

Prof. Dr. G. Tutz
M.Sc. M. Berger
Dipl.-Stat. W. Pößnecker
Institut für Statistik
LMU München

Bitte für die Korrektur freilassen!			
Aufgabe	1	2	Σ
Punkte	23	22	45
erzielt			

Übungs-Klausur-Aufgaben zur Vorlesung “Multivariate Verfahren” im Sommersemester 2015

07. Juli 2015

Hinweise zu den Übungsaufgaben:

- (a) Im Folgenden finden Sie zu Übungszwecken zwei ausgewählte Aufgaben aus alten MV-Klausuren.
- (b) Die Bearbeitungszeit der Bachelor-Klausur beträgt 120 Minuten. Bei der Bachelor-Klausur können 120 Punkte erreicht werden.
- (c) Die Bearbeitungszeit der Klausuren „Grundlagen der Multivariaten Verfahren“ und „Fortgeschrittene multivariate Verfahren“ beträgt jeweils 60 Minuten und es können jeweils 60 Punkte erreicht werden.
- (d) Da die Übungsaufgaben einen Umfang von 45 Punkten haben, sollte Sie unter Klausurbedingungen dementsprechend 45 Minuten für die beiden Übungsaufgaben benötigen.

Hinweise:

- (a) Als Hilfsmittel sind ausschließlich ein nicht programmierbarer Taschenrechner sowie die Formelsammlung zugelassen.
- (b) Nicht-Deutsch-Muttersprachler dürfen zusätzlich ein Wörterbuch verwenden der Form: Muttersprache – deutsch / deutsch – Muttersprache.
- (c) Bei Unterschleif gilt die Klausur als nicht bestanden und es erfolgt eine Meldung an das Prüfungsamt. Bei einem Verdacht auf Unterschleif sind Sie verpflichtet, bei der Aufklärung dieses Verdachts mitzuwirken.
- (d) Verwenden Sie für Ihre Notizen und Lösungen ausschließlich die Ihnen zur Verfügung gestellten Papierbögen.
- (e) Es darf nicht mit Bleistift geschrieben werden.
- (f) Geben Sie bitte am Ende der Klausur alle von Ihnen zur Korrektur vorgesehenen Blätter sowie diese Angabe ab und kennzeichnen Sie jedes abgegebene Blatt mit Namen und Matrikelnummer. Verlassen Sie anschließend aus Rücksicht auf Ihre Kommilitonen den Hörsaal zügig und leise.
- (g) Verlassen Sie den Prüfungsraum erst, nachdem Sie der Aufsicht die Klausur persönlich übergeben haben. Für den Eingang der kompletten Klausur bei der Aufsicht sind Sie selbst verantwortlich.

Ein bekannter Ernährungswissenschaftler behauptet, dass man mit seinem Diät-Programm (A) über 12 Wochen genau 6 kg Körpergewicht und 3 kg Körperfett abnimmt. Um dies zu überprüfen, führt er eine Studie mit 29 Kundinnen durch. Nach 12 Wochen Diät wird bei jeder Kundin gemessen, wie viel Körpergewicht (x_1) und Körperfett (x_2) sie abgenommen hat. Der Wissenschaftler erhält folgende empirische Größen (Mittelwertsvektor und Kovarianzmatrix):

$$\bar{x}_A = \begin{pmatrix} 5.89 \\ 2.42 \end{pmatrix} \quad S_A = \begin{pmatrix} 1.88 & 0.42 \\ & 0.53 \end{pmatrix}$$

Ein zweiter Ernährungswissenschaftler ist der Überzeugung, dass sein Diät-Programm (B) besser ist und führt dieselbe Studie mit 30 Kundinnen durch. Er erhält analog folgende Resultate:

$$\bar{x}_B = \begin{pmatrix} 4.70 \\ 3.27 \end{pmatrix} \quad S_B = \begin{pmatrix} 1.66 & 1.03 \\ & 1.41 \end{pmatrix}$$

Nehmen Sie für die folgenden Aufgaben an, dass die Ergebnisse in beiden Gruppen unabhängige Stichproben aus jeweils normalverteilten Grundgesamtheiten sind und dass die zugrunde liegenden, wahren Kovarianzmatrizen in beiden Gruppen gleich sind.

- (a) Testen Sie mit einem geeigneten multivariaten Test zum Signifikanzniveau von $\alpha=0.05$, ob der erste Ernährungswissenschaftler mit seiner Behauptung richtig liegt.

Quantile der F-Verteilung:

$$F_{0.95}(2, 29) = 3.327, F_{0.99}(2, 29) = 5.420, F_{0.95}(2, 28) = 3.340, F_{0.99}(2, 28) = 5.453$$

$$F_{0.95}(2, 27) = 3.354, F_{0.99}(2, 27) = 5.488, F_{0.95}(2, 26) = 3.369, F_{0.99}(2, 26) = 5.526$$

- (b) Testen Sie mit einem geeigneten multivariaten Test zum Signifikanzniveau von $\alpha=0.05$, ob sich die Ergebnisse der beiden Diät-Programme signifikant unterscheiden.

Quantile der F-Verteilung:

$$F_{0.95}(2, 59) = 3.153, F_{0.99}(2, 59) = 4.984, F_{0.95}(2, 58) = 3.156, F_{0.99}(2, 58) = 4.991$$

$$F_{0.95}(2, 57) = 3.159, F_{0.99}(2, 57) = 4.998, F_{0.95}(2, 56) = 3.162, F_{0.99}(2, 56) = 5.006$$

Eine Privatbank beurteilt die Kreditwürdigkeit ihrer Kunden anhand von drei Gehaltsstufen. Die Klassifikation zur Beurteilung der Kreditwürdigkeit von Kreditnehmern erfolgt in die Klassen kreditwürdig (Klasse 1) und nicht kreditwürdig (Klasse 2). Die Verteilung des trichotomen Merkmals X in den beiden Klassen und die apriori Wahrscheinlichkeiten für die jeweiligen Klassenzugehörigkeiten seien durch folgende Tabelle bestimmt:

	X			apriori-Wahrscheinlichkeit
	Gehalt niedrig	Gehalt mittel	Gehalt hoch	
Klasse1	0.1	0.2	0.7	p
Klasse2	0.8	0.1	0.1	$1-p$

- (a) Bestimmen Sie die ML-Zuordnung für alle drei Gehaltsstufen. Berechnen Sie anschließend für diese Zuordnung die individuellen Fehlerraten ϵ_{21} und ϵ_{13} .
- (b) Berechnen Sie für die Gehaltsstufe "hoch" die Bayes-Zuordnung in Abhängigkeit von p .
- (c) Es werden nun Fehlklassifikationskosten berücksichtigt. Bestimmen Sie die kostenoptimale Bayes-Zuordnung in Abhängigkeit von p für die Gehaltsstufe "niedrig" und den Kosten aus der folgenden Tabelle:

c_{ij}	j=1	j=2
i=1	0	1
i=2	2	0

- (d) Welcher Klasse würden Sie einen Kunden mit Gehaltsstufe "niedrig" unter Benutzung der Bayeszuordnungsregel mit Kosten zuordnen, wenn die apriori Wahrscheinlichkeit durch $p=0.7$ bestimmt ist?