

Seminar: Quantitatives Risikomanagement Multivariate Modelle I

Prof: Hanspeter.Schmidli
Betreuung: Julia Eisenberg
Zhou,Ying

3. Multivariate Modelle I

Finanzielle Risikomodelle für den Absatzmarkt oder für Kreditrisiken sind grundsätzlich multivariat.z.B Die Wertänderung des Bestandes einer gehandelten Ware über einen festen Zeitraum hängt von einem zufälligen Vektor von Risikofaktoren oder dem Ertrag ab. Im weiteren betrachten wir einige Modelle für Zufallsvektoren, die teilweise nützlich für finanzielle Daten sind.

3.1 Grundlagen der Multivariaten Modellierung

In diesem Abschnitt betrachten wir die multivariate Normalverteilung und ihre Eigenschaften. Diese Verteilung ist wichtig für klassische multivariate Analysis und ist der Start für Modellierung des Markrisikos.

3.1.1 Zufallsvektoren und ihre Verteilungen

- *Gemeinsame- und Randverteilung*

Betrachte einen d-dimensionalen Zufallsvektor von Risiko-Faktor –Veränderung (oder returns)

$X = (X_1, \dots, X_d)'$. Die Verteilung von X ist beschrieben durch die gemeinsame

Verteilungsfunktion (df)

$$F_X(x) = F_X(x_1, \dots, x_d) = P(X \leq x) = P(X_1 \leq x_1, \dots, X_d \leq x_d)$$

Die Randverteilungsfunktion von X_i kann leicht aus der gemeinsamen df berechnet werden:

$$F_i(x_i) = P(X_i \leq x_i) = F(\infty, \dots, \infty, x_i, \infty, \dots, \infty)$$

Falls $F_i(x)$ absolut stetig ist, dann nennen wir ihre Ableitung $f_i(x)$ **Randdichte** von X_i . Es ist möglich eine k-dimensionale Randverteilung von X für $2 \leq k \leq d-1$ zu definieren. Wir können X in (X_1', X_2') zerlegen wobei $X_1 = (X_1, \dots, X_k)'$, $X_2 = (X_{k+1}, \dots, X_d)'$, dann ist

die Randverteilungsfunktion von X_1 :

$$F_{X_1}(x_1) = P(X_1 \leq x_1) = F(x_1, \dots, x_k, \infty, \dots, \infty).$$

Die df von einem Zufallsvektor X heißt **absolut stetig**, falls

$$F(x_1, \dots, x_d) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_d} f(u_1, \dots, u_d) du_1 \dots du_d,$$

wobei f als gemeinsame Dichte von X nicht negativ ist.

Bemerkung. Existenz von gemeinsamer Dichte \Rightarrow Existenz von Randdichte für alle k -dim. Randwerte.

Definitionen :

Überlebensfunktion ist definiert als

$$\bar{F}_X(x) = \bar{F}_X(x_1, \dots, x_d) = P(X > x) = P(X_1 > x_1, \dots, X_d > x_d),$$

Die **Randüberlebensfunktion** von X_i ist definiert als:

$$\bar{F}_i(x_i) = P(X_i > x_i) = \bar{F}(-\infty, \dots, -\infty, x_i, -\infty, \dots, -\infty).$$

● **Bedingte Verteilung und Unabhängigkeit**

Falls wir ein multivariates Modell für Risiken haben in der Form einer gemeinsamen df, Überlebensfunktion oder Dichte, dann haben wir auch die Abhängigkeitsstruktur für Risiken beschrieben.

Zerlegen X in $(X'_1, X'_2)'$, und angenommen df von X ist absolut stetig. Sei f_{X_1} gemeinsame

Dichte von k -dim. Randverteilung F_{X_1} . Dann hat bedingte Verteilung von X_2 mit gegebenem

$X_1 = x_1$ Dichte :

$$f_{X_2|X_1}(x_2 | x_1) = \frac{f(x_1, x_2)}{f_{X_1}(x_1)}, \quad (3.2)$$

Und die zugehörige df:

$$F_{X_2|X_1}(x_2 | x_1) = \int_{u_{k+1}=-\infty}^{x_{k+1}} \dots \int_{u_d=-\infty}^{x_d} \frac{f(x_1, \dots, x_k, u_{k+1}, \dots, u_d)}{f_{X_1}(x_1)} du_{k+1} \dots du_d.$$

X_1 und X_2 unabhängig $\Leftrightarrow F(x) = F_{X_1}(x_1)F_{X_2}(x_2) \forall x$. Besitzt X eine gemeinsame Dichte,

$$f(x) = f_{X_1}(x_1)f_{X_2}(x_2).$$

Die Komponenten von X sind unabhängig voneinander $\Leftrightarrow F(x) = \prod_{i=1}^d F_i(x_i) \forall x \in R^d$, in

diesem Fall besitzt X eine Dichte $f(x) = \prod_{i=1}^d f_i(x_i)$.

● **Momente und charakteristische Funktion**

Erwartungsvektor von X , falls er existiert:

$$E(X) := (E(X_1), \dots, E(X_d))'$$

Kovarianzmatrix, falls sie existiert:

$$\text{cov}(X) := E((X - E(X))(X - E(X))'),$$

Das (i,j)te Element von Σ ist.

$$\sigma_{ij} = \text{cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j),$$

Die Diagonalelemente sind Varianzen $V(X_i)$ von Komponenten von X . D.h

$$\sigma_{11} = \text{Var}(X_1), \dots, \sigma_{dd} = \text{Var}(X_d)$$

Bezeichnen die Korrelationsmatrix von X mit $\rho(X)$.

Definieren Y mit $Y_i = X_i / \sqrt{\text{Var}(X_i)} \quad \forall i$ dann bezeichnen $P = \rho(X) = \text{Cov}(Y)$, und das (i,j) Element:

$$\rho_{ij} = \rho(X_i, X_j) = \frac{\text{cov}(X_i, X_j)}{\sqrt{\text{var}(X_i) \text{var}(X_j)}}, \quad (3.3)$$

Verhältnis zwischen Korrelation und Kovarianz-Matrix: $P = \wp(\Sigma)$

$$\Delta(\Sigma) := \text{diag}(\sqrt{\sigma_{11}}, \dots, \sqrt{\sigma_{dd}}), \quad (3.4)$$

$$\wp(\Sigma) := (\Delta(\Sigma))^{-1} \Sigma (\Delta(\Sigma))^{-1}. \quad (3.5)$$

Für $B \in R^{k \times d}$, $b \in R^k$

$$E(BX + b) = BE(X) + b, \quad (3.7)$$

$$\text{cov}(BX + b) = B \text{cov}(X) B'. \quad (3.8)$$

Kovarianzmatrix (und auch Korrelationsmatrix) sind also positiv semi definit;

Aus (3.8) folgt $\text{var}(a'X) = a' \Sigma a \geq 0 \quad \forall a \in R^d$

Wir können die Cholesky-Zerlegung der positiv-definiten Kovarianzmatrix häufig benutzen. Sei $\Sigma = AA'$, A sei untere Dreiecksmatrix mit positiven Diagonalelementen. A heißt **Cholesky**

Faktor: $\Sigma^{1/2}$, und die Inverse davon ist $\Sigma^{-1/2}$.

Charakteristische Funktion:

$$\phi_X(t) = E(\exp(it'X)) = E(e^{it'X}), \quad t \in \mathbb{R}^d.$$

3.1.2 Standard Schätzer von Kovarianz und Korrelation

Angenommen, wir haben n Beobachtungen eines d -dim Risiko-Faktor Return Vektors X_1, \dots, X_n .

Dies sind tägliche, wöchentliche, monatliche oder jährliche Beobachtungen, die eine multivariate Zeitreihenfolge bilden. In diesem Abschnitt nehmen wir an, dass die Beobachtungen identisch verteilt sind und entweder unabhängig oder zumindest unkorreliert.

Wir nehmen an, dass die Betrachtungen X_1, \dots, X_n vom Erwartungsvektor μ , endlicher

Kovarianzmatrix Σ und Korrelationsmatrix P entstammen.

Stichprobenmittelwertvektor und Stichprobenkovarianzmatrix:

$$\bar{X} := \frac{1}{n} \sum_{i=1}^n X_i, \quad S := \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})', \quad (3.9)$$

\bar{X} ist erwartungstreu, aber S nicht.

Setze $S_u := nS/(n-1)$, dann ist S erwartungstreu.

$$\begin{aligned} nE(S) &= E\left(\sum_{i=1}^n (X_i - \mu)(X_i - \mu)' - n(\bar{X} - \mu)(\bar{X} - \mu)'\right) \\ &= \sum_{i=1}^n \text{cov}(X_i) - n \text{cov}(\bar{X}) = n\Sigma - \Sigma, \end{aligned}$$

wobei $\text{cov}(\bar{X}) = n^{-1}\Sigma$, falls der Vektor iid oder identisch verteilt ist.

Stichprobenkorrelationsmatrix R mit Elementen r_{jk} kann leicht berechnet werden durch

Stichprobenkovarianzmatrix.

Ihr (j,k)tes Element ist $r_{jk} = s_{jk} / \sqrt{s_{jj}s_{kk}}$ und $R = \rho(s)$.

3.1.3 Die multivariate Normalverteilung

Definition 3.1 $X = (X_1, \dots, X_d)'$ hat eine **multivariate normal** oder **Gaußsche Verteilung**,

falls

$X = \mu + AZ$, wobei $Z = (Z_1, \dots, Z_k)'$ ein Vektor von iid. Univariat Standard

normalverteilten Zufallsvariablen (mit Erwartungswert 0 und Varianz 1), $A \in \mathbb{R}^{d \times k}$, $\mu \in \mathbb{R}^d$,

Es ist leicht zu beweisen, dass $E(X) = \mu$, und $\text{Cov}(X) = \Sigma$, wobei $\Sigma = A'A$ positiv semidefinit.

Wir wissen, dass $\phi_Z(t) = \exp(-\frac{1}{2}t^2)$ die charakteristische Funktion einer Standard univariat normalverteilten Z . Daraus folgt, dass die charakteristische Funktion von X durch

$$\phi_X(t) = E(\exp(it'X)) = \exp(it'\mu - \frac{1}{2}t'\Sigma t), \quad t \in \mathbb{R}^d. \quad (3.10)$$

gegeben ist.

Notation: $X \sim N_d(\mu, \Sigma)$

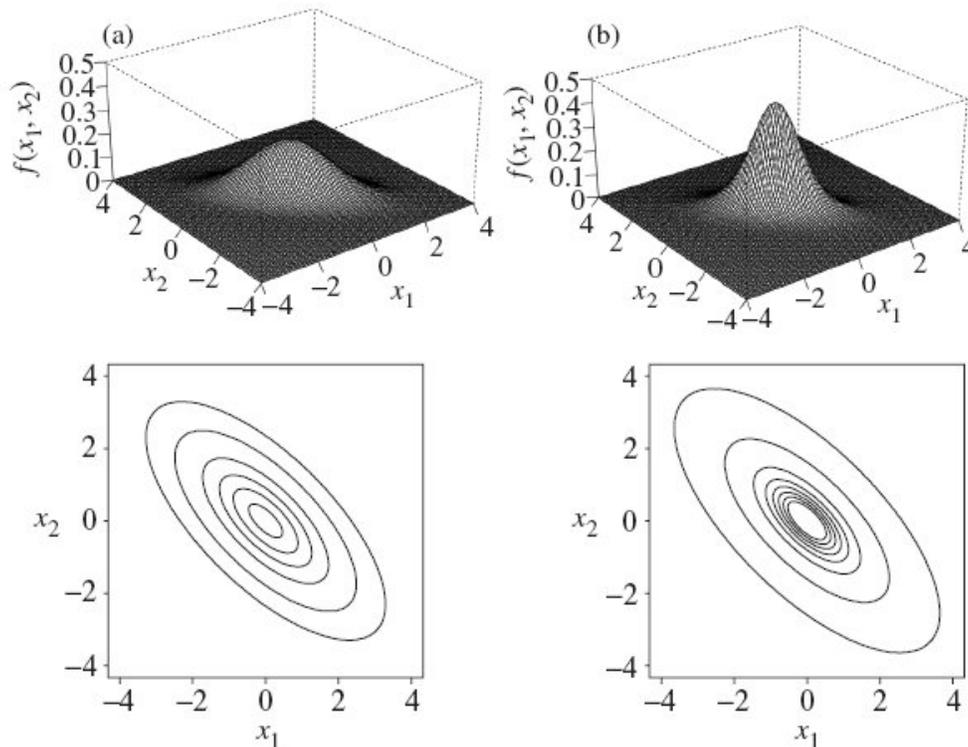
Bemerkung. Die Komponenten von X sind dann und nur dann unabhängig, wenn Σ eine Diagonalmatrix ist.

Betrachte $\text{Rang}(A) = d \leq k$, Σ hat den Rang d und deswegen invertierbar und positiv definit. X hat eine gemeinsame Dichte:

$$f(x) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\{-\frac{1}{2}(x - \mu)'\Sigma^{-1}(x - \mu)\}, \quad x \in \mathbb{R}^d, \quad (3.11)$$

Diese Form zeigt, dass sich Punkte mit gleichen Dichten auf der selben Ellipse befinden. Falls eine multivariate Dichte $f(x)$ von x nur von $(x - \mu)'\Sigma^{-1}(x - \mu)$ abhängt, dann ist $f(x)$ die Dichte einer elliptischen Verteilung. (Abschnitt 3.3). Im Bild 3.1 sehen wir Konturen gleicher Dichten sind Ellipsen.

Bild 3.1



(a) Perspektive und Kontur Plots für die Dichte einer bivariate Normalverteilung mit Standard Normalrandverteilung und Korrelation -70% . (b) die entsprechenden Plots für eine bivariate t-Dichte mit Freiheitsgrad 4 und gleichem Erwartungsvektor und Kovarianzmatrix.

Algorithmus 3.2 (Simulation der multivariaten Normalverteilung)

- (1) Zerlegen Σ durch Cholesky Zerlegung um $\Sigma^{1/2}$ zu erhalten.
- (2) Erzeugen $Z = (Z_1, \dots, Z_d)'$ von unabh. Standard normalverteilten ZV.
- (3) Setze $X = \mu + \Sigma^{1/2}Z$

Eigenschaften:

- **Lineare Kombination**

Lineare Kombination von multivariat normalen Zufallsvektoren bleibt multivariat normal.

Sei $X \sim N_d(\mu, \Sigma)$, $B \in R^{k \times d}$, $b \in R^k$ dann durch charakteristische Funktion (3.10) haben wir

$$BX + b \sim N_k(B\mu + b, B\Sigma B'). \quad (3.12)$$

Falls $a \in R^d$ dann

$$a'X \sim N(a'\mu, a'\Sigma a), \quad (3.13)$$

(Diese Eigenschaft ist anwendbar in der Varianz-Kovarianz Ansatz in Risikomanagement.)

Randverteilung: Die univariate Randverteilung von X muss univariat normal sein. (wegen 3.13)

- **Bedingte Verteilung**

Angenommen Σ ist positiv definit, dann ist die bedingte Verteilung von X_2 gegeben X_1 und von X_1 gegebenen X_2 auch multivariat Normal-verteilt.

- **Quadratische Formen**

Falls $X \sim N_d(\mu, \Sigma)$ mit Σ positiv definiert, dann

$$(X - \mu)' \Sigma^{-1} (X - \mu) \sim \chi_d^2, \quad (3.14)$$

eine χ^2 Verteilung mit d Freiheitsgraden.

Hier betrachten wir $Z = \Sigma^{-1/2}(X - \mu) \sim N_d(0, I_d)$ und $(X - \mu)' \Sigma^{-1} (X - \mu) = Z'Z \sim \chi_d^2$

- **Faltung**

Seien X und Y unabhängige d -dim. Zufallsvektoren mit $X \sim N_d(\mu, \Sigma)$ und $Y \sim N_d(\tilde{\mu}, \tilde{\Sigma})$,

dann mit Hilfe der char. Funktionen bekommen wir $X + Y \sim N_d(\mu + \tilde{\mu}, \Sigma + \tilde{\Sigma})$.

3.1.4 Tests für Normalität und multivariate Normalität

- **Univariate Tests**

Seien X_1, \dots, X_n iid multivariat normal, für $1 \leq j \leq d$, so muss die univariate Stichprobe

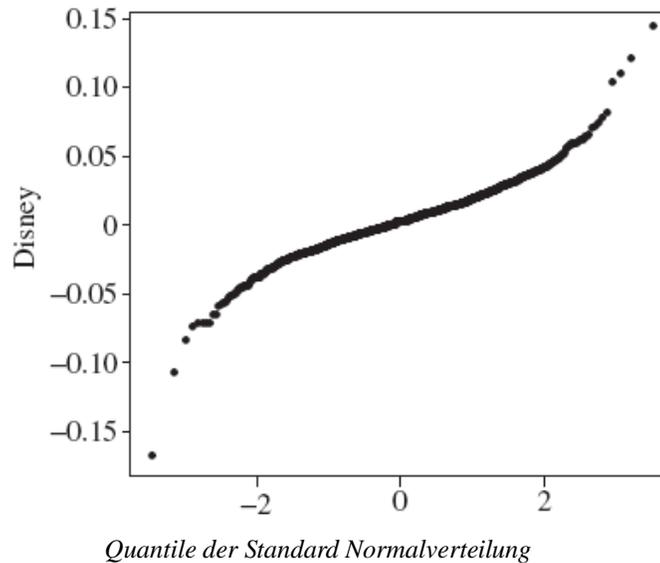
$X_{1,j}, \dots, X_{n,j}$ auch iid univariat normal sein. Wissen, jede univariate Stichprobe konstruiert durch

eine Linearkombination von der Form $a'X_1, \dots, a'X_n$ muss auch iid univariat normal sein. Das

kann auch mittels QQ-Plot grafisch gezeigt werden. (QQ-Plot ist ein grafisches Standardwerkzeug, das die Beziehung zwischen empirischen Quantilen und theoretischen Quantilen einer Referenzverteilung zeigt.)

In Bild 3.2 sehen wir ein QQ-Plot von täglichen Returns vom Disney Handelspreis von 1993 bis 2000 gegenüber einer Normal Referenzverteilung. Wir sehen dass die empirischen Quantile tendiert größer zu sein als die entsprechenden Quantile der Normalverteilung. D.h. die Normalverteilung ist ein schlechtes Modell für diese Returns.

Bild 3.2



Der anwendbare Jarque-Bera Test schätzt gleichzeitig ab ob die Schiefe und Wölbung der Verteilung mit dem gaußschen Modell übereinstimmen.

Der Schiefe- und Wölbungskoeffizient einer univariaten Stichprobe Z_1, \dots, Z_n sind definiert

durch

$$\sqrt{b} = \frac{(1/n) \sum_{i=1}^n (Z_i - \bar{Z})^3}{((1/n) \sum_{i=1}^n (Z_i - \bar{Z})^2)^{3/2}}, \quad k = \frac{(1/n) \sum_{i=1}^n (Z_i - \bar{Z})^4}{((1/n) \sum_{i=1}^n (Z_i - \bar{Z})^2)^2}. \quad (3.15)$$

Theoretische Schiefe und Wölbung sind definiert durch

$$\sqrt{\beta} = E(Z - \mu)^3 / \sigma^3 \quad \text{und} \quad \kappa = E(Z - \mu)^4 / \sigma^4$$

$\sqrt{\beta}$ und κ nehmen die Werte 0 und 3 entsprechend an für eine Normal-verteilt Z.

Jarque-Bera Test Statistik:

$$T = \frac{1}{6}n(b + \frac{1}{4}(k - 3)^2)$$

und hat eine asymptotisch chi-quadrat-Verteilung mit Freiheitsgrad 2 unter Nullhypothese von Normalität. Die Wölbung Werte der Stichprobe weichen stark von 3 ab und die Schiefe weicht stark von 0 ab, d.h. die Hypothese der Normalität trifft nicht zu.

● **Multivariate Tests:**

Wir testen gemeinsame Normalität.

Wissen: (3.14) hat eine chi-quadratische-Verteilung.

Wenden (3.9) an um μ und Σ zu schätzen und konstruieren

$$\{D_i^2 = (X_i - \bar{X})'S^{-1}(X_i - \bar{X}) : i = 1, \dots, n\}. \quad (3.16)$$

Weil die Schätzer von Erwartungsvektor und Kovarianzmatrix in jeder Konstruktion von D_i^2 benutzt werden, sind sie nicht unabhängig.

Wissen $n(n-1)^{-2}D_i^2 \sim \text{Beta}(\frac{1}{2}d, \frac{1}{2}(n-d-1))$ so dass die wahre Verteilung eine skalierte beta Verteilung ist. Wir erwarten D_1^2, \dots, D_n^2 verhalten sich wie eine iid Stichprobe von einer χ_d^2 -Verteilung, und wir konstruieren QQ-Plot gegen diese Verteilung.

Numerische Tests der multivariaten Normalität, basierend auf multivariaten Massen der Schiefe und Wölbung sind auch möglich.

Definiere:

$$b_d = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n D_{ij}^3, \quad k_d = \frac{1}{n} \sum_{i=1}^n D_i^4, \quad (3.17)$$

wobei D_i gegeben in (3.16) und bekannt als Mahalanobis Distanz zwischen X_i und \bar{X} und

$D_{ij} = (X_i - \bar{X})S^{-1}(X_j - \bar{X})$ als Mahalanobis Winkel zwischen $(X_i - \bar{X})$ und $(X_j - \bar{X})$.

Diese Masse werden zu univariaten Maßen b und k im Fall $d=1$. Unter der Nullhypothese der multivariaten Normalität ist die asymptotische Verteilung für $n \rightarrow \infty$

$$\frac{1}{6}nb_d \sim \chi_{d(d+1)(d+2)/6}^2, \quad \frac{k_d - d(d+2)}{\sqrt{8d(d+2)/n}} \sim N(0, 1). \quad (3.18)$$

Mardias Test enthält Vergleich von Schiefe und Wölbung Statistiken mit obigen theoretischen Referenzverteilungen.

Beispiel 3.3 (über die Normalität von Returns von Dow Jones 30 Aktien.)

Wir wenden den Test von Normalität zu einer bel. Teilmenge von 10 Aktien an. Wir betrachten 1993-2000 täglich, wöchentlich und vierteljährlich logarithmische Returns. Für jede Aktien berechnen wir die Stichproben-Schiefe und Wölbung und wenden Jarque-Beta Test zu der univariate Zeitfolge an. Für täglich und wöchentliche Returndaten fallen alle Tests durch.

Insbesondere, es gibt mache große Wert für Stichprobe Wölbung. Für monatliche Daten ist Nullhypothese nicht formal abgelehnt (ablehnen hier heißt p-value größer als 0.05) für 4 Aktien. Für vierteljährliche Daten ist nicht abgelehnt für 5 Aktien, obwohl die Stichprobegröße klein ist. Wir wenden Mardia's Tests von multinomalität an, die auf multivariat Schiefe und Wölbung für alle 10 Aktien basiert. Das Ergebnis ist in Tabelle 3.2 gezeigt. in Tabelle 3.3 verglichen D_i^2 Daten(3.16) zu einer χ_{10}^2 -Verteilung mit einer QQPlot.

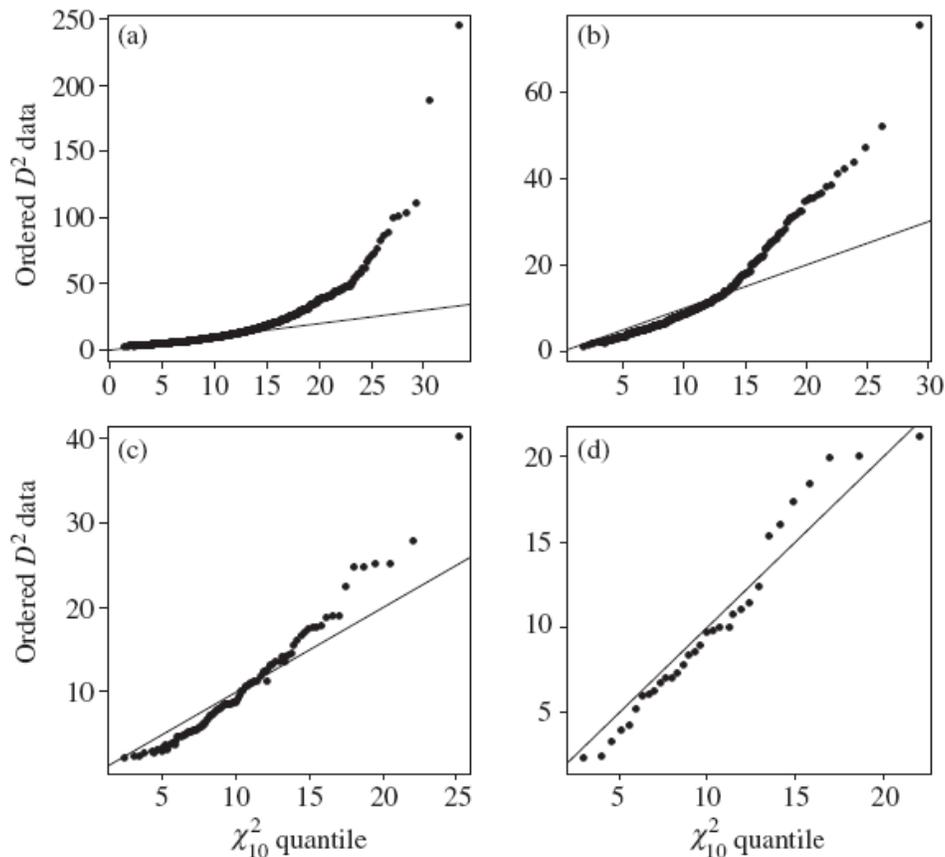
Table 3.1. Sample skewness (\sqrt{b}) and kurtosis (k) coefficients as well as p -values for Jarque–Bera tests of normality for an arbitrary set of 10 of the Dow Jones 30 stocks (see Example 3.3 for details).

Stock	\sqrt{b}	k	p -value	\sqrt{b}	k	p -value
	Daily returns, $n = 2020$			Weekly returns, $n = 416$		
AXP	0.05	5.09	0.00	-0.01	3.91	0.00
EK	-1.93	31.20	0.00	-1.13	14.40	0.00
BA	-0.34	10.89	0.00	-0.26	7.54	0.00
C	0.21	5.93	0.00	0.44	5.42	0.00
KO	-0.02	6.36	0.00	-0.21	4.37	0.00
MSFT	-0.22	8.04	0.00	-0.14	5.25	0.00
HWP	-0.23	6.69	0.00	-0.26	4.66	0.00
INTC	-0.56	8.29	0.00	-0.65	5.20	0.00
JPM	0.14	5.25	0.00	-0.20	4.93	0.00
DIS	-0.01	9.39	0.00	0.08	4.48	0.00
	Monthly returns, $n = 96$			Quarterly returns, $n = 32$		
AXP	-1.22	5.99	0.00	-1.04	4.88	0.01
EK	-1.52	10.37	0.00	-0.63	4.49	0.08
BA	-0.50	4.15	0.01	-0.15	6.23	0.00
C	-1.10	7.38	0.00	-1.61	7.13	0.00
KO	-0.49	3.68	0.06	-1.45	5.21	0.00
MSFT	-0.40	3.90	0.06	-0.56	2.90	0.43
HWP	-0.33	3.47	0.27	-0.38	3.64	0.52
INTC	-1.04	6.50	0.00	-0.42	3.10	0.62
JPM	-0.51	5.40	0.00	-0.78	7.26	0.00
DIS	0.04	3.26	0.87	-0.49	4.32	0.16

Table 3.2. Mardia's tests of multivariate normality based on the multivariate measures of skewness and kurtosis in (3.17) and the asymptotic distributions in (3.18) (see Example 3.3 for details).

	Daily	Weekly	Monthly	Quarterly
n	2020	416	96	32
b_{10}	9.31	9.91	21.10	50.10
p -value	0.00	0.00	0.00	0.02
k_{10}	242.45	177.04	142.65	120.83
p -value	0.00	0.00	0.00	0.44

Bild 3.3



Für die täglichen, wöchentlichen und monatlichen Returndaten fallen die multivariaten Tests von Normalität durch. Für vierteljährliche Return Daten lehnt die multivariate Wölbung Test die Nullhypothese nicht ab, aber der Schiefe Test schon. Die QQPlot im Bild 3.3(d) scheint leicht linear zu sein. Das ist Beweis dafür, dass Return über eine vierteljährliche Zeitspanne näher an der Normalverteilung ist.

Das Beispiel weist darauf hin, dass in vielen Risk-Management Anwendungen die multivariate Normalverteilung keine gute Beschreibung der Wirklichkeit ist. Die Nachteile sind:

- (1). Die Flanken der univariaten Randverteilungen sind zu dünn; Sie ordnen nicht genug "Gewicht" den extremfällen zu.
- (2). Die gemeinsame Flanken der Verteilung ordnet nicht genug Gewicht den gemeinsamen Extremfällen zu.
- (3). Diese Verteilung ist symmetrisch.(so genannte elliptische Symmetrie)

3.2 Gemischte Normalverteilungen

In diesem Abschnitt wird der Zufall zuerst in die Kovarianzmatrix und dann in den Erwartungsvektor einer multivariaten Normalverteilung durch eine positive Mischvariable W eingeführt.

3.2.1 Normal Varianz-Mischungen

Definition. Der Zufallsvektor X hat eine (multivariate) Normal Varianz Mischverteilung, falls

$$X \stackrel{d}{=} \mu + \sqrt{W}AZ, \quad (3.19)$$

Wobei (i) $Z \sim N_k(0, I_k)$;

(ii) $W \geq 0$ ist eine nicht negative, skalar wertige ZV, die unabhängig von Z ist

(iii) $A \in R^{k \times d}$, $\mu \in R^d$

Solche Verteilungen sind bekannt als Varianz Mischung.

Bei der multivariaten Normalverteilung interessiert man sich insbesondere für den Fall:

$\text{rang}(A) = d \leq k$, und Σ hat den Vollrang, positiv-definit; das bringt uns eine nicht-singuläre normal Varianz-Mischung.

Vorausgesetzt W hat eine endliche Erwartung, dann

$$E(X) = E(\mu + \sqrt{W}AZ) = \mu + E(\sqrt{W})AE(Z) = \mu$$

und

$$\text{cov}(X) = E((\sqrt{W}AZ)(\sqrt{W}AZ)') = E(W)AE(ZZ')A' = E(W)\Sigma. \quad (3.20)$$

Wir beziehen uns auf μ und Σ als Location Vector und dispersion matrix dieser Verteilung.

Bem. 1. Σ (Kovarianzmatrix von AZ) ist Kovarianzmatrix von X nur, falls $E(W) = 1$

2. μ ist der Erwartungsvektor nur, falls $E(X)$ definiert ist, d.h. $E(W^{1/2}) < \infty$

3. Diese Verteilungen sind ein gutes Beispiel der Modelle, wobei Korrelation = 0 nicht notwendigerweise die Unabhängigkeit von Komponenten von X bedeutet.

4. Die Korrelationsmatrizen von X und Z sind gleich, falls $E(W) < \infty$

Lemma 3.5 Hat (X_1, X_2) eine gemischte Normalverteilung mit $A=I_2$ und $E(W) < \infty$, so dass

$\text{cov}(X_1, X_2) = 0$. Dann sind X_1 und X_2 dann und nur dann unabhängig, wenn X fast sicher

konstant ist. d.h. (X_1, X_2) sind normal verteilt.

Beweis:

$$\begin{aligned} E(|X_1||X_2|) &= E(W|Z_1||Z_2|) = E(W)E(|Z_1|)E(|Z_2|) \\ &\geq (E(\sqrt{W}))^2 E(|Z_1|)E(|Z_2|) = E(|X_1|)E(|X_2|), \end{aligned}$$

Die Gleichheit gilt nur wenn W konstant ist.

Durch Anwendung von (3.10) können wir die charakteristische Funktion von einer normal Varianz Mischung berechnen.

$$\begin{aligned} \phi_X(t) &= E(E(\exp(it'X) | W)) = E(\exp(it'\mu - \frac{1}{2}Wt'\Sigma t)) \\ &= \exp(it'\mu) \hat{H}(\frac{1}{2}t'\Sigma t). \end{aligned} \quad (3.21)$$

wobei $\hat{H}(\theta) = \int_0^{\infty} e^{-\theta v} dH(v)$ Laplace-Stieltjes Transformation von df H von W. Aufgrund (3.21)

benutzen wir $X \sim M_d(\mu, \Sigma, \hat{H})$ für eine gemischte Normal Varianz.

Angenommen, Σ ist positiv definit, und die Verteilung von W hat kein Punktmass in 0, dann können wir die gemeinsame Dichte von der normal Varianz-Mischung Verteilung ableiten. Sei

$f_{X|W}$ die bedingte Dichte von X gegeben W. Die Dichte von X ist gegeben durch:

$$\begin{aligned} f(x) &= \int f_{X|W}(x | w) dH(w) \\ &= \int \frac{w^{-d/2}}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp \left\{ -\frac{(x - \mu)' \Sigma^{-1} (x - \mu)}{2w} \right\} dH(w), \end{aligned} \quad (3.22)$$

mit Lebesque-Stieltjes Integral; falls H Dichte h hat, meinen wir das Riemannsche Integral

$\int_0^{\infty} f_{X|W}(x | w) h(w) dw$. Alle solche Dichten werden von x nur durch die quadratische Form

$(x - \mu)' \Sigma^{-1} (x - \mu)$ abhängen. Das bedeutet, sie sind die Dichten von elliptischen Verteilungen.

Beispiel 3.6 (Gemischte zwei-Punkte Normal-Verteilungen)

Ein einfaches Beispiel von gemischten Normalverteilungen erhält man, falls W eine diskrete ZV ist. z.B das zwei-Punkte gemischte Normal modell wird dadurch erhalten, dass W in (3.19) eine diskrete rv wird, die nimmt an, dass die positive Werte k_1 und k_2 mit Wahrscheinlichkeit p und

1-p angenommen werden. Wähle k_2 relativ groß zu k_1 und wähle p groß, diese Verteilung

können angewendet werden um 2 Verläufe zu definieren.

- (1) Ein gewöhnliche Verlauf, der meiste Zeit gilt.
- (2) Ein Stress-Verlauf, der mit Wahrscheinlichkeit 1-p passiert.

Beispiel 3.7 (multivariate t Verteilung)

Setze W in (3.19) zu einer rv mit einer Inversen Gamma Verteilung $W \sim Ig(\frac{1}{2}v, \frac{1}{2}v)$, dann hat X eine mutivariate t Verteilung mit Freiheitsgrad 4.

Notation für die Multivariate t : $X \sim t_d(v, \mu, \Sigma)$. Σ ist nicht die Kovarianzmatrix von X in der Definition von Multivariate t. Wegen $E(W)=v/(v-2)$ haben wir $Cov(X)=(v/(v-2)) \Sigma$ und Kovarianzmatrix(und Korrelationsmatrix) von dieser Verteilung sind definiert falls $v > 2$.

Wir wenden (3.22) an, um die Dichte zu berechnen,

$$f(x) = \frac{\Gamma(\frac{1}{2}(v + d))}{\Gamma(\frac{1}{2}v)(\pi v)^{d/2} |\Sigma|^{1/2}} \left(1 + \frac{(x - \mu)' \Sigma^{-1} (x - \mu)}{v} \right)^{-(v+d)/2}. \quad (3.23)$$

Offensichtlich ist die Ortskurve von Punkten mit gleicher Dichte wieder ein Ellipsoid mit

Gleichung $(x - \mu)' \Sigma^{-1} (x - \mu) = c$ für $c > 0$. Im Vergleich mit der multivariaten normal Verteilung steigen die Konturen gleicher Dichte schnell in das Zentrum der Verteilung und verfallen schrittweise auf den „niedrigeren Hang“ von der Verteilung.

Beispiel 3.8 (symmetrische verallgemeinerte hyperbolische Verteilung) Eine flexible Familie von gemischter Normal Varianz ist enthalten durch einsetzen von W in (3.19) und wir bekommen eine verallgemeinerte Inverse Gaußsche Verteilung (GIG), $W \sim N^-(\lambda, \chi, \psi)$. Mit der Anwendung von (3.22) wird gezeigt, dass eine gemischte Normal Varianz konstruiert mit dieser gemischten Dichte eine gemeinsame Dichte hat.

$$f(x) = \frac{(\sqrt{\chi\psi})^{-\lambda} \psi^{d/2}}{(2\pi)^{d/2} |\Sigma|^{1/2} K_\lambda(\sqrt{\chi\psi})} \frac{K_{\lambda-(d/2)}(\sqrt{(\chi + (x - \mu)' \Sigma^{-1} (x - \mu))\psi})}{(\sqrt{(\chi + (x - \mu)' \Sigma^{-1} (x - \mu))\psi})^{(d/2)-\lambda}}, \quad (3.24)$$

wobei K_λ die modifizierte Bessel Funktion 3. Gattung ist.

Die GIG gemischte Verteilung ist sehr flexibel und enthält Gamma und Inverse Gamma Verteilung als besonderen Grenzfällen. (in Bezug auf $\lambda > 0, \chi = 0$, auf $\lambda < 0, \psi = 0$) In diesem Fall wird die Dichte in (3.24) als ein Limit interpretiert als $\chi \rightarrow 0$ oder $\psi \rightarrow 0$. Die gemischte Gamma Verteilung ergibt eine Laplace Verteilung oder ein sogenanntes symmetrisches Varianz-Gamma Modell und die Inverse Gamma ergibt t wie im Beispiel 3.7; t entspricht in dem Fall $\lambda = -v/2$ und $\chi = v$. Die besonderen Fälle $\lambda = -0.5$ und $\lambda = 1$ haben Aufmerksamkeit bei den finanziellen Modellen; Das zweite führt zu einer symmetrischen Multivariaten Verteilung, dessen ein-dim. Randverteilung als hyperbolische Verteilung bekannt ist.

Um die Kovarianz-Matrix von den Verteilungen in der verallgemeinerten symmetrischen Hyperbelfamilie zu berechnen, verlangen wir den Erwartungswert der GIG-Verteilung, die in (A9) für $\chi > 0$ und $\psi > 0$ gegeben ist. Die Kovarianzmatrix von der multivariaten Verteilung in (3.24) kommt vom (3.20),

Proposition 3.9 Falls $X \sim M_d(\mu, \Sigma, \hat{H})$ und $Y = BX + b$, wobei $B \in R^{k \times d}$ und $b \in R^k$, dann

$$Y \sim M_k(B\mu + b, B\Sigma B', \hat{H}):$$

Beweis: Wir wenden die charakteristische Funktion in (3.21) an,

$$\phi_Y(t) = E(e^{it'(BX+b)}) = e^{it'b} \phi_X(B't) = e^{it'(B\mu+b)} \hat{H}(\frac{1}{2}t' B \Sigma B' t).$$

So wird die Unterklasse der Mischverteilung, die durch \hat{H} angegeben ist, unter einer linearen Transformation geschlossen. z.B falls X eine multivariate t Verteilung mit Freiheitsgrad v ist, dann ist jede lineare Transformation von X auch so; Die lineare Kombination $a'X$ würde eine univariate t Verteilung mit Freiheitsgrad v haben.

Normal Varianz Mischverteilung werden leicht simuliert durch Def.3.4. Um eine Variete

$X \sim M_d(\mu, \Sigma, \hat{H})$ mit Σ positiv definit verwenden wir den folgenden Algorithmus.

Algorithmus 3.10 (Simulation von Normal Varianz-Mischung)

- (1) Erzeugen $Z \sim N_d(0, \Sigma)$ mit Algo.3.2.
- (2) Erzeugen unabhängig eine positive gemischte Variable W mit df H (entsprechend zur Laplace-Stieltjes Transformation \hat{H})
- (3) Setze $X = \mu + \sqrt{W}Z$

Um $X \sim t_d(v, \mu, \Sigma)$ zu erzeugen, sollte die gemischte variable W eine $\text{Ig}\left(\frac{1}{2}v, \frac{1}{2}v\right)$ Verteilung

haben; und in dem Fall $v/W \sim \chi_v^2$ hat eine chi-quadratische Verteilung Freiheitsgrad v . Um eine Stichprobe von einer verallgemeinerten hyperbolischen Verteilung mit Dichte (3.24) zu nehmen, müssen wir $W \sim N^-(\lambda, \chi, \Psi)$ erzeugen. Um eine Stichprobe von der GIG Verteilung zu nehmen, können wir eine Ablehnen-Algorithmus von Atkinson benutzen.

3.2.2 Gemischte Erwartungswert-Varianz Normalverteilung

Alle multivariaten Verteilungen, die wir bis jetzt betrachtet haben, haben elliptische Symmetrien und das kann ein vereinfachtes Modell für wahre risiko-Faktor return Daten sein.

Unter anderem bedeutet elliptische Symmetrie dass alle eindimensionalen Randverteilungen symmetrisch sind, der die häufige Beobachtung für Aktien Returns widerspricht, weil negative Returns dickere Flanken als positive Returns hat. Das Modell ,das wir jetzt einführen, versucht eine Asymmetrie zu der Klasse von gemischten Normalverteilungen hinzufügen, und der gibt die Klasse der multivariate normalen Erwartungs-Varianz Mischung.

Definition 3.11 Der Zufallsvektor X hat eine (multivariate) Normal Erwartungswert-Varianz Mischungverteilung, falls

$$X \stackrel{d}{=} m(W) + \sqrt{W}AZ, \quad (3.25)$$

Wobei

- (i) $Z \sim N_k(0, I_k)$
- (ii) $W \geq 0$ ist nicht negativ, Skalar-bewertet rv, die von Z unabhängig ist.
- (iii) $A \in R^{d \times k}$ ist eine Matrix
- (iv) $M: [0, \infty) \rightarrow R^d$ ist eine messbare Funktion

In diesem Fall haben wir $X|W = w \sim N_d(m(w), w\Sigma)$ (3.26)

Wobei $\Sigma = AA'$.

Im allgemein sind solche Verteilung nicht elliptisch.

Eine mögliche konkrete Spezifizierung für die Funktion $m(W)$ in (3.26) ist

$$m(W) = \mu + W\gamma, \quad (3.27)$$

wobei μ und γ sind Parametervektor in R^d . Wegen $E(X|W) = \mu + W\gamma$ und $\text{cov}(X|W) = W\Sigma$ haben wir

$$E(X) = E(E(X | W)) = \mu + E(W)\gamma, \quad (3.28)$$

$$\begin{aligned} \text{cov}(X) &= E(\text{cov}(X | W)) + \text{cov}(E(X | W)) \\ &= E(W)\Sigma + \text{var}(W)\gamma\gamma', \end{aligned} \quad (3.29)$$

falls die gemischte Variable W eine endliche Varianz hat.

Von (3.28) und (3.29) wissen wir dass μ und Σ sind der Erwartungsvektor und Kovarianzmatrix von X sind. Nur wenn $\gamma=0$, ist die Verteilung eine gemischte normale Varianz und die gegebene Moment-Formel in (3.20) gilt.

3.2.3 Verallgemeinerte hyperbolische Verteilungen

Im Beispiel 3.8 schauten wir auf die spezielle Unterklasse der verallgemeinerten hyperbolischen Verteilung, die aus dem elliptisch symmetrischen normalen Varianz Mischverteilung bestehen. Die volle verallgemeinerte Hyperbelfamilie wird dadurch erhalten, dass Erwartungswert-Varianz Mischung Konstruktion (3.25) und die bedingte Erwartungswert Spezifizierung (3.27). Für die Verteilungsmischung nahmen wir an $W \sim N^-(\lambda, \chi, \psi)$, eine GIG Verteilung mit Dichte (A.8).

Bemerkung 3.12. Diese Klasse der Verteilungen hat viel Aufmerksamkeit in der Literatur über finanziell Modelle erhalten, besonders im univariaten Fall. Ein wichtiger Grund ist ihre Verbindung

zu Levy-Prozessen mit unabhängigen und stationären Zuwächsen (wie die Brownsche Bewegung), die zum Preisprozessen modellieren verwendet werden in der stetiger Zeit. Für jeden verallgemeinerte Hyperbolische Verteilung ist es möglich um einen Levy Prozess zu bauen, so dass der Wert der Zunahme des Prozesses über ein festes Zeitintervall diese Verteilung hat; das ist nur möglich weil das verallgemeinerte Hyperbolische Gesetz eine sogenannte unendlich teilbare Verteilung ist, eine Eigenschaft die von der GIG gemischte Verteilung von W ist.

Die gemeinsame Dichte im nicht-sigulär Fall (Σ hat rang d):

$$\begin{aligned} f(x) &= \int_0^\infty \frac{e^{(x-\mu)'\Sigma^{-1}\gamma}}{(2\pi)^{d/2}|\Sigma|^{1/2}w^{d/2}} \\ &\quad \times \exp \left\{ -\frac{(x-\mu)'\Sigma^{-1}(x-\mu)}{2w} - \frac{\gamma'\Sigma^{-1}\gamma}{2/w} \right\} h(w) dw, \end{aligned}$$

wobei $h(w)$ die Dichte von W ist. Dadurch erhalten wir die verallgemeinerte hyperbolische Dichte

$$f(x) = c \frac{K_{\lambda-(d/2)}(\sqrt{(\chi + (x-\mu)'\Sigma^{-1}(x-\mu))(\psi + \gamma'\Sigma^{-1}\gamma)})e^{(x-\mu)'\Sigma^{-1}\gamma}}{(\sqrt{(\chi + (x-\mu)'\Sigma^{-1}(x-\mu))(\psi + \gamma'\Sigma^{-1}\gamma)})^{(d/2)-\lambda}}, \quad (3.30)$$

wobei die normalisierende Konstant ist

$$c = \frac{(\sqrt{\chi\psi})^{-\lambda}\psi^\lambda(\psi + \gamma'\Sigma^{-1}\gamma)^{(d/2)-\lambda}}{(2\pi)^{d/2}|\Sigma|^{1/2}K_\lambda(\sqrt{\chi\psi})}.$$

Falls $\gamma=0$, wird die Verteilung zu einer symmetrisch verallgemeinert hyperbolischen besonders im Fall von Beispiel 3.8. Im allgemeinen haben wir eine nicht-elliptische Verteilung mit asymmetrischer Randverteilung. Der Erwartungsvektor und Kovarianzmatrix von dieser Verteilung ist leicht berechenbar durch (3.28) und (3.29). Die char. Funktion von der verallgemeinerten hyperbolischen Verteilung kann berechnet werden durch (3.21):

$$\phi_X(t) = E(e^{it'X}) = e^{it'\mu} \hat{H}(\frac{1}{2}t'\Sigma t - it'\gamma), \quad (3.31)$$

wobei \hat{H} die Laplace-Stieltjes Transformation von der GIG Verteilung.

Wir übernehmen die Notation $X \sim GH_d(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$.

Bem. 1. Verteilungen $GH_d(\lambda, \chi/k, k\psi, \mu, k\Sigma, k\gamma)$ und $GH_d(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ sind identisch für $k>0$, welcher ein Identifizierbarkeit Problem verursacht, wenn wir versuchen die Parameter zu schätzen. Das kann gelöst werden, die Determinante $|\Sigma|$ mit besonderem Wert (wie 1) dazu beschränken, indem Sie passen.

2. Während solch eine Einschränkung wird eine Wirkung auf die zu schätzende Werte von χ und ψ haben, und es wird keine Wirkung auf $\chi\psi$, so ist dieses Produkt ein nützlicher zusammenfassender Parameter für den GH Verteilung.

● **Lineare Kombination.**

Die verallgemeinerte hyperbolische Klasse ist abgeschlossen unter dem linearen Operationen.

Proposition 3.13 Seien $X \sim GH_d(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ und $Y=BX+b$, wobei $B \in R^{k \times d}$, und

$b \in R^k$, dann $Y \sim GH_k(\lambda, \chi, \psi, B\mu + b, B\Sigma B', B\gamma)$.

Beweis. Wenden (3.31) und Proposition 3.9, dann

$$\phi_Y(t) = e^{it'(B\mu+b)} \hat{H}(\frac{1}{2}t' B \Sigma B' t - it' B \gamma).$$

So bleiben die von der GIG Mischverteilung übernommenen Parameter unverändert unter die linearen Operationen. D.h z,B die Randverteilung von X ist leicht

berechenbar : $X_i \sim GH_1(\lambda, \chi, \psi, \mu_i, \Sigma_{ii}, \gamma_i)$. Es bedeutet auch, es würde relativ leicht sein,

eine Version der Varianz-Kovarianz Methode darauf ein verallgemeinertes Hyperbolische Modell für Risikofaktoren zu stützen.

● **Parametrisierung.**

Es gibt eine verwirrende Reihe der alternativen parametrisierungen für die verallgemeinerte Hyperbolische Verteilung in der Literatur und es ist üblicher diesen Verteilung in einer wiederparametrisierten Form zu sehen. In einer allgemeiner Version von Kovarianzmatrix, die wir

Σ nennen, wird Δ umbenannt, und die Einschränkung wird auf $|\Delta| = 1$; das richtet das Identifizierbarkeit Problem von oben. Die Schiefe-Parameter χ wird ersetzt durch β , und die nichtnegativen χ und ψ werden ersetzt durch nicht negativ δ und α , gemäß:

$$\beta = \Delta^{-1}\gamma, \quad \delta = \sqrt{\chi}, \quad \alpha = \sqrt{\psi + \gamma' \Delta^{-1} \gamma}.$$

Diese Parameter müssen die Einschränkung $\delta \geq 0, \alpha^2 > \beta' \Delta \beta$, falls $\lambda > 0$; $\delta > 0$,

$\alpha^2 > \beta' \Delta \beta$, falls $\lambda = 0$; $\delta > 0, \alpha^2 \geq \beta' \Delta \beta$ falls $\lambda < 0$

Blaesild verwendet diese Parametrisierung, um zu zeigen, dass die verallgemeinerte Hyperbolische Verteilungsform eine geschlossene Klasse der Verteilungen unter linearen Operationen und Bedingen. Jedoch, Parametrisierung hat das Problem, daas die wichtigen Parameter α und δ nicht allgemein invariant unter jeder dieser Operationen sind. Es ist nützlich im Stande zu sein, sich leicht zwischen $\chi - \psi - \Sigma - \gamma$ zu bewegen wie in (3.30)

und $\alpha - \delta - \Delta - \beta$ Parametrisierung; Falls $\chi - \psi - \Sigma - \gamma$ verwendet, dann die andere Parametrisierung wird dadurch erreicht:

$$\Delta = |\Sigma|^{-1/d} \Sigma, \quad \beta = \Sigma^{-1} \gamma, \\ \delta = \sqrt{\chi |\Sigma|^{1/d}}, \quad \alpha = \sqrt{|\Sigma|^{-1/d} (\psi + \gamma' \Sigma^{-1} \gamma)}.$$

Falls $\alpha - \delta - \Delta - \beta$ verwendet ist, dann erhalten wir Parametrisierung duch einsetzen:

$$\Sigma = \Delta, \quad \gamma = \Delta \beta, \quad \chi = \delta^2, \quad \psi = (\alpha^2 - \beta' \Delta \beta).$$

● **Besonderer Fall.**

Der multivariate verallgemeinerte Hypobolische Familie ist flexibel und enthält viele spezielle Fälle.

- Falls $\lambda = \frac{1}{2}(d+1)$, dann lassen wir das Wort „Verallgemeinert“ fallen und beziehen uns auf die Verteilung als eine d-dim.hyperbolische Verteilung. Die univariate Randverteilung von dieser Verteilung hat $\lambda = \frac{1}{2}(d+1)$ und ist keine ein-dim. hyperbolische Verteilungen.
- Falls $\lambda = 1$, dann haben wir eine multivariate Verteilung, deren univariate Randverteilungen ein-dim. hyperbolische Verteilungen sind. Die ein-dim. hyperbolische Verteilung ist weit verwendet in der univariaten Analysis von Finanziellen Return Daten.
- Falls $\lambda = -\frac{1}{2}$, dann ist diese Verteilung eine NIG Verteilung. In dem univariaten Fall, wird dieses Modell für die Analyse von Return Daten verwendet; Seine funktionelle Form ist ähnlich zu der hyperbolischen mit einem ein bisschen schwereren Flanke . .
- Falls $\lambda > 0$ und $\chi = 0$ bekommen wir einen Begrenzungsfall von der Verteilung , der verschieden als eine verallgemeinerte Laplace, Bessel Funktion oder Varianz-Gamma Verteilung ist.
- Falls $\lambda = -\frac{1}{2} \nu$, $\chi = \nu$, und $\psi = 0$ bekommen wir einen anderen Begrenzungsfall, der weniger gut studiert worden zu sein scheint, aber er kann asymmetrische oder verdrehte t

Verteilung genannt werden. Durch Schätzen das Grenzwert von (3.30) für $\psi \rightarrow 0$

bekommen wir die multivariate Dichte

$$f(\mathbf{x}) = c \frac{K_{(v+d)/2}(\sqrt{(v+Q(\mathbf{x}))\boldsymbol{\gamma}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}}) \exp((\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}))}{(\sqrt{(v+Q(\mathbf{x}))\boldsymbol{\gamma}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}})^{-(v+d)/2} (1+(Q(\mathbf{x})/v))^{(v+d)/2}}, \quad (3.32)$$

Wobei $Q(\mathbf{x}) = (\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})$ und

$$c = \frac{2^{1-(v+d)/2}}{\Gamma(\frac{1}{2}v)(\pi v)^{d/2}|\boldsymbol{\Sigma}|^{1/2}}.$$

3.2.4 Fitten der verallgemeinerten hyperbolischen Verteilungen

Die univariaten verallgemeinerten hyperbolischen Modelle sind für Return Daten in vielen empirischen Studien geeignet, aber sie werden relativ wenig mit den multivariaten Verteilungen verwendet. Jedoch kann die gemischte Normalverteilung mit Algorithmus der Typ EM (expectation maximization) ausgerüstet werden.

Angenommen, wir haben iid X_1, \dots, X_n und möchten den multivariaten verallgemeinerten

hyperbolischen Fall passen. Wir fassen die Parameter zusammen in $\boldsymbol{\theta} = (\lambda, \chi, \psi, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\gamma})'$, und das Problem ist zu maximieren

$$\ln L(\boldsymbol{\theta}; X_1, \dots, X_n) = \sum_{i=1}^n \ln f_X(X_i; \boldsymbol{\theta}), \quad (3.33)$$

wobei $f_X(x; \boldsymbol{\theta})$ die verallgemeinerte hyperbolische Dichte in (3.30) ist.

Wir können die latent gemischte Variablen W_1, \dots, W_n vom (3.25) beobachten. Die

gemeinsame Dichte von beliebigen Paaren X_i und W_i ist gegeben durch

$$f_{X,W}(\mathbf{x}, \mathbf{w}; \boldsymbol{\theta}) = f_{X|W}(\mathbf{x} | \mathbf{w}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\gamma}) h_W(\mathbf{w}; \lambda, \chi, \psi), \quad (3.34)$$

dann können wir Likelihood konstruieren

$$\begin{aligned} \ln \tilde{L}(\boldsymbol{\theta}; X_1, \dots, X_n, W_1, \dots, W_n) \\ = \sum_{i=1}^n \ln f_{X|W}(X_i | W_i; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\gamma}) + \sum_{i=1}^n \ln h_W(W_i; \lambda, \chi, \psi), \end{aligned} \quad (3.35)$$

dabei können die zwei Terme getrennt maximiert werden. Der erste Term kann durch die Gaußschen Form leicht geschätzt werden.

Um die Latenz der W_i Daten zu überwinden, wird der Algorithmus von EM verwendet. Das

ist ein wiederholendes Verfahren, das aus einem E-Schritt und einem M-Schritt besteht.

Nehmen wir an, dass wir am Anfang des Schritts k Parameter Schätzer $\theta^{[k]}$ haben.

E-Schritt. Wir berechnen den bedingten Erwartungswert von der erweiterten Likelihood

(3.35) mit gegebenen X_1, \dots, X_n .

$$Q(\theta; \theta^{[k]}) = E(\ln \tilde{L}(\theta; X_1, \dots, X_n, W_1, \dots, W_n) \mid X_1, \dots, X_n; \theta^{[k]}).$$

M-Schritt. Wir maximieren die Zielfunktion mit Beachtung von θ und bekommen das

nächste $\theta^{[k+1]}$.

Durch abwechseln dieser Schritte erzeugt der Algorithmus von EM verbesserte Parameter Schätzer in jeden Schritt, und wir nähern uns an dem maximalen likelihood Schätzer an.

Druchführen von E-Step bedeutet tauschen funktionen $g(W_i)$ in (3.35) mit

$E(g(W_i) \mid X_i; \theta^{[k]})$ aus. Um es zu berechnen betrachten wir die bedingte Dichte von W_i

mit gegebenem X_i , die $f_{W_i|X}(w \mid x; \theta) \propto f_{W,X}(w \mid x; \theta)$ erfüllt. Dann haben wir aus (3.34)

$$W_i \mid X_i \sim N^-(\lambda - \frac{1}{2}d, (X_i - \mu)' \tilde{\Sigma}^{-1} (X_i - \mu) + \chi, \psi + \gamma' \Sigma^{-1} \gamma). \quad (3.36)$$

wenn wir likelihood (3.35) ausschreiben durch Verwenden von (3.26) für den ersten Term,

und die GIG Dichte für den zweiten Term, dann haben wir $g_1(w) = w$, $g_2(w) = 1/w$,

$g_3(w) = \ln(w)$. Und $E(\ln(W_i) \mid X_i; \theta^{[k]})$ schließt Ableitungen von einer

Bessel-Funktion in Bezug auf die Ordnung ein und muss numerisch approximiert werden.

Der folgend Notation

$$\delta_i^{[k]} = E(W_i^{-1} \mid X_i; \theta^{[k]}), \quad \eta_i^{[k]} = E(W_i \mid X_i; \theta^{[k]}), \quad \xi_i^{[k]} = E(\ln(W_i) \mid X_i; \theta^{[k]}) \quad (3.37)$$

erlaubt uns, das grundlegende Schema von EM zu beschreiben.

Im M-Step gibt es zwei Terme zu maximieren: $Q_1(\mu, \Sigma, \gamma; \theta^{[k]})$ und $Q_2(\lambda, \chi, \psi; \theta^{[k]})$.

Um das Identifizierbarkeit Problem in 3.2.3 anzugehen, beschränken wir die Determinante

von Σ ein fester Wert in der Maximierung von Q_1 zu sein. Die Maximierungswerte von

μ , Σ , γ können abgeleitet werden durch berechnen der partiellen Ableitungen, und setzen

die gleich 0. Die Maximierung von $Q_2(\lambda, \chi, \psi; \theta^{[k]})$ in Bezug auf Parameter von der

gemischten Verteilung wird numerisch durchgeführt. $Q_2(\lambda, \chi, \psi; \theta^{[k]})$ ist

$$\begin{aligned}
& (\lambda - 1) \sum_{i=1}^n \xi_i^{[1]} - \frac{1}{2} \chi \sum_{i=1}^n \delta_i^{[1]} - \frac{1}{2} \psi \sum_{i=1}^n \eta_i^{[1]} \\
& \quad - \frac{1}{2} n \lambda \ln(\chi) + \frac{1}{2} n \lambda \ln(\psi) - n \ln(2K_\lambda(\sqrt{\chi\psi})). \quad (3.38)
\end{aligned}$$

Wir nehmen an, dass die Parameter μ, Σ, γ zuerst in der Iteration aktualisiert werden. Wir definieren

$$\theta^{[k,2]} = (\lambda^{[k]}, \chi^{[k]}, \psi^{[k]}, \mu^{[k+1]}, \Sigma^{[k+1]}, \gamma^{[k+1]})'$$

Rechnen die Gewichte $\delta_i^{[k,2]}, \eta_i^{[k,2]}, \xi_i^{[k,2]}$ in (3.37) nach, dann maximieren

$Q_2(\lambda, \eta, \xi; \theta^{[k,2]})$ in (3.38). Die ergibt sich in einen sogenannten MCECM algorithmus.

Algorithmus 3.14 (EM Schätzen der verallgemeinerten hyperbolischen Verteilung)

(1). Setze $k=1$ und suche Startwert aus für $\theta^{[1]}$. Angemessene Startwert für μ, γ, Σ sind der Stichproben Erwartungswert, null Vektor und die Stichproben Kovarianz Matrix S .

(2) Berechne Gewichte $\delta_i^{[k]}, \eta_i^{[k]}$. Mitteln der Gewichte

$$\bar{\delta}^{[k]} = n^{-1} \sum_{i=1}^n \delta_i^{[k]} \quad \text{and} \quad \bar{\eta}^{[k]} = n^{-1} \sum_{i=1}^n \eta_i^{[k]}.$$

(3). Für eine symmetrisches Modell setzen wir $\gamma^{[k+1]}=0$. Sonst setzen

$$\gamma^{[k+1]} = \frac{n^{-1} \sum_{i=1}^n \delta_i^{[k]} (\bar{X} - X_i)}{\bar{\delta}^{[k]} \bar{\eta}^{[k]} - 1}.$$

(4) Aktualisieren den Schätzer von Mittelwertvektor und Varianzmatrix

$$\begin{aligned}
\mu^{[k+1]} &= \frac{n^{-1} \sum_{i=1}^n \delta_i^{[k]} X_i - \gamma^{[k+1]}}{\bar{\delta}^{[k]}}, \\
\Psi &= \frac{1}{n} \sum_{i=1}^n \delta_i^{[k]} (X_i - \mu^{[k+1]})(X_i - \mu^{[k+1]})' - \bar{\eta}^{[k]} \gamma^{[k+1]} \gamma^{[k+1]'}, \\
\Sigma^{[k+1]} &= \frac{|S|^{1/d} \Psi}{|\Psi|^{1/d}}.
\end{aligned}$$

(5.) setze

$$\theta^{[k,2]} = (\lambda^{[k]}, \chi^{[k]}, \psi^{[k]}, \mu^{[k+1]}, \Sigma^{[k+1]}, \gamma^{[k+1]})'$$

und berechnen Gewichte $\delta_i^{[k,2]}, \eta_i^{[k,2]}$ und $\xi_i^{[k,2]}$.

(6). Maximieren $Q_2(\lambda, \chi, \psi; \theta^{[k,2]})$ in Bezug auf λ, χ , und ψ um die Berechnung von

$\theta^{[k,2]}$ zu beenden. Erhöhen Iterationsschritt $k \rightarrow k+1$ und gehen zu Schritt (2).

3.2.5 Empirische Beispiele

In diesem Abschnitt passen wir die multivariate verallgemeinerte hyperbolische (GH) Verteilung an wahren Daten an und untersuchen welche der Unterklassen - wie t, hyperbolisch oder NIG - die nützlichsten sind. Wir untersuchen auch ob das allgemeine Mittelwert-Varianz Mischmodell durch (elliptisch symmetrisch) mixture Modell ersetzt werden kann. Unser erstes Beispiel bildet die Grundlage für multivariate Beispiele indem wir uns kurz die univariaten Modelle ansehen.

Beispiel 3.15 (univariate Aktien returns)

In der Literatur waren die NIG, hyperbolisch und t Modelle teilweise populäre Spezialfälle. Wir passen die symmetrischen und asymmetrischen Fälle dieser Verteilungen an die im Beispiel 3.3 verwendeten Daten an. Dabei beschränken wir unsere Aufmerksamkeit auf die täglichen und wöchentlichen returns bei denen reichlich Daten vorliegen ($n=2020$ und $n=468$). Die Modelle werden unter Verwendung des maximum likelihood, mit der einfachen Annahme das returns iid Beispiele formen, angepasst; Eine einfache quasi-Newton Methode ist eine gute Alternative zu dem EM Algorithmus im univariaten Fall.

In den beiden oberen Abschnitten von Tabelle 3.3 zeigen wir die Ergebnisse der Symmetrischen Modelle. Die NIG, hyperbolische und t Modelle können direkt unter Verwendung des log-likelihood beim Maximum verglichen werden, da alle die gleiche Anzahl an Parametern besitzen. Für tägliche Daten haben wir herausgefunden das 8 von 10 returns die t Verteilung der hyperbolischen und NIG Verteilung vorziehen; Für wöchentliche returns wird die t Verteilung in 6 von 10 Fällen bevorzugt. Zusammengefasst scheint die NIG Verteilung das zweitbeste Modell zu sein. Die gemixten Modelle passen in allen Fällen besser als das Gaußsche Modell und es kann einfach mittels des Akaike Informations Kriterium (AIC) verifiziert werden.

Für Asymmetrische Modelle zeigen wir nur die Fälle in denen mindestens eines der asymmetrischen NIG, hyperbolisch und t Modelle eindeutige Verbesserungen ($p < 0.05$) gegenüber dem entsprechenden symmetrischen Modell, basierend auf dem likelihood ratio Test, zeigt. Dieses passierte für wöchentliche returns von der Citigroup (C) und Intel (INTC), jedoch nicht für tägliche returns. Für die Citigroup waren die p-Werte der Tests 0.06, 0.04 und 0.04 für die t, NIG und hyperbolischen Fälle; Für Intel waren die p-Werte in allen Fällen 0.01, was eine starke Asymmetrie erkennen lässt.

Im Fall von Intel haben wir die Dichten von verschiedenen angepassten asymmetrischen Verteilungen in einem Histogramm (Abbildung 3.4) überlagernd dargestellt. Ein Druck der log-Dichten längsseits zeigt die Unterschiede zwischen den Verteilungen im Endbereich. Das linke Ende (entsprechend der Verluste) scheint stärker für diese Daten zu sein und die best passende Verteilung des likelihood Vergleiches ist die asymmetrische t Verteilung.

Beispiel 3.16 (multivariate stock returns)

Wir haben die multivariaten Modelle an das volle 10 dimensionale Datenset der log-returns aus dem vorherigen Beispiel angepasst. Die Ergebnisse des maximierten log-likelihood werden in Tabelle 3.4, zusammen mit den p-Werten für einen likelihood ratio test mit allen speziellen Fällen gegenüber dem (asymmetrisch) verallgemeinerte hyperbolische (GH), gezeigt. Die Anzahl der Parameter jedes Modells ist also gegeben. Beachten Sie das das allgemeine d-dimensionale GH Model $1/2d(d+1)$ dispersion parameter, d Schiefe parameter und drei parameter von der GIG

Mischerteilung hat, jedoch Gegenstand zu einer erkennbaren Einschränkung ist; Das sind $1/2(d(d+5)+4)$ freie Parameter.

Table 3.3

Stock	Gauss	<i>t</i> model		NIG model		Hyperbolic model	
	$\ln L$	ν	$\ln L$	$\sqrt{\chi\psi}$	$\ln L$	$\sqrt{\chi\psi}$	$\ln L$
<i>Daily returns: symmetric models</i>							
AXP	4945.7	5.8	5001.8	1.6	5002.4	1.3	5002.1
EK	5112.9	3.8	5396.2	0.8	5382.5	0.6	5366.0
BA	5054.9	3.8	5233.5	0.8	5229.1	0.5	5221.2
C	4746.6	6.3	4809.5	1.9	4806.8	1.7	4805.0
KO	5319.6	5.1	5411.0	1.4	5407.3	1.3	5403.3
MSFT	4724.3	5.8	4814.6	1.6	4809.5	1.5	4806.4
HWP	4480.1	4.5	4588.8	1.1	4587.2	0.9	4583.4
INTC	4392.3	5.4	4492.2	1.5	4486.7	1.4	4482.4
JPM	4898.3	5.1	4967.8	1.3	4969.5	0.9	4969.7
DIS	5047.2	4.4	5188.3	1	5183.8	0.8	5177.6
<i>Weekly returns: symmetric models</i>							
AXP	719.9	8.8	724.2	3.0	724.3	2.8	724.3
EK	718.7	3.6	765.6	0.7	764.0	0.5	761.3
BA	732.4	4.4	759.2	1.0	758.3	0.8	757.2
C	656.0	5.7	669.6	1.6	669.3	1.3	669
KO	757.1	6.0	765.7	1.7	766.2	1.3	766.3
MSFT	671.5	6.3	683.9	1.9	683.2	1.8	682.9
HWP	627.1	6.0	637.3	1.8	637.3	1.5	637.1
INTC	595.8	5.2	611.0	1.5	610.6	1.3	610
JPM	681.7	5.9	693.0	1.7	692.9	1.5	692.6
DIS	734.1	6.4	742.7	1.9	742.8	1.7	742.7
<i>Weekly returns: asymmetric models</i>							
C	NA	6.1	671.4	1.7	671.3	1.3	671.2
INTC	NA	6.3	614.2	1.8	613.9	1.7	613.3

Für die täglichen Daten ist der beste Spezialfall die verdrehte *t* Verteilung, welche einen Wert für das maximum likelihood ergibt der nicht erkennbar, durch das allgemeinere Modell mit seinen zusätzlichen Parametern, verbessert werden kann. Allen anderen nicht elliptischen symmetrischen Untermodelle werden bei einem likelihood ratio Test abgelehnt. Jedoch kann die elliptische symmetrische *t* Verteilung nicht abgelehnt werden wenn sie mit den meisten gängigen Modellen verglichen wird, so das es ein einfaches sparsames Modell für diese Daten zu sein scheint (der geschätzte Freiheitsgrad ist 6.0).

Für wöchentliche Daten ist der beste Spezialfall die NIG Verteilung, die gefolgt von der verdrehte *t* Verteilung. Die hyperbolische und variance gamma werden abgelehnt. Der beste elliptische symmetrische Spezialfall scheint die *t* Verteilung zu sein (mit geschätztem Freiheitsgrad 6.2).

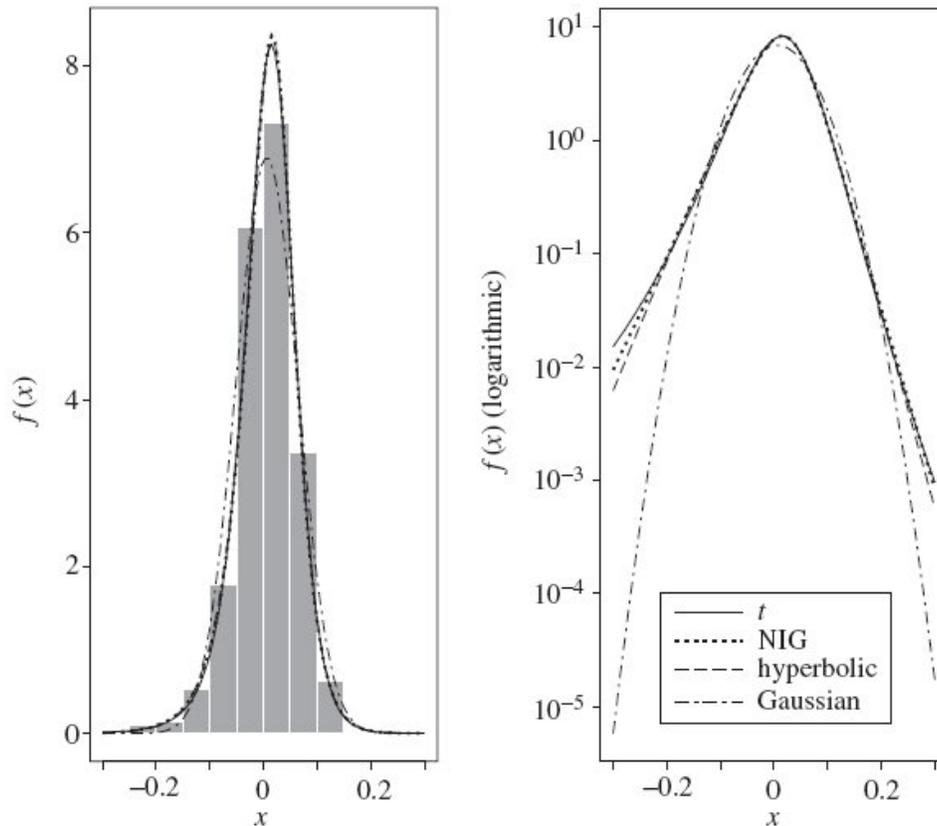


Figure 3.4. Models for weekly returns on Intel (INTC).

Beispiel 3.17 (multivariate Währungskurs returns)

Wir haben die selben multivariate Modelle an ein vier-dimensionales Datenset von Währungskurs log-returns angepasst. Dies waren GB Pfund, Euro, japanische Yen und schweizer Franken gegenüber dem US Dollar für den Zeitraum von Januar 2000 bis Ende März 2004 (1067 tägliche returns und 222 wöchentliche returns). Die Ergebnisse des maximierten log-likelihood sind in Tabelle 3.5 abgebildet.

Für die täglichen Daten ist der beste Spezialfall (im allgemeinen und sogar begrenzt auf symmetrische Modelle) die NIG Verteilung gefolgt von der hyperbolischen t und variance-gamma (VG) Verteilung. In einem likelihood ratio test von den Spezialfällen gegenüber einer allgemeinen GH Verteilung musste nur das VG Modell auf dem 5% Level abgewiesen werden. Das verdrehte t Modell wird auf dem 10% Level abgelehnt. Im erneuten Vergleich zum vollen Modell konnten sichere elliptische Modelle nicht zurückgewiesen werden dabei ist das beste von ihnen das NIG Modell.

Für wöchentliche Daten ist der beste Spezialfall die t Verteilung gefolgt von der NIG, hyperbolisch und variance gamma Verteilung. Keiner von diesen Spezialfällen kann durch einen 5% Level Test abgelehnt werden obwohl das VG Modell mit dem 10% Level Test abgelehnt werden kann.

Unter den elliptisch symmetrischen Verteilungen ist die Gaußsche Verteilung klar abgelehnt und die VG ist erneut auf dem 10% Level abgelehnt, andererseits jedoch werden die elliptischen Spezialfälle akzeptiert; Die beste unter ihnen scheint die t Verteilung zu sein (mit einem geschätzten Freiheitsgrad von 5.99).

Table 3.4. A comparison of models in the GH family for 10-dimensional stock-return data. For each model, the table shows the value of the log-likelihood at the maximum ($\ln L$), the numbers of parameters (# par.) and the p -value for a likelihood ratio test against the general GH model. The log-likelihood values for the general model, the best special case and the best elliptically symmetric special case are in bold type. See Example 3.16 for details.

	GH	NIG	Hyperbolic	t	VG	Gauss
<i>Daily returns: asymmetric models</i>						
$\ln L$	52 174.62	52 141.45	52 111.65	52 174.62	52 063.44	
# par.	77	76	76	76	76	
p -value		0.00	0.00	1.00	0.00	
<i>Daily returns: symmetric models</i>						
$\ln L$	52 170.14	52 136.55	52 106.34	52 170.14	52 057.38	50 805.28
# par.	67	66	66	66	66	65
p -value	0.54	0.00	0.00	0.63	0.00	0.00
<i>Weekly returns: asymmetric models</i>						
$\ln L$	7 639.32	7 638.59	7 636.49	7 638.56	7 631.33	
p -value		0.23	0.02	0.22	0.00	
<i>Weekly returns: symmetric models</i>						
$\ln L$	7 633.65	7 632.68	7 630.44	7 633.11	7 625.4	7 433.77
p -value	0.33	0.27	0.09	0.33	0.00	0.00

Table 3.5. A comparison of models in the GH family for four-dimensional exchange-rate return data. For each model, the table shows the value of the log-likelihood at the maximum ($\ln L$), the numbers of parameters (# par.) and the p -value for a likelihood ratio test against the general GH model. The log-likelihood values for the general model, the best special case and the best elliptically symmetric special case are in bold type. See Example 3.17 for details.

	GH	NIG	Hyperbolic	t	VG	Gauss
<i>Daily returns: asymmetric models</i>						
$\ln L$	17 306.44	17 306.43	17 305.61	17 304.97	17 302.5	
# par.	20	19	19	19	19	
p -value		0.85	0.20	0.09	0.00	
<i>Daily returns: symmetric models</i>						
$\ln L$	17 303.10	17 303.06	17 302.15	17 301.85	17 299.15	17 144.38
# par.	16	15	15	15	15	14
p -value	0.15	0.24	0.13	0.10	0.01	0.00
<i>Weekly returns: asymmetric models</i>						
$\ln L$	2 890.65	2 889.90	2 889.65	2 890.65	2 888.98	
p -value		0.22	0.16	1.00	0.07	
<i>Weekly returns: symmetric models</i>						
$\ln L$	2 887.52	2 886.74	2 886.48	2 887.52	2 885.86	2 872.36
p -value	0.18	0.17	0.14	0.28	0.09	0.00