

Wahrscheinlichkeitsrechnung und **Statistik** für Informatik

Grossmann / Neuwirth / Cenker

16. Dezember 2003

Name	
Matrikelnummer	
Studienkennzahl	

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.

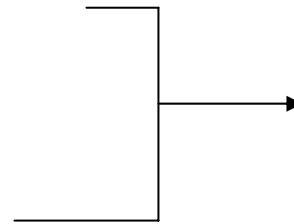
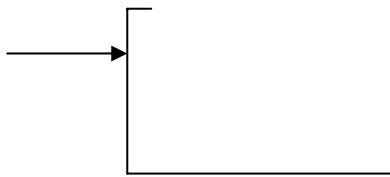
Viel Glück!

1. Berechnen Sie die Wahrscheinlichkeit, bei 10-maligem Wurf mit einem fairen Würfel mindestens 2-mal einen Sechser zu bekommen!

2. Erklären Sie den Unterschied zwischen Lagemaßen, Mittelwerten und Streuungsmaßen. Welche Invarianzeigenschaften haben diese Maßzahlen? Welches mit Invarianz verwandte Konzept wird noch zur Charakterisierung von statistischen Maßzahlen verwendet?

3. Gegeben ist unten skizziertes Netzwerk.

- a) Berechnen Sie dessen Zuverlässigkeit (Ansatz, Berechnung, Ergebnis)! Alle Komponenten haben dieselbe Ausfallswahrscheinlichkeit p . Rechnen Sie zuerst allgemein, dann mit $p=0.2$.
- b) Wenn jede Komponente eine Lebensdauerverteilung $G(x)$ hat, wie lautet dann die Gesamtlebensdauerverteilung dieses Netzwerkes?



4. Folgende Tabellen enthalten eine statistische Auswertung der Erträge von Getreidefeldern vor und nach Düngung mit einem Spezialdünger (Werte in t). Als Signifikanzniveau wurde 95% gewählt.

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
102	125	23
234	265	31
129	135	6
123	155	32
345	320	-25
678	657	-21
234	250	16
234	240	6
123	123	0
345	341	-4

Zweistichproben t-Test bei abhängigen Stichproben (Paarvergleichstest)

	<i>Vorher</i>	<i>Nachher</i>
Mittelwert	254,7	261,1
Varianz	30064,9	25769,6556
Beobachtungen	10	10
Hypothetische Differenz der Mittelwerte	0	
Freiheitsgrade (df)	9	
t-Statistik	-1,02417845	
P(T<=t) einseitig	0,16624503	
Kritischer t-Wert bei einseitigem t-Test	1,83311386	
P(T<=t) zweiseitig	0,33249007	
Kritischer t-Wert bei zweiseitigem t-Test	2,26215889	

5. Gegeben ist folgende unvollständige Varianzanalysetabelle

	Quadrat-summen	Freiheits-grade	Mittl. Qua-dratsumme	F-Wert	p-Wert
zwischen d. Gruppen	190				$7,7 \cdot 10^{-11}$
innerhalb d. Gruppen		41			
Total	241	44			

- Vervollständige die Tabelle und bestimme den Wert der F-Statistik zum Test der Nullhypothese, dass die Mittelwerte der Gruppen gleich sind.
- Liegt auf Grund der Daten genügend Evidenz vor, dass zwischen den Gruppen ein Unterschied besteht (Signifikanzniveau $\alpha = 0,01$)?
- Wie groß ist die Anzahl der Beobachtungen in jeder Gruppe unter der Annahme, dass die Anzahl der Beobachtungen in allen Gruppen gleich ist?
- Welche Möglichkeiten zur Darstellung der Daten würden Sie bei der Varianzanalyse empfehlen?
- Welche deskriptiven Statistiken sollte man jedenfalls bei der Varianzanalyse betrachten?

6. Zwei Gruppen, die aus jeweils 150 Personen bestehen leiden an einer Krankheit. Die Gruppe A wird mit einem neuen Mittel behandelt, die Gruppe B mit dem herkömmlichen Mittel. In Gruppe A werden dabei 105 Personen gesund, in Gruppe B werden 45 Personen gesund.

- a) Man teste mittels des χ^2 -Test die Hypothese, dass das zwischen der Wirkung der Medikamente ein Unterschied besteht. (Vergleichswert $\chi^2_{1; 0.95} = 3,84$)
- b) Bestimme die Odds-Ratio für die Wirkung der beiden Medikamente.
- c) Bestimme ein 95% Konfidenzintervall für die Differenz der Anteile der durch die zwei Behandlungen genesenen Patienten. ($z_{0,975} = 1,96$)
- d) Welche Voraussetzungen müssen für die Berechnungen nach a) und c) gelten?

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

- f) Welche grafischen Darstellung der Daten würden Sie empfehlen (Absolutwerte, verschiedene Prozentwerte, ...)?
- g) Liegt auf Grund der Daten genügend Evidenz vor, dass abends mehr Männer in das Lokal kommen? (Signifikanzniveau $\alpha = 0,05$; kritischer Wert = 3,84)
- h) 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)