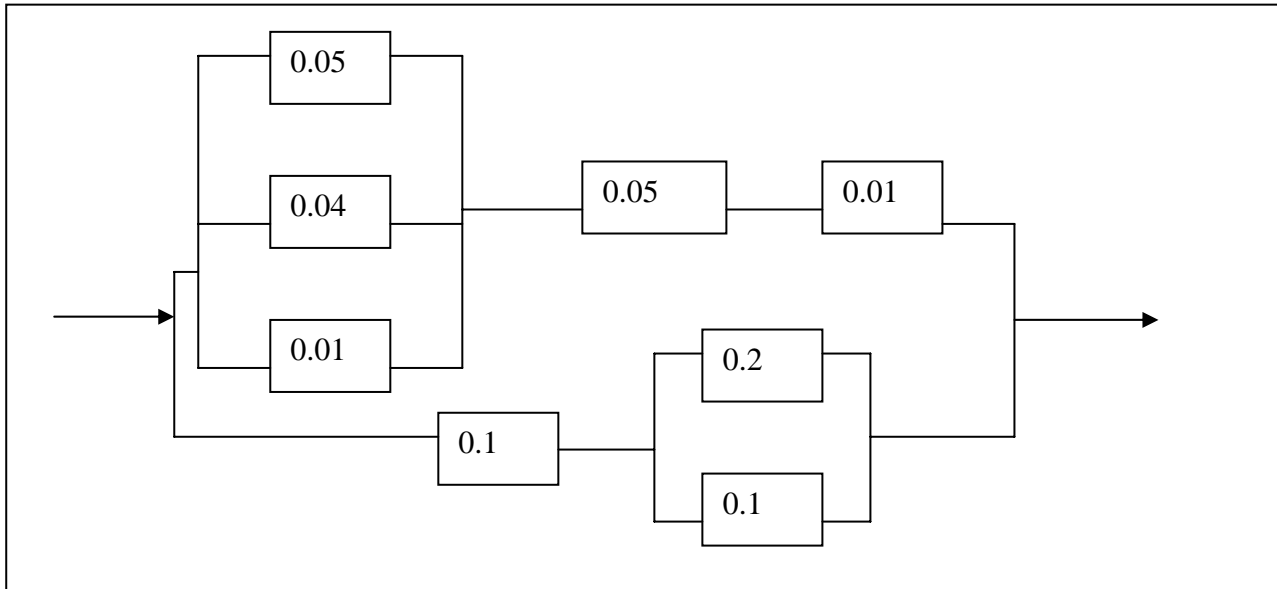
1. Welche grafischen Darstellungsformen für Häufigkeiten kennen Sie?

Erklären sie insbesondere auch, wofür man Kerndichteschätzer und Mosaic-Plots verwenden kann.

**2.** Sie werfen einen (fairen) Würfel 20 Mal. Wie groß ist die Wahrscheinlichkeit, eine Augensumme von höchstens 22 oder mindestens 118 zu erhalten? Wir fragen also nach der Wahrscheinlichkeit „extremer“ Ergebnisse, d.h., besonders hoher oder besonders niedriger Augensummen.

3. Gegeben ist unten skizziertes Netzwerk.

- Berechnen Sie dessen Zuverlässigkeit (Ansatz, Berechnung, Ergebnis)! In den Komponenten ist jeweils die Ausfallswahrscheinlichkeit angegeben.
- Wenn jede Komponente eine Lebensdauerverteilung  $G(x)$  hat, wie lautet dann die Gesamtlebensdauerverteilung dieses Netzwerkes?



4. Folgende Tabellen enthalten eine statistische Auswertung der Wirkung eines Blutdruck senkenden Medikaments. Als Signifikanzniveau wurde 95% gewählt. Um den Test zu vereinfachen, wurde nur der diastolische Wert betrachtet. Die erste Spalte enthält den Wert ohne Medikation, die zweite Spalte bei laufender Medikation. Die Daten eines Patienten sind zeilenweise eingetragen.

Welche Annahmen wurden gemacht? Worauf muss geachtet werden?

Wie lauten sinnvolle Nullhypothesen bzw. Alternativen, wenn Sie die unten angeführte statistische Auswertung verwenden wollen?

Was ergibt sich als Aussage dieser statistischen Auswertung? Warum?

Ist dieser Test sinnvoll?

Vorher	Nachher	Differenz
210	160	-50
180	150	-30
189	160	-29
195	150	-45
160	160	0
150	160	10
155	120	-35
165	180	15
140	160	20

Zweistichproben t-Test bei abhängigen Stichproben (Paarvergleichstest)

	<i>Variable 1</i>	<i>Variable 2</i>
Mittelwert	171,555556	155,555556
Varianz	539,277778	252,777778
Beobachtungen	9	9
Pearson Korrelation	0,07184973	
Hypothetische Differenz der Mittelwerte	0	
Freiheitsgrade (df)	8	
t-Statistik	1,76570854	
P(T<=t) einseitig	0,05771931	
Kritischer t-Wert bei einseitigem t-Test	1,85954832	
P(T<=t) zweiseitig	0,11543863	
Kritischer t-Wert bei zweiseitigem t-Test	2,30600563	

5. Ein Kellner arbeitet sowohl tagsüber als auch abends in einem Lokal und führt eine Statistik über das Geschlecht der Besucher. Im Laufe einer Woche ergeben sich die folgenden Daten:

	Tagsüber	Abends	Gesamt
Weiblich	35	52	87
Männlich	33	124	157
Gesamt	68	176	244

- Welche grafischen Darstellung der Daten würden Sie empfehlen (Absolutwerte, verschiedene Prozentwerte, ...)?
- Liegt auf Grund der Daten genügend Evidenz vor, dass abends mehr Männer in das Lokal kommen? (Signifikanzniveau  $\alpha = 0,05$ ; kritischer Wert = 3,84)
- Berechnen Sie die Odds-Ratio und interpretieren Sie diese. Welchen Vorteil hat die Odds-Ratio gegenüber der Differenz der Anteile?

6. Neben der Statistik der Besucher führt der Kellner auch noch Aufzeichnungen über die Höhe des Rechnungsbetrages, gegliedert nach Tageszeit und nach Geschlecht. Die folgende Tabelle gibt eine Zusammenfassung der Daten:

#### Deskriptive Statistiken

Abhängige Variable: Rechnungsbetrag

Tageszeit	Geschlecht des Zahlers	Mittelwert	Standardabweichung	N
tagsüber	maennlich	18,0512	7,9497	33
	weiblich	16,3391	7,5008	35
	Gesamt	17,1700	7,7122	68
abends	maennlich	21,5923	9,5363	124
	weiblich	19,0869	8,2857	52
	Gesamt	20,8521	9,2326	176
Gesamt	maennlich	20,8480	9,3146	157
	weiblich	17,9815	8,0494	87
	Gesamt	19,8259	8,9735	244

Um Unterschiede zwischen den einzelnen Rechnungsbeträgen zu den verschiedenen Zeiten zu untersuchen, führt er eine zweifache Varianzanalyse durch. Die Ergebnisse sind in folgender Tabelle enthalten:

	Quadratsummen	Freiheitsgrade	Mittlere Quadratsumme	F	Signifikanz
Gesamtmodell	944,76	3	?	4,06	0,008
ZEIT	458,98	1	458,98	?	0,016
GESCHL	206,42	1	?	2,66	0,104
ZEIT * GESCHL	7,30	1	7,30	0,09	0,759
Fehler	18622,33	240	77,59		
Gesamt	19567,09	243			

Man beantworte folgende Fragen:

- Ergänze die Fragezeichen in der Tabelle.
- Kann man daraus schließen, dass der Rechnungsbetrag tagsüber und abends unterschiedlich ist? (Begründung)
- Kann man daraus schließen, dass der Rechnungsbetrag bei Männern und Frauen unterschiedlich ist? (Begründung)?
- Kann man daraus schließen, dass es beim Rechnungsbetrag einen Effekt gibt, der sich aus der Kombination von Tageszeit und Geschlecht zusammen setzt? (Begründung)