Prof. Dr. Hermann Singer

Modul 31821
Multivariate Verfahren

Kurs 00883
Kurseinheit 1:
Multivariate Statistik
Kurseinheit 2:
Aufgaben und Lösungen

LESEPROBE

Fakultät für
Wirtschafts-
wissenschaft

FernUniversität in Hagen

Kurseinheit 1

Inhaltsverzeichnis

Kurseinheit 2

Inhaltsverzeichnis

Kapitel 3

Tests und Konfidenzintervalle

3.1 Allgemeine Bemerkungen

Im univariaten Fall konnten Hypothesen über skalare Parameter, etwa $H_0 : \mu = \mu_0$, mit Hilfe von Gauß- oder t-Tests überprüft werden. Im multivariaten Fall hat man das Problem, daß mehrere Mittelwerte $\mu_1, ..., \mu_p$ gleichzeitig (simultan) getestet werden müssen. In Kap. 1.5 wurde erwähnt, daß bei der Untersuchung auf korrelative Zusammenhänge mehrere Korrelationen gleichzeitig auf Signifikanz (d.h. $H_0 : \rho_{ij} = 0$ ablehnen) untersucht werden müssen. Setzt man die Nullhypothese aus dem Schnitt mehrerer univariater Hypothesen zusammen, so stellt sich bei einer Abfolge von univariaten Tests heraus, daß das simultane Signifikanzniveau (Fehler 1. Art) $\alpha^* = P(H_0 \text{ ablehnen}|H_0 \text{ richtig})$ größer als das α der Einzeltests werden kann. Eine Adjustierung der Einzeltests führte zu einer recht einfachen Lösung (Bonferroni-Ungleichung), jedoch kann der Test konservativ sein (bestehende Unterschiede werden durch die Testprozedur nicht entdeckt, da das gesamte Signifikanzniveau α^* zu klein ist.)

Daher ist es sinnvoll, multivariate Tests durchzuführen, die das geforderte simultane Signifikanzniveau (Fehler 1. Art) $\alpha^* = P(H_0 \text{ ablehnen}|H_0 \text{ richtig}) = \alpha$ exakt einhalten.

Als *Hypothesen* werden im folgenden Teilräume des Parameter-Raums $\mathbf{\Theta}$ bezeichnet.

Hypothesen

Etwa ist $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ ein einzelner Punkt im u-dimensionalen Raum $\mathbf{\Theta} = \mathbb{R}^u$.

Man kann auch allgemeiner $H_0 = \{\boldsymbol{\theta}|\boldsymbol{\theta} \in \boldsymbol{\Theta}_0\} \subset \boldsymbol{\Theta}$ schreiben, also die Menge der Parameterwerte, die eine bestimmte Bedingung erfüllen, z.B. ein einzelner Punkt $\boldsymbol{\Theta}_0 = \{\boldsymbol{\theta}_0\} = \{[\theta_{01}, ..., \theta_{0u}]'\}$.

Systematische Prinzipien zur Konstruktion von Tests sind das Likelihood-Quotienten- sowie das Union-Intersection-Prinzip. Im ersten Fall wird der Likelihood-Quotient (LQ)

Likelihood-Quotient

$$\lambda = \frac{L(\hat{\boldsymbol{\theta}}_0; \mathbf{X})}{L(\hat{\boldsymbol{\theta}}_1; \mathbf{X})} \tag{3.1}$$

unter der H_0 sowie der H_1 berechnet. Hierbei sind $\hat{\boldsymbol{\theta}}_0, \hat{\boldsymbol{\theta}}_1$ die Parameter-Schätzwerte, bei denen die Likelihood unter den Hypothesen maximal wird (vgl. Abs. 2.4). Daten, die eher für die H_1 sprechen, führen also zu kleinen Werten der LQ-Statistik.

Union-Intersection-Prinzip

Beim Union-Intersection-Prinzip wird die multivariate Nullhypothese als Schnittmenge

$$H_0 = \bigcap_{\mathbf{a}} H_{0\mathbf{a}} \tag{3.2}$$

univariater Hypothesen (Komponenten) geschrieben. Etwa ist H_{0a} : $\mathbf{a}'\boldsymbol{\mu} = \mathbf{a}'\boldsymbol{\mu}_0$ eine solche Hypothese.

Wählt man als $\mathbf{a} = \mathbf{e}_1 = [1, 0]'$ (Einheitsvektor in x-Richtung), so ist $H_{0\mathbf{e}_1} : \mathbf{e}_1'\boldsymbol{\mu} = \mathbf{e}_1'\boldsymbol{\mu}_0$ bzw. $\mu_1 = \mu_{10}$ eine univariate Hypothese für die 1. Komponente. Die Wahl $\mathbf{a} = [1, 1]'$ führt zu einer Linearkombination $\mu_1 + \mu_2 = \mu_{10} + \mu_{20}$. Derartige Linearkombinationen sind oft den Daten besser angepaßt, wenn die Variablen korreliert sind.

Die Nullhypothese wird beibehalten, wenn alle Komponenten $H_{0\mathbf{a}}$ beibehalten werden. Dagegen führt die Ablehnung schon einer Komponenten-Hypothese $H_{0\mathbf{a}}$ zur Ablehnung der H_0. Der Test von $H_{0\mathbf{a}}$ wird mit einer geeigneten univariaten Teststatistik durchgeführt.

Beispiel 3.1 (Mittelwerts-Test, $\boldsymbol{\Sigma}$ bekannt)

Will man die Hypothese $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0 = [6, 6]'$ testen, so kann man die Komponenten $H_{0\mathbf{a}} : \mathbf{a}'\boldsymbol{\mu} = \mathbf{a}'\boldsymbol{\mu}_0 = 6a_1 + 6a_2$ einzeln abprüfen. Der Mittelwert $\bar{\mathbf{x}}$ ist normalverteilt $N(\boldsymbol{\mu}, \boldsymbol{\Sigma}/N)$. Daher gilt für die Projektionen auf den Vektor $\mathbf{a}$: $\mathbf{a}'\bar{\mathbf{x}} \sim N(\mathbf{a}'\boldsymbol{\mu}, \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}/N)$.

Somit ist $Z_{\mathbf{a}} = (\mathbf{a}'\bar{\mathbf{x}} - \mathbf{a}'\boldsymbol{\mu}_0)/\sqrt{\mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}/N}$ standardnormalverteilt. H_0 wird nur beibehalten, wenn alle Einzeltests nicht signifikant sind, also $|Z_{\mathbf{a}}| \leq z(1 - \alpha/2)$. Maximiert man über alle denkbaren $\mathbf{a}$, so muß auch

$\max_{\mathbf{a}} |Z_{\mathbf{a}}| \leq z(1-\alpha/2)$ gelten. Die Maximierung führt direkt zur Teststatistik nach dem Union-Intersection-Prinzip.

■

3.2 Ein-Stichproben-Fall

3.2.1 Test für den Erwartungswert $\boldsymbol{\mu}$ ($\boldsymbol{\Sigma}$ bekannt)

Dies ist die direkte Verallgemeinerung des Gauß-Tests bei univariaten normalverteilten Stichproben.

Bekanntlich ist der Mittelwert $\bar{X}$ der unabhängigen Daten $X_n \sim N(\mu, \sigma^2), n = 1, ..., N$ unter der Nullhypothese $\mu = \mu_0$ auch normalverteilt $N(\mu_0, \sigma^2/N)$. Quadrate von normalverteilten Größen sind χ^2-verteilt. Schreibt man $t^2 = (\bar{X} - \mu_0)^2/(\sigma^2/N) = Z^2$ mit der (unter H_0) standardisierten Variable $Z = (\bar{X} - \mu_0)/\sqrt{\sigma^2/N}$, so kann man die $\chi^2(1)$-Verteilung zum Hypothesentest verwenden.

Analog schreibt man im multivariaten Fall (p-dimensionale Daten):

1. Daten: $\mathbf{x}_n \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}), n = 1, ..., N$
   (unabhängig und identisch verteilt).
2. Hypothesen: $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ gegen $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$.
3. Teststatistik: $T^2 = N(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0) \sim \chi^2(p)$ unter H_0.
4. Kritischer Wert: $\chi^2(1-\alpha, p)$.
5. Testentscheidung: Falls $T^2 > \chi^2(1-\alpha, p)$, H_0 ablehnen.

In der Tat ist $\mathbf{z} = \sqrt{N}\boldsymbol{\Gamma}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)$, $\boldsymbol{\Sigma} = \boldsymbol{\Gamma}\boldsymbol{\Gamma}'$ (vgl. (2.65)) ein normalverteilter Zufallsvektor mit $\text{Var}(\mathbf{z}) = N\boldsymbol{\Gamma}^{-1}(\boldsymbol{\Sigma}/N)(\boldsymbol{\Gamma}^{-1})' = \boldsymbol{\Gamma}^{-1}\boldsymbol{\Gamma}\boldsymbol{\Gamma}'(\boldsymbol{\Gamma}^{-1})' = \mathbf{I}$. Hierbei wurde $\mathbf{I} = \boldsymbol{\Gamma}^{-1}\boldsymbol{\Gamma} = \boldsymbol{\Gamma}'(\boldsymbol{\Gamma}^{-1})'$ und $\text{Var}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0) = \boldsymbol{\Sigma}/N$ eingesetzt.

Daher sind die Komponenten von $\mathbf{z}$ standardnormalverteilt $N(0, 1)$ und es gilt $T^2 = \mathbf{z}'\mathbf{z} = \sum z_i^2 \sim \chi^2(p)$.

Beispiel 3.2 (Mittelwerts-Test)

Die Variablen `Lifeexpectancy, Selfreportedhealth` sollen bivariat auf den Erwartungswert $\boldsymbol{\mu}_0 = [6, 6]'$ getestet werden. Abb. 3.1 kann man die Mittelwerte sowie die empirischen Kovarianzen entnehmen. Die wahre Kovarianzmatrix $\boldsymbol{\Sigma}$ ist nicht bekannt. Wir nehmen daher zunächst an, daß $\boldsymbol{\Sigma}$ konstant und numerisch gleich der empirischen Kovarianzmatrix $\mathbf{S}$ ist (vgl. aber Abs. 3.2.3).

Somit hat man $N = 34$, $\bar{\mathbf{x}} = [6.154, 6.447]'$, $\boldsymbol{\Sigma} = \begin{bmatrix} 7.701 & 2.647 \\ 2.647 & 6.397 \end{bmatrix}$.

Die Teststatistik ist

$T^2 = N(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0) \sim \chi^2(2)$.

Man benötigt noch die Inverse der Kovarianz, d.h.

$\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} 0.151 & -0.063 \\ -0.063 & 0.182 \end{bmatrix}$.

Dies ergibt sich aus der Formel (2.10). Die Determinante ist $\det(\boldsymbol{\Sigma}) = 42.257$.

Insgesamt hat man also (die Realisation von T^2 wird klein geschrieben)

$t^2 = 34\, [0.154, 0.447] \begin{bmatrix} 0.151 & -0.063 \\ -0.063 & 0.182 \end{bmatrix} \begin{bmatrix} 0.154 \\ 0.447 \end{bmatrix} = 1.067$.

Der kritische Wert ist aber $\chi^2(0.95, 2) = 5.991$.

Damit muß $H_0 : \boldsymbol{\mu} = [6, 6]'$ auf dem 5%-Niveau beibehalten werden.

■

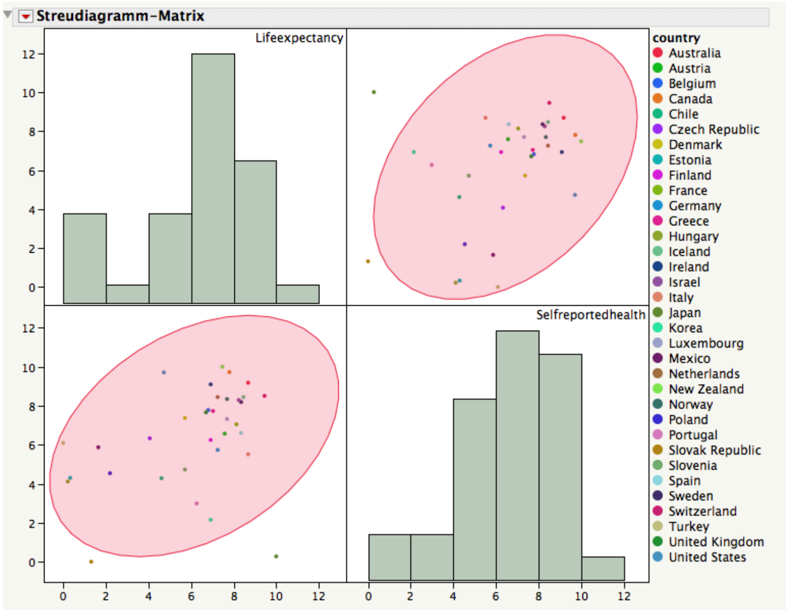
Der Stoff wird in Aufgabe 3.1 vertieft.

3.2.2 Konfidenzintervall für den Erwartungswert $\boldsymbol{\mu}$ ($\boldsymbol{\Sigma}$ bekannt)

Aus der Teststatistik T^2 läßt sich die Wahrscheinlichkeitsaussage

$$P\{N(\bar{\mathbf{x}} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}) \leq \chi^2(1 - \alpha, p)\} = 1 - \alpha \qquad (3.3)$$

herleiten. Die quadratische Form definiert ein Konfidenz-Ellipsoid (Ellipse für $p = 2$) im Parameterraum mit Zentrum $\bar{\mathbf{x}}$ und Konfidenz-Niveau



Deskriptive Statistiken

	Mittelwert	Standardabweichung	N
Life expectancy	6,153846	2,7750179	34
Self-reported health	6,446748	2,5291737	34

Korrelationen

		Life expectancy	Self-reported health
Life expectancy	Korrelation nach Pearson	1	,377*
	Signifikanz (2-seitig)		,028
	Quadratsummen und Kreuzprodukte	254,124	87,363
	Kovarianz	7,701	2,647
	N	34	34
Self-reported health	Korrelation nach Pearson	,377*	1
	Signifikanz (2-seitig)	,028	
	Quadratsummen und Kreuzprodukte	87,363	211,092
	Kovarianz	2,647	6,397
	N	34	34

*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Abbildung 3.1: OECD-Daten. Oben (JMP): Streudiagramm und Histogramm der Variablen `Lifeexpectancy`, `Selfreportedhealth`, sowie 95%-Ellipsen der empirischen Verteilung $N(\bar{\mathbf{x}}, \mathbf{S})$. unten (SPSS): Mittelwerte und Kovarianzen.

$1 - \alpha$.

Liegt der Punkt $\boldsymbol{\mu}_0$ innerhalb der Ellipse, so wird H_0 beibehalten, ansonsten abgelehnt.

Beispiel 3.3 (Mittelwerts-Test, Fortsetzung)

Betrachtet man Abb. 3.2, so würde H_0 bei einem Signifikanzniveau von $\alpha = 60\%$ (Konfidenzniveau von $1 - \alpha = 40\%$) abgelehnt (rote Ellipse), jedoch bei kleineren Niveaus beibehalten.

Die Überschreitungswahrscheinlichkeit (p-Wert) des Tests ist

p-**Wert**

$$p = P(T^2 > t^2 = 1.067) = 0.587. \tag{3.4}$$

Daher müßte $\alpha > p$ gewählt werden (etwa $\alpha = 60\%$), um ein signifikantes Ergebnis zu erhalten. Entsprechend gilt für das Konfidenzniveau $1 - \alpha < 1 - p = 0.413$.

Das Signifikanzniveau α muß jedoch *vor* Ausführung des Tests fixiert werden.
Übliche Praxis ist jedoch, den Test auszuführen und *nach* Betrachtung des p-Werts das Signifikanzniveau so zu wählen, daß man das gewünschte Ergebnis erhält.
Leider wird der p-Wert von SPSS als *Signifikanz* bezeichnet. Daher erscheinen die obigen Bemerkungen als wirkungslos.

Will man im Nachhinein die H_0 verwerfen, so wäre das zu wählende Signifikanzniveau ($\alpha > p = 0.587$, z.B. $\alpha = 0.6$) außerhalb der üblichen Werte (0.1, 0.05, 0.01). Mehr Spielraum zum „Erreichen" des gewünschten Testergebnisses bleibt bei p-Werten, die kleiner als 0.1 sind. Es handelt sich jedoch bei diesem Vorgehen um eine Verfälschung der Testprozedur.

■

3.2.3 Test für den Erwartungswert $\boldsymbol{\mu}$ ($\boldsymbol{\Sigma}$ unbekannt)

In der Praxis ist $\boldsymbol{\Sigma}$ meistens unbekannt. Wird es durch $\mathbf{S}$ geschätzt, so ist die Teststatistik $T^2 = N(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)'\mathbf{S}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)$ unter H_0 nicht mehr $\chi^2(p)$-verteilt.

Im skalaren Fall $p = 1$ hat man den Quotient aus dem Quadrat einer normalverteilten Größe (d.h. $\chi^2(1)$-verteilt) und einer $\chi^2(N-1)$-verteilten

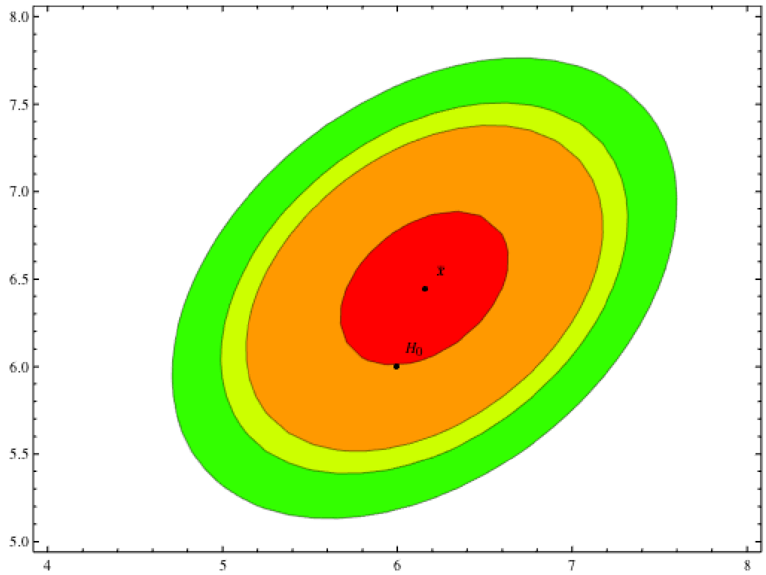


Abbildung 3.2: OECD-Daten. Bekanntes $\mathbf{\Sigma}$. Konfidenz-Ellipsen zu den Niveaus $1-\alpha = 0.4, 0.9, 0.95, 0.99$. Außerdem ist die Nullhypothese $H_0 : \boldsymbol{\mu}_0 = [6, 6]'$ eingezeichnet.

Stichprobenvarianz im Nenner. Dies führt also auf eine $F(1, N-1)$-Verteilung. Im skalaren Fall wird jedoch meistens der t-Test benutzt. In der Tat ist das Quadrat einer $t(N-1)$-verteilten Variable $F(1, N-1)$-verteilt.

Die Zusammenhänge zwischen den Testverteilungen N, χ^2, t und F sollten Ihnen bekannt sein. Siehe Kap. 10.1

Die Verteilung der Statistik

$$T^2 = N(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)'\mathbf{S}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0) \sim T^2(p, N-1) \tag{3.5}$$

Hotelling-T^2-Verteilung

wird als Hotelling-T^2-Verteilung bezeichnet. Sie hat, wie oben motiviert, einen engen Zusammenhang zur F-Verteilung. Es gilt der Zusammenhang

$$T^2(p,m) = \frac{mp}{m-p+1} F(p, m-p+1) \tag{3.6}$$

Setzt man $m = N - 1$, so läßt sich der Test mit Hilfe der $F(p, N - p)$-Verteilung durchführen. Der Testwert T^2 muß nur mit einem Faktor multipliziert werden:

$$\frac{N-p}{(N-1)p} \cdot T^2 := \tilde{T}^2 \sim F(p, N-p). \tag{3.7}$$

Beispiel 3.4 (Mittelwerts-Test, Fortsetzung)

In diesem Fall ist $N = 34$, $\bar{\mathbf{x}} = [6.154, 6.447]'$, $\mathbf{S} = \begin{bmatrix} 7.701 & 2.647 \\ 2.647 & 6.397 \end{bmatrix}$.

Die Teststatistik hat die Form

$$\tilde{T}^2 = \frac{N-p}{(N-1)p} \cdot N \cdot (\bar{\mathbf{x}} - \boldsymbol{\mu}_0)'\mathbf{S}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0) \sim F(p, N-p).$$

Man benötigt die Inverse der Stichproben-Kovarianz, d.h.

$\mathbf{S}^{-1} = \begin{bmatrix} 0.151 & -0.063 \\ -0.063 & 0.182 \end{bmatrix}$.

Dies ergibt sich aus der Formel (2.10). Die Determinante ist $\det(\mathbf{S}) = 42.257$.

Insgesamt hat man also (die Realisation von T^2 wird klein geschrieben)

$$\begin{aligned} \tilde{t}^2 &= \frac{34-2}{(34-1)2} \cdot t^2 \\ &= \frac{32}{66} \cdot 34[0.154, 0.447] \begin{bmatrix} 0.151 & -0.063 \\ -0.063 & 0.182 \end{bmatrix} \begin{bmatrix} 0.154 \\ 0.447 \end{bmatrix} = 0.517. \end{aligned}$$

Der kritische Wert (95%-Quantil) ist aber $F(0.95, 2, 32) = 3.295$.

Damit muß $H_0 : \boldsymbol{\mu}_0 = [6, 6]'$ auf dem 5%-Niveau beibehalten werden.

Der p-Wert $p = 0.601$ ist nun etwas größer und $1 - p = 0.399$.

Ein Signifikanzniveau von $\alpha = 60\%$ reicht nun nicht mehr aus, um H_0 abzulehnen (vgl. Abb. 3.3). Der Unterschied zwischen den Quantilen der

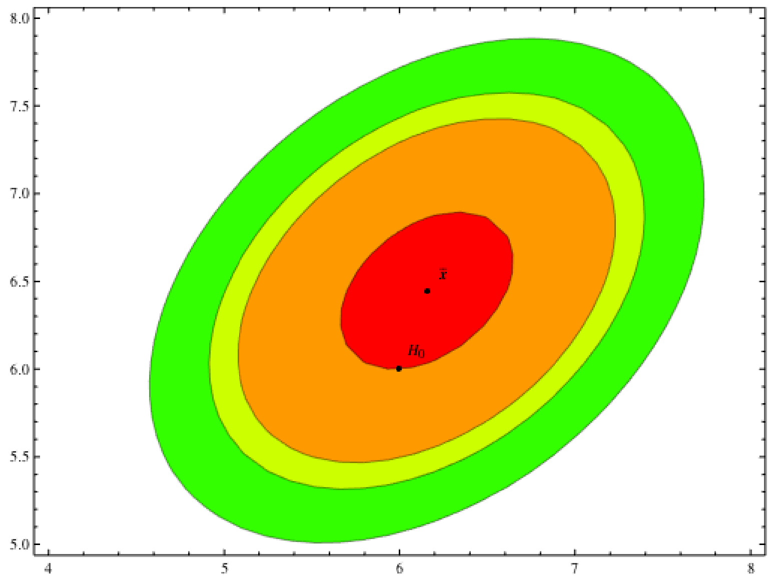


Abbildung 3.3: OECD-Daten. Unbekanntes $\boldsymbol{\Sigma}$. Konfidenz-Ellipsen zu den Niveaus $1-\alpha = 0.4, 0.9, 0.95, 0.99$. Außerdem ist die Nullhypothese H_0 : $\boldsymbol{\mu}_0 = [6, 6]'$ eingezeichnet.

χ^2-Verteilung und der Hotelling-T^2-Verteilung ist in Abb. 3.4, unten) zu sehen. Die Quantile der Hotelling-T^2-Verteilung sind immer größer, da ja $\boldsymbol{\Sigma}$ nur geschätzt wurde (analog zur Normal- und t-Verteilung).

Wählt man als Nullhypothese $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0 = [7, 5.5]'$, so ergibt sich

$\tilde{t}^2 = \frac{32}{66} \cdot 34\, [-0.846, 0.947] \begin{bmatrix} 0.151 & -0.063 \\ -0.063 & 0.182 \end{bmatrix} \begin{bmatrix} -0.846 \\ 0.947 \end{bmatrix} = 6.135.$

Damit muß H_0 auf dem 5%-Niveau abgelehnt werden (vgl. Abb. 3.6).

Der Stoff wird in Aufgabe 3.2 vertieft.

■

A.3 Aufgaben zu Kapitel 3

A.3.1

Verwenden Sie den Datensatz `mineral.sav`. Testen Sie für die Variablen `natrium` und `magnesium`, ob der Mittelwert von $\boldsymbol{\mu}_0 = [85,\ 31]$ zum Niveau $\alpha = 0.05$ verschieden ist! Gehen Sie dabei davon aus, dass die Daten unabhängig normalverteilt sind und dass die Inverse der (bekannten) Kovarianzmatrix gegeben ist:

$$\boldsymbol{\Sigma}^{-1} = \frac{1}{100\ 000}\begin{bmatrix} 3.6 & -5.6 \\ -5.6 & 119.9 \end{bmatrix} = 0.1166$$

A.3.2

Greifen Sie auf die Daten aus Aufgabe 3.1 zurück. Gehen Sie diesmal aber davon aus, dass die tatsächliche Kovarianzmatrix unbekannt ist!

Hinweis: Sie können die Stichproben-Konvarianzmatrix von SPSS im Menü „Bivariate Korrelationen/Optionen“ ausgeben lassen.

Prüfen Sie wiederum, ob sich der Mittelwert der Variablen `natrium` und `magnesium` zum Niveau $\alpha = 0.05$ von $\boldsymbol{\mu}_0 = [85,\ 31]$ unterscheidet.

L.3 Lösungen zu den Aufgaben (Kapitel 3)

L.3.1

Bestimmen Sie zunächst den Mittelwert der Daten und die Anzahl der Daten (z.B. über das Menü „Deskriptive Statistiken"):

$$\boldsymbol{\mu} = [85.8,\ 32]' \quad N = 103$$

Für die Teststatistik ergibt sich dann:

$$t^2 = \frac{103}{100\,000}[0.8,\ 1]\begin{bmatrix} 3.6 & -5.6 \\ -5.6 & 119.9 \end{bmatrix}\begin{bmatrix} 0.8 \\ 1 \end{bmatrix} = 0.117$$

Da $t^2 = 0.117 < 5.991 = \chi^2(0.95, 2)$ gilt, muss die Nullhypothese $H_0 : \boldsymbol{\mu}_0 = [85,\ 31]$ zum 5%-Niveau beibehalten werden.

L.3.2

Zunächst wird die Inverse der Stichproben-Kovarianzmatrix bestimmt:

$$\begin{aligned} \mathbf{S} &= \begin{bmatrix} 31\,137 & 1429 \\ 1429 & 976 \end{bmatrix} \\ \det(\mathbf{S}) &= 28\,347\,671 \\ \mathbf{S}^{-1} &= \frac{1}{28\,347\,671}\begin{bmatrix} 976 & -1\,429 \\ -1\,429 & 31\,137 \end{bmatrix} \end{aligned}$$

Die Teststatistik berechnet sich zu:

$$t^2 = \frac{101}{204} \cdot 103 \frac{1}{28\,347\,671}[0.8,\ 1]\begin{bmatrix} 976 & -1429 \\ -1429 & 31\,137 \end{bmatrix}\begin{bmatrix} 0.8 \\ 1 \end{bmatrix} = 0.053$$

Wegen $t^2 = .053 < 3,09 = F(0.95, 2, 101)$ muss die Nullhypothese wiederum zum 5%-Niveau beibehalten werden.