Das fQ-System

DISSERTATION

Zur Erlangung des Grades
eines Doktors der Naturwissenschaften
der Universität Dortmund

Dem Fachbereich Statistik der Universität Dortmund
vorgelegt von
AXEL SCHEFFNER
aus Hagen

Dortmund, 1998

Tag der mündlichen Prüfung: 14. 5. 1998

Inhaltsverzeichnis

Kapitel 1

Einleitung und Überblick

Die Kernidee dieser Arbeit besteht aus der Zusammenführung des Konzepts der *Verteilungssysteme* mit dem der *Dichte-Quantilfunktion* zu einem neuen System, dem *fQ-System*, welches eine parametrische Darstellung der fQ-Funktion nutzt.
Die Dichte-Quantilfunktion, kurz die *fQ-Funktion*, wurde erstmals von Parzen (1979b), der sie zur Klassifikation des Randverhaltens von Verteilungen nutzte, untersucht. Darüber hinaus liefern die fQ-Funktionen, die aufgrund ihrer Konstruktion allgemein als Definitionsbereich das Intervall $[0, 1]$ besitzen, auf kanonische Weise eine Umskalierung, mit der die Gestalt unterschiedlichster Verteilungen verglichen werden können.
Verteilungssysteme sind Familien von Verteilungsfunktionen, in denen durch flexible, vielparametrische funktionale Darstellungen eine große Klasse unterschiedlicher Verteilungstypen vereint werden können. Diese Systeme stellen eine Zwischenstufe der parametrischen und nichtparametrischen Statistik dar.
Die Vereinigung dieser Begriffe erfolgt bei der Definition des *fQ-Systems* mittels eines parametrischen Ansatzes für die fQ-Funktion in der Form einer verallgemeinerten Beta-Verteilung. Dadurch wird eine neue Familie von Verteilungen konstruiert, mit der die Darstellung stetiger, unimodaler Verteilungen anhand ihrer fQ-Funktionen ermöglicht wird. Wie das motivierende Beispiel der Normalverteilung im zweiten Kapitel zeigt, kann dabei die exakte fQ-Funktion sehr gut approximiert werden.
In den Kapiteln des ersten Teils dieser Arbeit werden theoretische Aussagen über das fQ-System als eigenständige Klasse von Verteilungen getroffen. Bei

der Beschreibung der elementaren Eigenschaften im dritten Kapitel fällt dabei besonders die Tatsache ins Gewicht, daß es keine direkte Darstellung der Dichte-, Quantil- und Verteilungsfunktion gibt und demzufolge die Aussagen über diese Funktionen von allgemeinerer Form (Existenz von Momenten, Randverhalten) sind.

Diesem Umstand wird ebenfalls im vierten Kapitel Rechnung getragen, in dem die notwendigen numerischen Verfahren zur expliziten Darstellung der Quantil-, Verteilungs- und Dichtefunktion entwickelt werden.

Im fünften Kapitel werden die strukturellen Eigenschaften des fQ-Systems untersucht und die dabei vorgeschlagenen Ordnungsbegriffe, Metriken und Kennmaße durch Bedingungen an die fQ-Parameter ausgedrückt.

Die Ausführungen in den folgenden beiden Kapiteln lösen sich von Darstellung des fQ-Systems als solchem und erläutern zunächst in Kapitel 6, wie bekannte unimodale Verteilungen formal im fQ-System dargestellt werden können.

Im anschließenden wesentlichen siebten Kapitel wird die Methodik zur Anpassung von fQ-Verteilungen an empirische Daten geschildert. Es zeigt sich, daß das vorgeschlagene Verfahren durch eine Trennung in einen linearen und einen nichtlinearen Anteil der Schätzung zu einer Erweiterung der bekannten Theorie der approximativen L-Schätzer führt.

Bei den Anwendungen im zweiten, praktischen Teil wird nach einer einführenden Übersicht im achten Kapitel anhand dreier ausgewählter Verteilungen zunächst im neunten Kapitel die Qualität der Approximation der fQ-Darstellung mit graphischen Mitteln verdeutlicht und quantifiziert.

Unter Ausnutzung dieser Ergebnisse können im Kapitel 10 unter Verwendung eines theoretischen Ergebnisses von Blom (1958) Schätzungen für die Erwartungswerte von Ordnungsstatistiken mit Hilfe des fQ-Systems konstruiert werden.

Die für die Praxis wichtigen Resultate sind in den beiden abschließenden Kapiteln zu finden. Dabei wird in Kapitel 11 zunächst anhand von Simulationsstudien die Effizienz der fQ-ABLU-Schätzer mit denen anderer Verfahren verglichen. Schließlich wird die Qualität der fQ-Dichteschätzung in Kapitel 12 durch einen Vergleich mit den Resultaten einer Kernschätzung für eine Reihe von Beispielen verdeutlicht.

Teil I

Theoretischer Teil

Kapitel 2

Genesis

2.1 Graphische Motivation des fQ-Systems

Beim Vergleich stetiger Verteilungen ist zu beobachten, daß sich einerseits ihre Dichtekurven graphisch kaum unterscheiden, während andererseits die zugehörigen funktionalen Darstellungen vollkommen verschiedene Formen haben.
Als Beispiele dazu mögen hier die Standardnormalverteilung und die t-Verteilung mit 6 Freiheitsgraden dienen:

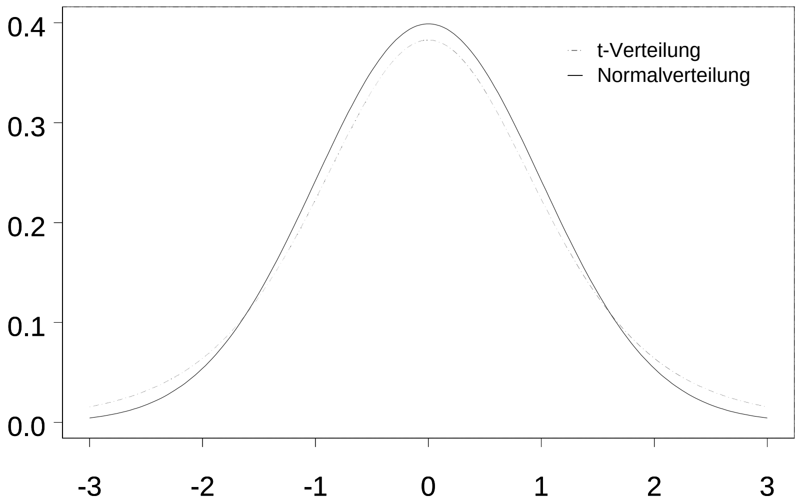

Abbildung 2.1: Dichte der Normalverteilung und der t_6-Verteilung

Die Ähnlichkeit der beiden Kurven steht im Gegensatz zu den entsprechenden Dichtefunktionen

$$\begin{aligned} f_{NV}(x) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad \text{und} \\ f_{t;6}(x) &= \frac{\Gamma(\frac{7}{2})}{\Gamma(3)\sqrt{6\pi}} \left(1 + \frac{x^2}{6}\right)^{-\frac{7}{2}}, \end{aligned}$$

die keine offensichtliche Äquivalenz aufweisen.

Um dennoch derartige Dichtefunktionen analytisch miteinander vergleichen zu können, wurden in der Vergangenheit verschiedene Ansätze verfolgt, durch Umparametrisierungen und zusätzliche Parameter verallgemeinerte Dichtefunktionen zu erhalten, die eine möglichst große Klasse von Verteilungen umfassen. Die bekanntesten Vetreter dieser Methode sind die Systeme von Pearson (1895), Johnson (1949) und die verallgemeinerten λ-Verteilung von Tukey (1962), die in Abschnitt 2.3.4 näher erläutert wird.

Gemeinsam ist diesen Systemen, daß sie von einer möglichst flexiblen funktionalen Darstellung der Dichtefunktion ausgehen und die Parameter der Verteilungen anhand von charakteristischen Merkmalen - wie etwa den Definitionsbereich oder die Existenz von Momenten - in verschiedene Typen einteilen.

So ist etwa das System von Pearson durch die Dichtefunktionen $y = f(x)$, die folgender Differentialgleichung genügen, definiert:

$$y' = \frac{x+m}{b_0 + b_1 x + b_2 x^2} y.$$

Für die Unterscheidung verschiedener Dichtetypen in diesem System werden die Koeffizienten b_i und die Nullstellen a_1, a_2 des Polynoms

$$b_0 + b_1 x + b_2 x^2$$

herangezogen, wobei Pearson insgesamt 9 verschiedene Klassen unterscheidet. Eine vollständige Auflistung dieser Klassen in Verbindung mit einer graphischen Darstellung der Parameterbereiche ist zum Beispiel bei Johnson, Kotz und Balakrishnan [1994, S. 16ff] zu finden.

Der wesentliche Vorteil des Systems liegt darin, daß viele Standardverteilungen exakt dargestellt werden können; so gilt beispielsweise

$$\begin{aligned} b_1 = b_2 = 0 &\ :\ \text{Normalverteilung,} \\ a_1 < 0 < a_2 &\ :\ \text{Betaverteilung,} \\ b_1 \neq 0, b_2 = 0 &\ :\ \text{Gammaverteilung.} \end{aligned}$$

Das prinzipielle Problem bei allen genannten Systemen liegt darin, daß die Parametrisierungen beziehungsweise die Bedingungen an die Parameter für bestimmte Dichtetypen keine intuitive Vorstellung von der Gestalt der Kurve vermitteln können (siehe dazu auch die Anmerkungen zur λ-Verteilung in Abschnitt 2.3.4).
Demgegenüber wird bei dem hier vorgestellten fQ-System der umgekehrte Weg eingeschlagen: die funktionale Darstellung des Systems orientiert sich an der Zielsetzung, primär die Gestalt der Dichtekurven mittels geeigneter Parameter zu beschreiben; optisch ähnliche Kurven sollen also durch ähnliche Parameter repräsentiert werden. Zwangsläufig ergibt sich aus dieser Konstruktion, daß sich aus den Parametern direkte Aussagen über die Form der zugehörigen Dichte ableiten lassen.
Im folgenden seien hier kurz die Überlegungen, die zur Definition des fQ-Systems führen, beschrieben:
Die erste Schwierigkeit beim Vergleich von Dichtefunktionen ist der Umstand, daß die Träger der Funktionen

$$D_f := \{x \in \mathbb{R} : f(x) > 0\} \tag{2.1}$$

mit den Grenzen

$$\alpha(f) := \inf \{x \in \mathbb{R} : f(x) > 0\}, \tag{2.2}$$

$$\omega(f) := \sup \{x \in \mathbb{R} : f(x) > 0\} \tag{2.3}$$

abgeschlossene, halboffene oder offene Intervalle umfassen können. Somit ist ein direkter Vergleich der Kurven schon aufgrund der unterschiedlichen Definitionsbereiche schwierig.
Da für stetige Verteilungen F die Quantilfunktion

$$Q(p) := F^{-1}(p)$$

ebenfalls stetig ist, bietet diese die Möglichkeit, auf kanonische Weise die Definitionsbereiche aller stetigen Verteilungen auf das Intervall $[0,1]$ umzuskalieren, da gilt

$$\alpha(f) = Q(0) \quad \text{und} \quad \omega(f) = Q(1).$$

Statt der Dichtefunktion $f(x)$ soll also nun die Dichte-Quantilfunktion $f(Q(p))$ betrachtet werden, die im folgenden *fQ-Funktion* genannt und mit der abkürzenden Schreibweise $fQ(p)$ bezeichnet wird. Diese Terminologie geht zurück auf Parzen (1979a) (vgl. Abschnitt 2.3.2).
Zur Illustration seien hier die fQ-Funktionen von vier bekannten Verteilungen abgebildet:

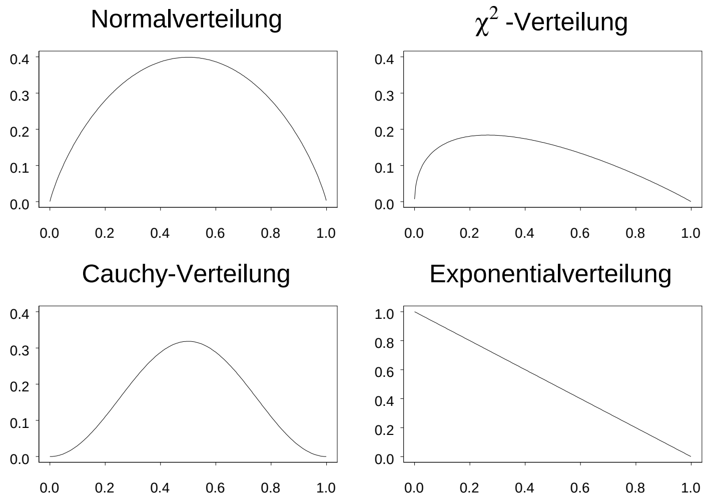

Abbildung 2.2: fQ-Funktionen von 4 Standardverteilungen

Ein optischer Vergleich dieser Kurven mit den Dichtefunktionen der Betaverteilung (siehe Johnson, Kotz und Balakrishnan [1995, S. 220]) legt die Idee nahe, die fQ-Funktionen durch die funktionale Form der Betadichte darzustellen:

$$fQ(p) = \sigma p^{\alpha}(1-p)^{\beta}, \quad p \in [0,1], \tag{2.4}$$

wobei der zusätzliche Parameter $\sigma > 0$ notwendig ist, da die Kurven der fQ-Funktion aufgrund der Umskalierung keinen Flächeninhalt von 1 aufweisen.

Um diese Darstellung mit den tatsächlichen fQ-Funktionen vergleichen zu können, sei hier eine Tabelle mit den fQ-Funktionen ausgewählter Verteilungen aufgeführt:

Name	$f(x)$	$fQ(p)$
Rechteck	$1, \quad x \in [0,1]$	$1, \quad p \in [0,1]$
Potenz	$c(x-1)^{c-1}, \quad x \in [0,1], c > 0$	$c(1-p)^{1-1/c}$
Exponential	$e^{-x}, \quad x > 0$	$1-p$
Logistisch	$\frac{e^x}{(1+e^x)^2}, \quad x \in \mathbb{R}$	$p(1-p)$
Weibull	$cx^{c-1}e^{-x^c}, \quad x > 0, c > 0$	$c(1-p)\log(\frac{1}{1-p})^{1-1/c}$
Extremwert	$e^{-x}e^{-e^x}, \quad x \in \mathbb{R}$	$(1-p)\log(\frac{1}{1-p})$
Normal	$(2\pi)^{-1/2}\mathrm{e}^{-x^2/2}$	$(2\pi)^{-1/2}e^{\left(-1/2\vert\Phi^{-1}(p)\vert^2\right)}$
Cauchy	$(\pi(1+x^2))^{-1}, \quad x \in \mathbb{R}$	$\pi^{-1}(\cos(\pi(p-1/2)))^2$
Pareto	$(cx^{1+1/c})^{-1}, \quad x > 1, c > 0$	$c^{-1}(1-p)^{1+c}$

Tabelle 2.1: fQ-Funktionen ausgewählter Verteilungen

Aus der letzten Spalte dieser Tabelle ist offensichtlich zu erkennen, daß sich die 3-parametrische Klasse von Funktionen gemäß (2.4) nicht dazu eignet, beliebige fQ-Funktionen exakt darzustellen. Es zeigt sich aber, daß es möglich ist, Approximationen der Form (2.4) zu finden, die eine hohe Anpassungsgüte besitzen.

Als Beispiel sei die Approximation der fQ-Funktion der Normalverteilung fQ_{NV} durch eine Funktion fQ^* der Gestalt (2.4) betrachtet, die durch Minimierung des fQ-Abstandes

$$fQ^* = \arg\min_{fQ} \left(\int_0^1 (fQ_{NV} - fQ)^2 \, dp \right)^{1/2}$$

gewonnen wurde (die Methodik dieser Vorgehensweise wird im Abschnitt 5.4.2 und in Kapitel 7 erläutert).

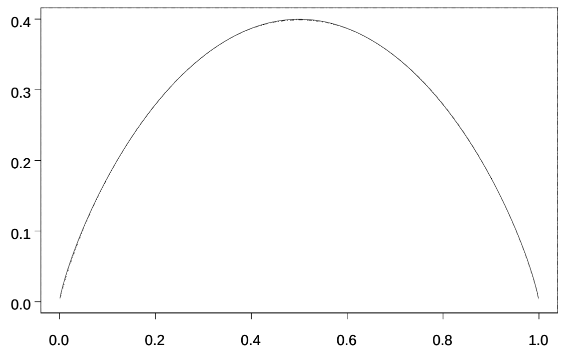

Abbildung 2.3: fQ-Funktion der Normalverteilung mit Approximation

Die Qualität der Approximation ist so hoch, daß nur ein weiteres Diagramm mit den Abständen der Kurven die Details offenbaren kann.

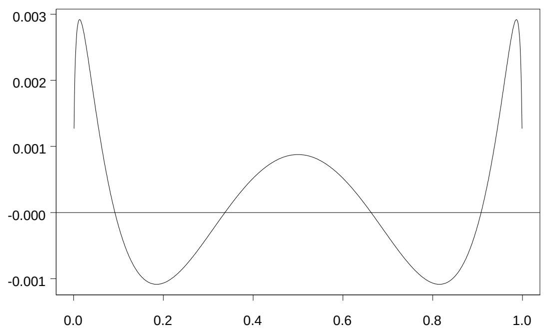

Abbildung 2.4: Güte der Approximation der Normalverteilung

Da sich dieses Bild einer guten Anpassungsqualität durchgehend für alle stetigen, unimodalen Verteilungen ergibt, ist damit die Rechtfertigung gegeben, sich von der Idee der reinen Approximation von beliebigen fQ-Funktionen zu trennen und alle Verteilungen, die sich in der Form (2.4) darstellen lassen, als eigenständiges System zu betrachten.
Im letzten Schritt der Vorüberlegungen muß berücksichtigt werden, daß für die Quantilfunktionen zweier Zufallsvariablen X und Y mit $Y := a + X, a \in \mathbb{R}$

$$Q_Y(p) = a + Q_X(p)$$

gilt, woraus

$$Q'_Y(p) = Q'_X(p)$$

folgt.
Aufgrund der Differentiationsregeln für Umkehrfunktionen gilt für die Ableitung der Quantilfunktion

$$Q'(p) = \frac{1}{f(Q(p))}, \tag{2.5}$$

und somit

$$fQ_Y(p) = fQ_X(p), \quad \forall a \in \mathbb{R}.$$

Somit ist die fQ-Funktion lokationsinvariant, d.h. unabhängig von einer Translation der Verteilung, also ist zur Fixierung der Dichte die Angabe eines Lokationsparameters μ notwendig. Da bei der definierenden Gleichung (2.4) die Quantilfunktion als innere Funktion der fQ-Funktion zuerst zur Anwendung kommt, ist die Wahl $\mu = Q(1/2)$ des Medians der Verteilung an dieser Stelle zumindest anschaulich gerechtfertigt. Daß diese Wahl auch analytisch sinnvoll ist, wird sich bei der konkreten Darstellung der Quantilfunktion in Abschnitt 3.2 erweisen.

2.2 Definition des fQ-Systems

Insgesamt lassen sich die Überlegungen im vorausgehenden Abschnitt zu folgender Definition zusammenfassen:

Definition 2.1
Eine reellwertige Zufallsvariable X heißt fQ-verteilt *mit Parametern $\sigma > 0$, $\alpha, \beta, \mu \in \mathbb{R}$*

$$X \sim \mathcal{F}(\mu, \sigma, \alpha, \beta),$$

wenn ihre fQ-Funktion die Gestalt

$$fQ(p) = \sigma p^{\alpha}(1-p)^{\beta}, \quad p \in [0,1] \tag{2.6}$$

hat und zusätzlich gilt

$$Q(1/2) = \mu. \tag{2.7}$$

Die Menge aller fQ-Verteilungen

$$\mathcal{FQ} := \{F = \mathcal{F}(\mu, \sigma, \alpha, \beta), \sigma > 0, \alpha, \beta, \mu \in \mathbb{R}\} \tag{2.8}$$

heißt das fQ-System.

Man beachte, daß die Gleichung (2.6) die Stetigkeit der fQ- Funktion und damit die der Dichtefunktion f impliziert. Somit gilt

$$\mathcal{FQ} \subset \mathcal{F}_S,$$

wobei $\mathcal{F}_S$ die Menge der stetigen reellwertigen Verteilungen bezeichnet. Weiterhin folgt aus (2.6) die Existenz der Ableitung der fQ- Funktion. Dies führt zu folgender Definition:

Definition 2.2
Sei $F \in \mathcal{FQ}$, dann heißt die Funktion

$$J(p) = -(fQ)'(p), \quad p \in (0,1) \tag{2.9}$$

die Scorefunktion *von F.*

Bemerkung 2.3
Durch nähere Betrachtung der Ableitung der fQ-Funktion

$$\begin{aligned} J(p) &= -(fQ)'(p) \\ &= -\frac{f'(Q(p))}{f(Q(p))} \end{aligned}$$

ergibt sich die in der Literatur übliche Form der Scorefunktion

$$J = -\frac{f'}{f}.$$

Die besondere Bedeutung dieser Funktion in vielen Bereichen der nichtparametrischen Statistik zeigt der Artikel von Joiner und Hall (1983).
Häufig jedoch führen die mit der Scorefunktion verbundenen Aussagen für theoretische Verteilungen zu komplizierten Ausdrücken, und die Schätzverfahren für empirische Untersuchungen zeigen oftmals schlechte Konvergenzeigenschaften. Demgegenüber ist innerhalb des fQ-Systems die analytische Darstellung der Scorefunktion trivial (siehe Lemma 3.5).

Terminologie

Um Verwechslungen bei den verwendeten Notationen zu vermeiden, seien hier die wichtigsten Bezeichnungen noch einmal zusammengefaßt:

- Die *fQ-Funktion* ist ein allgemeiner Begriff und gilt für beliebige stetige Verteilungen. Sie kann allerdings nicht für alle Verteilungen in geschlossener Form dargestellt werden.

- Eine *fQ-Verteilung* ist eine Verteilung, deren fQ-Funktion die speziellen Eigenschaften (2.6) und (2.7) erfüllt.

- Das *fQ-System* $\mathcal{FQ}$ besteht aus der Gesamtheit aller fQ-Verteilungen.

- Zur Beschreibung der Eigenschaften einer fQ-Verteilung werden unter anderem ihre funktionalen Darstellungen verwendet. Dies sind die *fQ-Dichte-*, die *fQ-Quantil-* und die *fQ-Verteilungsfunktion*. Der Begriff fQ-Verteilung bezeichnet somit die gesamte Verteilung als solche, im Gegensatz zur fQ-Verteilungsfunktion im üblichen Sinn.

Parametrisierungen

Elementare Umformungen der in Definition 2.1 verwendeten Parametrisierung führen zu alternativen Darstellungen der charakterisierenden Gleichung für die fQ-Verteilung, die in späteren Kapiteln benötigt werden:

$$\begin{aligned} fQ(p) &= \sigma p^{\alpha}(1-p)^{\beta}, p \in [0,1], && \text{(A)} \\ F'(x) &= \sigma F(x)^{\alpha}(1-F(x))^{\beta}, x \in \mathbb{R}, && \text{(B)} \\ Q'(p) &= \sigma^{-1}p^{-\alpha}(1-p)^{-\beta}, p \in [0,1]. && \text{(C)} \end{aligned}$$

Die Darstellung (A) wurde bei der Herleitung der Definiton des fQ-Systems benutzt, (B) wird zur Berechnung der Verteilungsfunktion F und der Dichte $f = F'$ und (C) zur Darstellung der Quantilfunktion herangezogen. In Kapitel 6.2 wird eine weitere Umparametrisierung vorgestellt, die auf einer äquivalenten Darstellung der fQ-Funktion durch Gestaltparameter (Schiefe und Dispersion) der Verteilung beruht. Da diese Begriffe spezielle Eigenschaften des fQ-Systems ausnutzen, die erst später beschrieben werden, sei hier auf das entsprechende Kapitel verwiesen.

Notationen

Da im folgenden häufig Aussagen getroffen werden, die nur für Verteilungen mit bestimmten analytischen Eigenschaften gelten, seien hier einige Notationen für die entsprechenden Verteilungsmengen eingeführt.
Gemeinsame Obermenge aller betrachteten Verteilungen ist die Menge der stetigen, reellen Verteilungsfunktionen $\mathcal{F}_S$. Eine wichtige Untermenge von $\mathcal{F}_S$ wird durch die *unimodalen* Verteilungen $\mathcal{F}_U$ gebildet, wobei der Begriff der Unimodalität hier rein graphisch verstanden wird (siehe auch Handl [1985, S. 65]):

Definition 2.4
Eine Verteilung $F \in \mathcal{F}_S$ heißt unimodal, *wenn ein eindeutig bestimmter Punkt x_{mod} im Inneren des Trägers D_f von Fexistiert, so daß die Dichte f für $x < x_{mod}$ streng monoton steigend und für $x > x_{mod}$ streng monoton fallend ist.*
Der Wert x_{mod} heißt der Modus *von f und $f(x_{mod})$ der* Modalwert *von f.*

Durch die Einschränkung des Modus auf das Innere des Trägers werden Verteilungen ausgeschlossen, die wie die Exponentialverteilung ihr Maximum am Rand annehmen. Um auch diese Verteilungen zu erfassen, wird folgende, schwächere Variante definiert:

Definition 2.5
Eine Verteilung $F \in \mathcal{F}_S$ heißt schwach unimodal, *wenn ein Punkt $x_{mod} \in D_f$ existiert, so daß die Dichte f für $x \leq x_{mod}$ monoton steigend und für $x \geq x_{mod}$ monoton fallend ist. Die Menge der schwach unimodalen Verteilungen sei mit $\mathcal{F}_{WU}$ bezeichnet.*

Neben der Exponentialverteilung ist in dieser Menge ebenfalls die Rechteckverteilung enthalten. Man beachte, daß der hier verwendete Begriff der Unimodalität sich von der folgenden Definition unterscheidet, in der eine analytische Eigenschaft benutzt wird (siehe Hájek und Šidák [1967, S. 15]):

Definition 2.6
Eine Verteilung $F \in \mathcal{F}_S$ heißt streng unimodal, *wenn für die Dichte f gilt*

$$\log f \text{ ist streng konvex.} \tag{2.10}$$

Die methodische Bedeutung dieser Definition liegt in der Aussage, daß die Faltung zweier streng unimodaler Verteilungen wieder streng unimodal ist (siehe Ibragimov [1956]). Allerdings werden durch diese schärfere Bedingung (2.10) wichtige Verteilungen wie zum Beispiel die Cauchyverteilung ausgeschlossen, deren Dichte in obigem graphischen Sinne unimodal sind. Aus diesem Grund wird im folgenden der Begriff der strengen Unimodalität gemäß Definition 2.6 nicht weiter untersucht.

Bei Aussagen im Zusammenhang mit der Scorefunktion J wird die Menge der stetigen Verteilungen mit endlicher Fisher-Information

$$I(F) := \int_{-\infty}^{\infty} \left(\frac{f'(x)}{f(x)} \right)^2 f(x)\,dx = \int_0^1 (J(p))^2\,dp \tag{2.11}$$

von Bedeutung sein, die mit dem Symbol $\mathcal{F}_I$ bezeichnet sei.

Ein wesentlicher Bestandteil dieser Arbeit besteht in der Untersuchung dieser Verteilungsmengen mit der zusätzlichen Einschränkung auf die fQ-Verteilungen. Demzufolge werden im folgenden die Untermengen

$$\mathcal{FQ}_U := \mathcal{FQ} \cap \mathcal{F}_U \tag{2.12}$$

$$\mathcal{FQ}_{WU} := \mathcal{FQ} \cap \mathcal{F}_{WU} \tag{2.13}$$

$$\mathcal{FQ}_I := \mathcal{FQ} \cap \mathcal{F}_I \tag{2.14}$$

benötigt.

2.3 Historische Einordnung

In diesem Abschnitt sollen zunächst zwei wesentliche Vorgängerarbeiten von Blom (1958) und Parzen (1979a, 1979b) zitiert werden, die sich mit der fQ-Funktion befassen. Aufgrund der bereits erwähnten besonderen Rolle der Scorefunktion gibt es viele weitere Artikel, die sich mit dieser Thematik auseinandersetzen, wobei aber bisher keine allgemeinen parametrischen Ansätze für die fQ-Funktion wie im fQ-System gemacht wurden.
Im Anschluß daran werden zwei bekannte Verteilungssysteme vorgestellt, deren definierenden Gleichungen Ähnlichkeit mit der des fQ-System aufweisen, die aber deutliche strukturelle Unterschiede haben.

2.3.1 Die transformierten Beta-Variablen

Die vermutlich erste Erwähnung der Gestalt des fQ-Systems erfolgt im Rahmen einer Untersuchung von Blom (1958) über das Verhalten von *transformierten Beta-Variablen*, wobei Blom darunter zunächst ganz allgemein monotone Funktionen G versteht, die auf (betaverteilte) Ordnungsstatistiken von rechteckverteilten Zufallsvariablen angewendet werden. Neben verschiedenen anderen speziellen Klassen für die Funktion G führt er den Begriff der *B-Transformierten (beta-transform)* ein, den er durch die Beziehung

$$G'(u) = c_0 u^{-\tau_1}(1-u)^{-\tau_2}, \quad \tau_1, \tau_2 \in [0,1]$$

definiert.

Offensichtlich entspricht dies der Parametrisierung (C) des fQ-Systems, wobei Blom diese Gleichung jedoch nicht als Definition eines Verteilungssystems interpretiert und vor allem nicht die Beziehung zur Dichte-Quantilfunktion nutzt. Er erwähnt lediglich, daß mit obiger Form einige bekannte Verteilungen (Rechteck-, Expontential-, Logistische und Arcus-Sinus-Verteilung) erfaßt werden können.

Da Bloms Arbeit aus dem Jahr 1958 stammt, und er demzufolge mangels notwendiger Computerunterstützung nicht in der Lage war, die numerischen Probleme der Differentialgleichung in der Praxis zu lösen, zieht er die B-Transformierten nur für wenige, theoretische Aussagen heran, von denen eine im Abschnitt 6.3 dieser Arbeit für die Erwartungswerte von Ordnungsstatistiken direkt genutzt wird.

2.3.2 Die Untersuchungen von Parzen

Die umfangreichen Arbeiten von Parzen (1979a, 1979b) stellen die eigentliche Geburtsstunde der fQ-Funktion dar und liefern für weite Teile der vorliegenden Arbeit sowohl technische Hilfsmittel als auch viele methodische Anregungen, die in den entsprechenden Kapiteln näher aufgeschlüsselt werden.

Der wesentliche Unterschied des Ansatzes von Parzen zum fQ-System besteht darin, daß Parzen von keiner parametrischen Grundannahme für die fQ-Funktion ausgeht. Demzufolge schlägt er zur Schätzung der fQ-Funktion und deren Ableitung die nichtparametrischen Methoden der Kernschätzung und der Schätzung einer autoregressiven Darstellung von fQ' vor.

Beide Verfahren liefern aufgrund des nichtparametrischen Charakters der Schätzmethode keine geschlossene Darstellung der fQ-Funktion und können somit nur in Verbindung mit weiteren Teststatistiken genutzt werden.

2.3.3 S-Verteilungen

Die Klasse der *S-Verteilungen* (Voit [1992]) basiert auf einer Vereinfachung der *S-Systeme*, die ursprünglich konzipiert wurden, um komplizierte biologische und biochemische Systeme zu beschreiben. Die dort auftretenden Interaktionen werden dabei durch eine Menge von n nichtlinearen Differentialgleichungen erster

Ordnung modelliert:

$$\frac{dF_i}{dt} = a_i \prod_{j=1}^{n} F_j^{\beta_{ij}} - b_i \prod_{j=1}^{n} F_j^{\gamma_{ij}}, \quad i = 1, \ldots, n,$$

wobei $a_i, b_i > 0, \beta_{ij}, \gamma_{ij}$ reelle Konstanten und F_i stetige Funktionale sind.
Voit kann zeigen, daß nahezu jedes System gewöhnlicher Differentialgleichungen auf ein S-System zurückgeführt werden kann. Durch die spezielle Wahl der Ableitungen

$$F_i = F^{(i-1)}$$

für eine Verteilungsfunktion F ist damit ebenfalls eine exakte Darstellung von Standardverteilungen möglich, wobei die resultierende Parametermenge zum Teil sehr umfangreich und für die Praxis unbrauchbar ist.
Die naheliegende Einschränkung auf nur eine Differentialgleichung führt dann zu folgender Darstellung

$$f(x) = \alpha(F(x)^{\beta} - F(x)^{\gamma}), \tag{2.15}$$

wobei wie auch beim fQ-System die Angabe eines Lokationsparameters durch Vorgabe eines Funktionswertes $F(x_0) =: F_0$ notwendig ist.
Trotz der Ähnlichkeit von (2.15) mit der Definiton des fQ-Systems in der Darstellung (B) weisen diese beiden Verteilungsklassen deutliche Unterschiede auf, die auf den verschiedenen Konstruktionsprinzipien und Parametrisierungsarten beruhen. Insbesondere ist eine einfache Interpretation der S-Parameter schon aufgrund der Definitonsgleichung erschwert:

- Aus analytischen Überlegungen (Voit [1992, S. 858]) folgt, daß für die Parameter der S-Verteilung $\beta < \gamma$ gelten muß, damit eine Dichtefunktion vorliegt. Dies schränkt auf unnatürliche Weise den Grundraum der Parameter ein.

- Die Symmetriebedingung für eine S-Verteilung lautet (Voit [1992, S. 860])

$$\left(\frac{\beta}{\gamma}\right)^{\frac{1}{\gamma-\beta}} = 0.5.$$

  Diese Gleichung liefert auch noch für wenig symmetrisch erscheinende Parameterkonstellation wie beispielsweise ($\beta = 0.5484, \gamma = 3$) einen Wert von 0.49999 und ist somit wenig intuitiv.

2.3.4 Die verallgemeinerte λ-Verteilung

Die eindimensionale λ-Verteilung geht auf einen Vorschlag von Tukey (1962) zurück und wird von Ramberg und Schmeiser (1972, 1974) durch die Einführung weiterer Parameter zu folgender Form der Quantilfunktion

$$Q_\lambda(p) = \lambda_1 + \frac{p^{\lambda_3} - (1-p)^{\lambda_4}}{\lambda_2}$$

erweitert. Durch Differentiation folgt für die fQ-Funktion

$$fQ_\lambda(p) = \frac{\lambda_2}{\lambda_3 p^{\lambda_3 - 1} + \lambda_4 (1-p)^{\lambda_4 - 1}}. \tag{2.16}$$

Wie auch im fQ-System wird die fQ-Funktion hier durch Terme der Form p^a und $(1-p)^b$ ausgedrückt, wobei die Summation dieser Ausdrücke wie auch schon bei den S-Verteilungen zu komplizierten Parameterkonstellation führt. So zeigen Karian et. al (1996), daß (2.16) genau dann eine Dichtefunktion darstellt, wenn die Bedingung

$$fQ_\lambda(p) \geq 0$$

erfüllt ist. Eine vereinfachte äquivalente Charakterisierung dieser Aussage durch Bedingungen an einzelne Parameter λ_i ist allerdings nur mittels einer Graphik möglich, in der die zulässigen Paramterbereiche markiert sind (Karian et. al [1996]).

Zudem bedingt die vorgeschlagene Methode zur Parameterschätzung durch die Momentenmethode große technische Schwierigkeiten, da auch hier die Existenz der benötigten Momente von den zu schätzenden Parametern abhängt.

Trotz dieser methodischen Probleme werden die verallgemeinerten λ-Verteilungen häufig im Bereich der nichtparametrischen Statistik eingesetzt, da sie eine Schätzung der fQ-Funktion und der Scorefunktion erlauben (Handl [1985], Büning [1991]).

Kapitel 3

Elementare Eigenschaften

3.1 Die fQ-Funktion

Bevor auf die speziellen Eigenschaften der fQ-Funktion innerhalb des fQ-Systems eingegangen wird, seien hier einige Ergebnisse zusammengetragen, die für die fQ-Funktionen beliebiger Verteilungen gelten. Für die explizite Darstellung von fQ-Funktionen ausgewählter Verteilungen sei auf Tabelle 2.1 verwiesen.

Die allgemeine fQ-Funktion

Das folgende Lemma stellt zunächst alle notwendigen Beziehungen zwischen den zentralen Funktionen f,F,Q und J zusammen:

Lemma 3.1
Sei $F \in \mathcal{F}_S$, dann gelten für $p \in (0,1)$ folgende Aussagen:

$$F(Q(p)) = p \tag{3.1}$$

$$fQ(p) = Q'(p)^{-1}, \tag{3.2}$$

$$J(p) = \frac{-f'(Q(p))}{fQ(p)} = \frac{-f'(F^{-1}(p))}{f(F^{-1}(p))}. \tag{3.3}$$

Beweis:

Die erste Aussage folgt aus der Tatsache, daß für eine Zufallsvariable X mit stetiger Verteilungsfunktion gilt:

$$X \sim Q(U),$$

wobei U eine auf $[0, 1]$ rechteckverteilte Zufallsvariable ist. (3.2) und (3.3) folgen aus den Regeln für die Ableitungen von Umkehrfunktionen beziehungsweise zusammengesetzten Funktionen. □

Da für viele Umformungen das Verhalten der Quantilfunktion unter Transformationen von besonderer Bedeutung ist, sind hier die wichtigsten Ergebnisse aufgelistet:

Lemma 3.2

Sei $X \sim F \in \mathcal{F}_S$ eine stetige Zufallsvariable mit Quantilfunktion Q_X und g eine montone, reellwertige Funktion. Dann gilt für die Quantilfunktion Q_Y der Zufallsvariablen $Y = g(X)$:

$$Q_Y(p) = \begin{cases} g\left(Q_X(p)\right) & , \text{ falls } g \text{ steigend} \\ g\left(Q_X(1-p)\right) & , \text{ falls } g \text{ fallend.} \end{cases}$$

Ist g stetig, so gilt:

$$E[g(X)] = E[g(Q(U))] = \int_0^1 g(Q(p))\, dp$$

und somit insbesondere

$$\mu_k := E[X^k] = \int_0^1 (Q(p))^k \, dp. \tag{3.4}$$

Beweis:

Alle Aussagen folgen direkt aus den entsprechenden Transformationssätzen für Verteilungsfunktionen beziehungsweise durch elementare Substitution. □

Die Beziehung (3.4) wird es später ermöglichen, Existenzaussagen für die Momente von fQ-verteilten Zufallsvariablen auf Untersuchungen des Randverhaltens der Quantilfunktion zurückzuführen.

Für den Speziallfall einer affinen Transformation gilt für die Scorefunktion:

Lemma 3.3
Seien $F, G \in \mathcal{F}_S$ mit

$$F(x) = G\left(\frac{x-\mu}{\sigma}\right), \ \sigma > 0, \mu \in \mathbb{R}.$$

Dann gilt:

$$J_F(p) = \frac{J_G(p)}{\sigma}, \quad p \in (0,1). \tag{3.5}$$

Beweis:
Zunächst gelten folgende Beziehungen für die zugehörigen Dichtefunktionen f und g beziehungsweise die Quantilfunktionen Q_F und Q_G:

$$\begin{aligned} f(x) &= \frac{1}{\sigma}\, g\left(\frac{x-\mu}{\sigma}\right), \\ f'(x) &= \frac{1}{\sigma^2}\, g'\left(\frac{x-\mu}{\sigma}\right), \\ Q_F(p) &= \mu + \sigma Q_G(p). \end{aligned}$$

Daraus folgt:

$$\begin{aligned} J_F(p) &= -\frac{f'(Q_F(p))}{f(Q_F(p))} &&= -\frac{1}{\sigma}\frac{g'\left(\frac{Q_F(p)-\mu}{\sigma}\right)}{g\left(\frac{Q_F(p)-\mu}{\sigma}\right)} \\ &= -\frac{1}{\sigma}\frac{g'(Q_G(p))}{g(Q_G(p))} &&= \frac{1}{\sigma} J_G(p). \end{aligned}$$

□

Die fQ-Funktion innerhalb von $\mathcal{FQ}$

Im folgenden werden nun die speziellen Eigenschaften der fQ-Funktion innerhalb des fQ-Systems untersucht. Zunächst gilt für eine monotone lineare Transformation einer fQ-verteilten Zufallsvariable:

Lemma 3.4

Sei $X \sim \mathcal{F}(\mu, \sigma, \alpha, \beta) \in \mathcal{FQ}$ *und* $Y := aX + b$ *mit* $a > 0$, $b \in \mathbb{R}$. *Dann gilt*

$$Y \sim \mathcal{F}(\mu + b, \sigma/a, \alpha, \beta).$$

Beweis:

Mit Lemma 3.2 folgt

$$\begin{aligned} f_Y(Q_Y(p)) &= f_Y(aQ_X(p) + b) \\ &= \frac{1}{a} f_X(Q_X(p)) \\ &= \frac{\sigma}{a} p^\alpha (1-p)^\beta \end{aligned}$$

und

$$Q_Y(1/2) = Q_X(1/2) + b = \mu + b.$$

□

Lemma 3.5

Sei $F = \mathcal{F}(\mu, \sigma, \alpha, \beta) \in \mathcal{FQ}$, *dann gilt für* $p \in (0, 1)$

$$\begin{aligned} -J(p) = fQ'(p) &= fQ(p)\left(\frac{\alpha}{p} - \frac{\beta}{1-p}\right) &(3.6) \\ &= \sigma\left(\alpha p^{\alpha-1}(1-p)^\beta - \beta p^\alpha (1-p)^{\beta-1}\right), &(3.7) \\ \frac{J(p)}{fQ(p)} &= \frac{\beta}{1-p} - \frac{\alpha}{p}, &(3.8) \\ \int_0^1 fQ(p)\,dp &= \sigma B(\alpha+1, \beta+1), &(3.9) \end{aligned}$$

wobei $B(a, b)$ *die vollständige Betafunktion bezeichnet.*

Beweis:

Die Aussagen (3.6)- (3.8) folgen durch elementare Differentiation der fQ-Funktion und (3.9) aus der Definition der Betafunktion. □

Bemerkung 3.6
Die Gleichung (3.6) bedeutet insbesondere, daß die Dichte-Quantilfunktion einer fQ-Verteilung genau wie die Verteilungsfunktion einer einfachen Differentialgleichung genügt.
Mittels der Gleichung (3.7) ist es möglich, die Scorefunktion für fQ-Verteilungen explizit zu berechnen.

Die besondere Struktur der Scorefunktion (3.7) erlaubt folgende einfache Darstellung der Fisher-Information $I(F)$ der Dichtefunktion einer fQ-Verteilung F:

Satz 3.7
Sei $X \sim \mathcal{F}(\mu, \sigma, \alpha, \beta) \in \mathcal{FQ}$ mit $\alpha, \beta > 1/2$, dann gilt für die Fisher-Information der Dichte f von X

$$
\begin{aligned}
I(F) &= \int_{-\infty}^{\infty} \left(\frac{f'(x)}{f(x)} \right)^2 f(x)\, dx && (3.10) \\
&= 2\sigma^2 \alpha\beta \frac{\alpha + \beta - 1}{(2\beta - 1)(2\alpha - 1)} B(2\alpha, 2\beta). && (3.11)
\end{aligned}
$$

Beweis:
Der Beweis beruht zunächst auf der Umformung

$$I(F) = \int_0^1 (J(p))^2\, dp,$$

die aus der Substitution $x = Q(p)$ in (3.10) und der Definition (2.2) der Scorefunktion folgt. Durch Ausnutzung der Darstellung für die Scorefunktion J aus Lemma 3.5 ergibt sich dann

$$J^2(p) = \sigma^2 \left(\alpha p^{\alpha-1}(1-p)^{\beta} - \beta p^{\alpha}(1-p)^{\beta-1} \right)^2.$$

Durch Ausquadrieren dieses Ausdrucks und anschließende Integration erhält man:

$$
\begin{aligned}
\frac{I(F)}{\sigma^2} &= \alpha^2 \int_0^1 p^{2\alpha-2}(1-p)^{2\beta}\, dp - 2\alpha\beta \int_0^1 p^{2\alpha-1}(1-p)^{2\beta-1}\, dp \\
&+ \beta^2 \int_0^1 p^{2\alpha}(1-p)^{2\beta-2}\, dp.
\end{aligned}
$$

Da das Integral

$$\int_0^1 p^a \, dp$$

nur für $a > -1$ existiert (Bronstein [1984a, S. 308]), ergeben sich mit der Forderung $I(F) < \infty$ aus dem ersten und letzten Integral folgende notwendige Bedingungen:

$$\begin{aligned} 2\alpha - 2 > -1 \quad &\Leftrightarrow \quad \alpha > \frac{1}{2}, \\ 2\beta - 2 > -1 \quad &\Leftrightarrow \quad \beta > \frac{1}{2}. \end{aligned}$$

Sind diese Kriterien erfüllt, so lassen sich die Integrale gemäß (3.9) auflösen und es folgt

$$\begin{aligned} I(F) \quad &= \quad \sigma^2 \left(\alpha^2 B(2\alpha - 1, 2\beta + 1) - 2\alpha\beta B(2\alpha, 2\beta) \right. \\ &+ \quad \left. \beta^2 B(2\alpha + 1, 2\beta - 1) \right). \end{aligned} \tag{3.12}$$

Nach mehrfacher Anwendung der Beziehung

$$B(a+1, b) = \frac{a}{a+b} B(a, b)$$

folgt

$$\begin{aligned} B(2\alpha + 1, 2\beta - 1) \quad &= \quad \frac{2\alpha}{2\beta - 1} B(2\alpha, 2\beta), \\ B(2\alpha - 1, 2\beta + 1) \quad &= \quad \frac{2\beta}{2\alpha - 1} B(2\alpha, 2\beta). \end{aligned}$$

und damit die Behauptung. □

Als Teilergebnis dieses Satzes kann somit die folgende Charakterisierung der Menge $\mathcal{FQ}_I$, der fQ-Verteilungen mit endlicher Fisher-Information, festgehalten werden:

Lemma 3.8

Sei $F = \mathcal{F}(\mu, \sigma, \alpha, \beta) \in \mathcal{FQ}$, dann gilt

$$F \in \mathcal{FQ}_I \Leftrightarrow \alpha, \beta > 1/2. \tag{3.13}$$

3.2 Die fQ-Quantilfunktion

Da keine der drei definierenden Gleichungen des fQ-Systems eine direkte Berechnung der Quantilfunktion erlaubt, ist es notwendig, eine geeignete Umformung vorzunehmen, die eine numerische Behandlung ermöglicht.
Diese Umformung basiert dabei auf der Darstellung (C) aus Kapitel 2.1

$$Q'(x) = \frac{1}{\sigma} x^{-\alpha}(1-x)^{-\beta}.$$

Aufgrund der Stetigkeit von Q' im Intervall $(0,1)$ gilt mit Hilfe des ersten Hauptsatzes der Differential- und Integralrechnung:

$$Q(b) - Q(a) = \int_a^b Q'(x)\,dx, \qquad 0 < a < b < 1.$$

Unter Berücksichtigung des Lokationsparameters $\mu = Q(1/2)$ folgt dann für $b = 1/2, a = p$

$$Q(p) = \mu - \frac{1}{\sigma} \int_p^{1/2} x^{-\alpha}(1-x)^{-\beta}\,dx, \qquad p \in (0, 1/2\,],$$

und für $a = 1/2, b = p$

$$Q(p) = \mu + \frac{1}{\sigma} \int_{1/2}^{p} x^{-\alpha}(1-x)^{-\beta}\,dx, \qquad p \in (1/2, 1).$$

Mit der üblichen Konvention $\int_b^a \,dx = -\int_a^b \,dx$ ergibt sich also folgender Satz:

Satz 3.9
Die Quantilfunktion einer fQ-Verteilung $F = \mathcal{F}(\mu,\sigma,\alpha,\beta) \in \mathcal{FQ}$ ist gegeben durch:

$$Q(p) = \mu + \frac{1}{\sigma} \int_{1/2}^{p} x^{-\alpha}(1-x)^{-\beta}\,dx, \qquad p \in (0,1). \tag{3.14}$$

Bemerkung 3.10
Bei der Darstellung (3.14) der Quantilfunktion ist ein wesentlicher Punkt bei der numerischen Berechnung (s. Abschnitt 4.2) von Bedeutung:

Die Funktion Q wird im Gegensatz zur Verteilungsfunktion F nicht vom Rand des Definitionsbereichs $(Q(0), Q(1))$ aus berechnet, sondern von der Stelle $1/2$, also vom Median der Verteilung ausgehend. Dies ermöglicht eine gut konditionierte Berechnung von Q im zentralen Bereich der Verteilung.
Nachträglich wird durch diese Definition nun auch die spezielle Wahl des Lokationsparameters $\mu = Q(1/2)$ motiviert, da damit auf kanonische Weise eine im analytischen Sinne praktikable Größe verwendet wird.

Für den Fall $\alpha, \beta < 1$ kann das Integral in (3.14) als gestutzte Verteilungsfunktion einer Beta-Verteilung aufgefaßt werden. In diesem Fall erhält man also eine Darstellung der Quantilfunktion einer fQ-Verteilung durch die Verteilungsfunktion einer Beta-Verteilung:

Satz 3.11
Die Quantilfunktion einer fQ-Verteilung $F = \mathcal{F}(\mu, \sigma, \alpha, \beta) \in \mathcal{FQ}$ mit $\alpha, \beta < 1$ läßt sich schreiben als

$$Q(p) = \mu + \frac{B(a,b)}{\sigma} \left(F_{a,b}(p) - F_{a,b}(1/2)\right), \tag{3.15}$$

wobei $F_{a,b}$ die Verteilungsfunktion einer Beta-Verteilung mit Parametern $a = 1 - \alpha$ und $b = 1 - \beta$ ist.

Beweis:
Da allgemein die Verteilungsfunktion einer Beta-Verteilung $F_{a,b}$ durch das unvollständige Betaintegral

$$F_{a,b}(p) = \frac{1}{B(a,b)} \int_0^p x^{a-1}(1-x)^{b-1}\,dx$$

gegeben ist, folgt die Darstellung (3.15) aus:

$$\begin{aligned} Q(p) &= \mu + \frac{1}{\sigma} \int_{1/2}^p x^{-\alpha}(1-x)^{-\beta}\,dx \\ &= \mu + \frac{1}{\sigma} \left(\int_0^p x^{-\alpha}(1-x)^{-\beta}\,dx - \int_0^{1/2} x^{-\alpha}(1-x)^{-\beta}\,dx \right) \\ &= \mu + \frac{B(a,b)}{\sigma} \left(F_{a,b}(p) - F_{a,b}(1/2)\right). \end{aligned}$$

□

3.3 Randverhalten der Quantilfunktion

Der folgende Satz ist für alle weiteren Untersuchungen des Randverhaltens einer fQ-Verteilung von zentraler Bedeutung. Da in den praktischen Anwendungen lediglich die unimodalen Dichten betrachtet werden und diese durch die Bedingung $\alpha, \beta \geq 0$ bestimmt sind (Kapitel 3.4), erfolgt an dieser Stelle eine Beschränkung auf eben diese Klasse von Verteilungen.

Satz 3.12
Sei $F = \mathcal{F}(\mu, \sigma, \alpha, \beta) \in \mathcal{FQ}$ mit $\alpha, \beta \geq 0$, dann gelten für die Quantilfunktion folgende Abschätzungen:
1.Fall: $p > \frac{1}{2}$
a) $\beta \neq 1$

$$\frac{(1-p)^{1-\beta}}{\sigma(\beta-1)} - \frac{2^{\beta-1}}{\sigma(\beta-1)} \leq Q(p) - \mu \leq \frac{2^{\alpha}(1-p)^{1-\beta}}{\sigma(\beta-1)} - \frac{2^{\alpha+\beta-1}}{\sigma(\beta-1)} \tag{3.16}$$

b) $\beta = 1$

$$\frac{\log(1/2) - \log(1-p)}{\sigma} \leq Q(p) - \mu \leq \frac{2^{\alpha}}{\sigma}(\log(1/2) - \log(1-p)) \tag{3.17}$$

2.Fall: $p < \frac{1}{2}$
a) $\alpha \neq 1$

$$\frac{2^{\alpha-1}}{\sigma(\alpha-1)} - \frac{p^{1-\alpha}}{\sigma(\alpha-1)} \geq Q(p) - \mu \geq \frac{2^{\alpha+\beta-1}}{\sigma(\alpha-1)} - \frac{2^{\beta}p^{1-\alpha}}{\sigma(\alpha-1)} \tag{3.18}$$

b) $\alpha = 1$

$$\frac{\log(p) - \log(1/2)}{\sigma} \geq Q(p) - \mu \geq \frac{2^{\beta}}{\sigma}(\log(p) - \log(1/2)) \tag{3.19}$$

Beweis:
Alle Ungleichungen ergeben sich durch Abschätzungen für den Integranden in der Darstellung (3.14)

$$Q(p) = \mu + \frac{1}{\sigma}\int_{1/2}^{p} x^{-\alpha}(1-x)^{-\beta}\,dx$$

der Quantilfunktion.

1.Fall: $p > 1/2$.

Da die Funktion $p^{-\alpha}$ für $p \to 1$ monoton fallend ist, gilt für $x > 1/2$ folgende Abschätzung

$$1 \leq x^{-\alpha} \leq 2^{\alpha}.$$

Daraus folgt unmittelbar

$$\frac{1}{\sigma}\int_{1/2}^{p}(1-x)^{-\beta}\,dx \;\leq\; Q(p)-\mu \;\leq\; \frac{2^{\alpha}}{\sigma}\int_{1/2}^{p}(1-x)^{-\beta}\,dx.$$

Damit folgen die Aussagen für den ersten Fall durch Einsetzen des Integrals

$$\int_{1/2}^{p}(1-u)^{-\beta}\,du = \begin{cases} \frac{(1-p)^{1-\beta}-(1/2)^{1-\beta}}{\beta-1} & \beta \neq 1 \\ \log(1/2)-\log(1-p) & \beta = 1 \end{cases}$$

2.Fall: $p < 1/2$

Der Beweis verläuft vollkommen analog zum ersten Fall, wobei berücksichtigt werden muß, daß sich hier die Integrationsgrenzen und damit das Vorzeichen des Integrals in der Darstellung der Quantilfunktion ändert. Im übrigen sind nur die Rollen der Parameter α und β beziehungsweise der Argumente $1-p$ und p zu tauschen. □

Als erste Anwendung dieses Satzes betrachte man folgende vereinfachte Darstellung für den ersten Fall für $p > 1/2$ mit $\beta > 1$:

$$a_1 + b_1(1-p)^{1-\beta} \;\leq\; Q(p) \;\leq\; a_2 + b_2(1-p)^{1-\beta}, \tag{3.20}$$

mit

$$\begin{aligned} a_1 &= \mu - \frac{2^{\beta-1}}{\sigma(\beta-1)} \quad , \quad a_2 = \mu - \frac{2^{\alpha+\beta-1}}{\sigma(\beta-1)} \\ 0 < b_1 &= \frac{1}{\sigma(\beta-1)} \;<\; b_2 = \frac{2^{\alpha}}{\sigma(\beta-1)}. \end{aligned}$$

Damit ist eine Minorante und eine Majorante für die Quantilfunktion Q mit einem gemeinsamen funktionalen Kern $(1-p)^{1-\beta}$ gegeben, die später eine vereinfachte Untersuchung des Randverhaltens der Quantilfunktion und damit auch der Dichtefunktion erlauben.

3.4 Die fQ-Dichtefunktion

Da sich die Dichtefunktion f nicht in geschlossener Form darstellen läßt, werden in diesem Abschnitt nur Aussagen zur allgemeinen Gestalt von f gemacht. Dabei wird auf die entsprechenden Eigenschaften der Beta-Verteilung zurückgegriffen, die zusammenfassend in Johnson, Kotz und Balakrishnan [1995, Kapitel 25] dargestellt werden.

Zunächst wird untersucht, wie sich der Begriff der Unimodalität aus den Definitionen 2.4 und 2.5 durch Bedingungen an die fQ-Parameter ausdrücken läßt.

Satz 3.13

Sei $X \sim F = \mathcal{F}(\mu, \sigma, \alpha, \beta) \in \mathcal{FQ}$, dann gilt:

$$\alpha, \beta > 0 \quad \Leftrightarrow \quad F \in \mathcal{FQ}_U, \tag{3.21}$$

$$\alpha, \beta \geq 0 \quad \Leftrightarrow \quad F \in \mathcal{FQ}_{WU}. \tag{3.22}$$

Beweis:

Beide Behauptungen ergeben sich unmittelbar durch Anwendung der Aussagen für die Unimodalität einer entsprechenden Betadichte der Form

$$f(x) = \frac{x^\alpha (1-x)^\beta}{B(\alpha+1, \beta+1)},$$

siehe Johnson, Kotz und Balakrishnan [1995, S. 229]. □

Mit Hilfe von Satz 3.12 über das Randverhalten der Quantilfunktion können folgende Aussagen über den Träger einer fQ-Verteilung gemacht werden:

Satz 3.14

Sei $F = \mathcal{F}(\mu, \sigma, \alpha, \beta) \in \mathcal{FQ}$. Dann gelten folgende Aussagen:

$$\begin{aligned}
\alpha \geq 1 &\quad \Rightarrow \quad \alpha(F) = -\infty, \\
\alpha < 1 &\quad \Rightarrow \quad \alpha(F) > -\infty, \\
\beta \geq 1 &\quad \Rightarrow \quad \omega(F) = \infty, \\
\beta < 1 &\quad \Rightarrow \quad \omega(F) < \infty.
\end{aligned}$$

Für $\alpha, \beta < 1$ gilt speziell:

$$\begin{aligned} \alpha(F) &= \mu - \frac{B(1-\alpha, 1-\beta)}{\sigma} F_{1-\alpha,1-\beta}(1/2), \\ \omega(F) &= \mu + \frac{B(1-\alpha, 1-\beta)}{\sigma} \left(1 - F_{1-\alpha,1-\beta}(1/2)\right), \end{aligned} \tag{3.23}$$

wobei $F_{1-\alpha,1-\beta}$ die Verteilungsfunktion einer Beta-Verteilung mit Parametern $1-\alpha$ und $1-\beta$ bezeichnet.

Beweis:
Alle Aussagen folgen direkt aus Satz 3.12, wenn man in der vereinfachten Darstellung (3.20) die Grenzwerte für $p \to 1$ beziehungsweise $p \to 0$ betrachtet.
So gilt z.B. für $\beta > 1$

$$\lim_{p\to 1} Q(p) \geq \lim_{p\to 1}(a_1 + b_1(1-p)^{1-\beta}) = \infty, \quad \text{mit } b_1 > 0$$

und für $\beta < 1$

$$\lim_{p\to 1} Q(p) \leq \lim_{p\to 1}(a_2 + b_2(1-p)^{1-\beta}) = a_2 < \infty, \quad \text{mit } b_2 > 0.$$

Die Darstellung (3.23) schließlich folgt aus Satz 3.11 durch Einsetzen von $Q(0)$ beziehungsweise $Q(1)$. □

Als Alternative zu dieser *analytischen* Bedingung liefert der folgende Satz eine Charakterisierung, mit der anhand eines *optisch* überprüfbaren Kriteriums, der Existenz von Wendepunkten der fQ-Funktion, die Endlichkeit des Trägers von F überprüft werden kann:

Satz 3.15
Sei $F = \mathcal{F}(\mu, \sigma, \alpha, \beta) \in \mathcal{FQ}$, dann hat die fQ-Funktion von F genau dann einen Wendepunkt in $p > 1/2$ ($p < 1/2$), wenn gilt $\beta \geq 1$ ($\alpha \geq 1$).

Beweis:
Faßt man die fQ-Funktion

$$fQ(p) = \sigma p^{\alpha}(1-p)^{\beta}$$

als eine umskalierte Dichtefunktion einer Beta-Verteilung auf, so folgt aufgrund der entsprechenden Eigenschaften der Betadichte unter Berücksichtigung der unterschiedlichen Parametrisierung nach Johnson, Kotz und Balakrishnan [1994, S. 219], daß die fQ-Funktion Wendepunkte in den Punkten

$$p_{1;2} = \frac{\alpha}{\alpha+\beta} \pm \frac{\sqrt{\alpha\beta}}{\sqrt{\alpha+\beta-1}\,(\alpha+\beta)}$$

hat, sofern gilt

$$p_{1;2} \in [0\,,1].$$

Nun gilt

$$\begin{aligned} p_1 &\leq 1 \\ \Leftrightarrow \quad \alpha + \frac{\sqrt{\alpha\beta}}{\sqrt{\alpha+\beta-1}} &\leq \alpha+\beta \\ \Leftrightarrow \quad \frac{\alpha}{\alpha+\beta-1} &\leq \beta \\ \Leftrightarrow \quad 1 &\leq \beta \end{aligned}$$

und analog für p_2

$$\begin{aligned} p_2 &\geq 0 \\ \Leftrightarrow \quad \alpha &\geq \frac{\sqrt{\alpha\beta}}{\sqrt{\alpha+\beta-1}} \\ \Leftrightarrow \quad \alpha &\geq 1. \end{aligned}$$

□

Die Frage nach der Endlichkeit der Ränder $Q(0)$ und $Q(1)$ kann somit daran entschieden werden, ob die fQ-Funktion für $p < 1/2$ beziehungsweise $p > 1/2$ einen Wendepunkt hat (vgl. dazu auch die Abbildung 2.2).
In vielen statistischen Anwendungen ist die Existenz von Momenten eine notwendige Bedingung für weitere Untersuchungen. Mit Hilfe der nachfolgenden einfachen Kriterien können diese Informationen aus den Parametern einer fQ-Anpassung gewonnen werden.

Satz 3.16
Sei $X \sim \mathcal{F}(\mu, \sigma, \alpha, \beta) \in \mathcal{FQ}$ eine fQ-verteilte Zufallsvariable. Dann gilt

$$E(X^k) < \infty \Leftrightarrow \max(\alpha, \beta) < \frac{k+1}{k}, \quad k \in \mathbb{N}. \tag{3.24}$$

Beweis:

Da eine affine Transformation nichts an der Existenz der Momente ändert, sei o.B.d.A. $\mu = 0$ und $\sigma = 1$.

Zunächst läßt sich das k-te Moment von X durch die Quantilfunktion ausdrücken:

$$\begin{aligned} E(X^k) &= \int_{-\infty}^{\infty} x^k f(x)\, dx \\ &= \int_0^1 Q(p)^k\, dp \\ &= \int_0^{1/2} Q(p)^k\, dp + \int_{1/2}^1 Q(p)^k\, dp. \end{aligned}$$

Dann gilt:

$$\begin{aligned} E(X^k) < \infty \quad &\Leftrightarrow \quad \int_0^{1/2} |Q(p)^k|\, dp + \int_{1/2}^1 |Q(p)^k|\, dp < \infty \\ &\Leftrightarrow \quad \int_0^{1/2} |Q(p)^k|\, dp + \int_0^{1/2} |Q(1-p)^k|\, dp < \infty. \end{aligned}$$

Da weiter gilt

$$\begin{aligned} Q(1-p) &= \int_{1/2}^{1-p} x^{-\alpha}(1-x)^{-\beta}\, dx \\ &= -\int_{1/2}^{p} (1-x)^{-\alpha} x^{-\beta}\, dx := -\tilde{Q}(p), \end{aligned}$$

wobei $\tilde{Q}$ die Quantilfunktion einer fQ-Verteilung mit Parametern $(0, 1, \beta, \alpha)$ ist, reicht es aus, im folgenden das Integral

$$I := \int_0^{1/2} |Q(p)^k|\, dp$$

zu betrachten und anschließend die resultierenden Bedingungen an (α, β) auf (β, α) zu übertragen.

Der folgende Beweis orientiert sich an der Darstellung des Randverhaltens der Quantilfunktion in Satz 3.12 und unterscheidet dementsprechend mehrere Fälle.

1.Fall: $\alpha < 1$
Nach Satz 3.14 gilt hier

$$\alpha(F) > -\infty \Rightarrow \lim_{p \to 0+} Q(p) := C_1 > -\infty,$$

d.h., Q ist linksseitig beschränkt und I läßt sich durch

$$\int_0^{1/2} |Q(p)^k|\, dp < \left| \frac{|Q(1/2)^k| - |C_1^k|}{2} \right| < \infty$$

abschätzen.

2.Fall: $\alpha = 1$
Nach Satz 3.12 (Fall 2b) gilt für $p < 1/2$

$$|Q(p)| \leq 2^{\alpha}(\log(1/2) - \log(p))$$

und somit

$$\begin{aligned} \int_0^{1/2} |Q(p)|^k\, dp &\leq 2^{\alpha} \int_0^{1/2} (\log(1/2) - \log(p))^k\, dp \\ &= 2^{\alpha-1} \int_0^{\infty} u^k\, e^{-u}\, du \qquad \text{mit } u := \log(1/2) - \log(p) \\ &= 2^{\alpha-1}\Gamma(k-1) < \infty. \end{aligned}$$

3.Fall: $\alpha > 1$
Hier gilt nach Satz 3.12 (Fall 2a) für $p < 1/2$

$$|Q(p)| \leq C_2 + C_3\, p^{1-\alpha}, \quad C_2, C_3 > 0$$

und somit

$$\int_t^{1/2} |Q(p)|^k\, dp \leq \int_t^{1/2} (C_2 + C_3\, p^{1-\alpha})^k\, dp.$$

Entwickelt man nun dieses Polynom gemäß der binomischen Formel, so ist für die Konvergenzaussage nur der Term mit der höchsten Potenz von p relevant, und es folgt

$$\int_t^{1/2} |Q(p)|^k\, dp \propto C_4 \int_t^{1/2} p^{(1-\alpha)k}\, dp, \quad C_4 > 0.$$

Da das Integral

$$\lim_{t\to 0}\int_t^{1/2} p^a \, dp$$

nur für $a > -1$ konvergent ist (Bronstein, Semendjajew [1984a, S. 308]), folgt

$$\begin{aligned}\lim_{t\to 0}\int_t^{1/2} |Q(p)|^k \, dp < \infty \quad &\Leftrightarrow \quad (1-\alpha)k > -1\\ &\Leftrightarrow \quad 1 + \frac{1}{k} > \alpha.\end{aligned}$$

□

Bemerkung 3.17

Aufgrund der Darstellung

$$E(X^k) = \int_0^1 Q(p)^k \, dp \tag{3.25}$$

lassen sich die Momente einer fQ-Verteilung im Falle ihrer Existenz durch eine numerische Integration von (3.25) berechnen, wobei zu berücksichtigen ist, daß die Quantilfunktion Q ihrerseits auch nur in einer Integraldarstellung

$$Q(p) = \mu + \frac{1}{\sigma}\int_{1/2}^{p} x^{-\alpha}(1-x)^{-\beta} \, dx$$

vorliegt.

Die Komplexität dieser geschachtelten Integrale läßt keine explizite Angabe der Momente zu.

Kapitel 4

Numerische Verfahren

4.1 Motivation

Aufgrund der speziellen Gestalt der definierenden Gleichungen (A), (B) und (C) der fQ-Verteilung ist es nicht möglich, die Verteilungs-, die Quantil- oder die Dichtefunktion in einer geschlossenen Form anzugeben. Stattdessen muß für jede dieser drei Funktionen ein spezieller Algorithmus angewendet werden, wobei die unterschiedlichen Parametrisierungen der Verteilung ausgenutzt werden.

Dies impliziert, daß es unmöglich ist, Berechnungen *von Hand* durchzuführen, daß also die Anwendung des fQ-Systems nur mit Hilfe eines Computers realisiert werden kann. Für den praktischen Gebrauch der fQ-Verteilung scheint die Notwendigkeit von komplizierten Verfahren zur Berechnung selbst elementarer Dinge wie der Dichtefunktion zunächst ein großes Hindernis zu sein. Dabei sollte man sich aber vor Augen halten, daß dies grundsätzlich auch für andere Verteilungen gilt und etwa für die Berechnung der Quantile der Normalverteilung ebenfalls komplizierte Algorithmen notwendig sind.

Die besondere Bedeutung dieses Kapitels liegt somit darin, daß hier die zentralen numerischen Probleme bei der konkreten Darstellung der fQ-Verteilung behandelt werden, ohne deren Lösung ein sinnvoller Einsatz dieses Systems nicht denkbar wäre.

4.2 Berechnung der Quantilfunktion

Die konkrete Berechnung der Quantilfunktion beruht auf der Untersuchung des Ausdrucks für Q gemäß Satz 3.9:

$$Q(p) = \mu + \frac{1}{\sigma} \int_{\frac{1}{2}}^{p} x^{-\alpha}(1-x)^{-\beta}\, dx, \qquad p \in (0,1).$$

Wie aus der Analysis bekannt ist, gibt es keine allgemeine geschlossene Darstellung für das Integral

$$\int_a^b x^c(1-x)^d\, dx, \qquad 0 < a < b < 1, \quad c, d < 0,$$

und es muß statt dessen ein numerisches Verfahren benutzt werden.
Aufgrund der Tatsache, daß der Integrand

$$g(x) = x^{-\alpha}(1-x)^{-\beta}$$

eine stetige und lokal monotone Funktion ist, die zudem in den kritischen Randbereichen durch einen polynomialen Term dominiert wird (siehe auch Satz 3.12), bieten sich dabei Verfahren an, die für diesen funktionalen Typ geeignet sind.
Konkret wird dazu die Methode der *adaptiven Simpson-Regel* verwendet, die ein Spezialfall der *Newton-Cotes-Formeln* ist und folgendermaßen dargestellt werden kann (Björk, Dahlquist [1972, S. 196ff]):

$$\int_a^b g(x)\, dx = \frac{b-a}{6}\left(g(a) + 4\, g\left(\frac{a+b}{2}\right) + g(b)\right) + R, \tag{4.1}$$

mit einem Restterm der Form

$$R = -\frac{(b-a)^5}{90} g^{(4)}(\xi), \quad \xi \in [a,b]. \tag{4.2}$$

Üblicherweise wird zur Verbesserung der Genauigkeit das Intervall $[a,b]$ in Teile der Form

$$x_j = a + jh, \quad j = 0, \ldots, n, h := \frac{b-a}{n}, \tag{4.3}$$

geteilt, wobei h (bzw. n) gerade so gewählt wird, daß die Fehlerschranke R in (4.2) durch Abschätzungen für die Ableitung $g^{(4)}$ unter einer vorgegebenen Schranke bleibt.

Da sich im Falle der Quantilfunktion bei der Ableitung des Integranden $g(x)$ durch die wiederholte partielle Differentiation ein komplizierter Term für $g^{(4)}$ ergibt, dessen Abschätzung für einen lokalen Bereich zu aufwendigen Analysen führen würde, wird statt dessen ein Verfahren benutzt, welches das allgemeine Prinzip der Intervallhalbierung nutzt (vgl. auch Sedgewick [1992, S. 638ff]).

Die Idee dieses adaptiven Verfahrens besteht darin, durch eine rekursive Teilung des Integrationsbereichs die Schrittweite h zu verringern, wobei die Entscheidung, ob ein Teilintervall noch weiter unterteilt wird, von lokalen Informationen innerhalb dieses Intervalls abhängig gemacht wird. Sedgewick schlägt als Kriterium für eine weitere Teilung die Differenz zweier Simpson-Näherungen der Form (4.3) mit verschiedenen Schrittzahlen $n_1 < n_2$ vor, wobei die Teilung dann abgebrochen wird, wenn diese Differenz hinreichend klein ist. Der Nachteil dieser Methdode besteht darin, daß in jedem Fall zumindest eine überflüssige Simpson-Integration durchgeführt wird, nämlich die mit der kleineren Schrittweite n_1.

Demgegenüber orientiert sich die hier vorgeschlagene Vorgehensweise an dem Randverhalten der Funktion g. Dies bedeutet konkret, daß ein Intervall $[x_j, x_{j+1}]$ genau dann unterteilt wird, wenn gilt

$$|g(x_{j+1}) - g(x_j)| > \epsilon$$

für ein vorgegebenes $\epsilon > 0$. Somit werden gerade im Randbereich der Funktion g aufgrund des steileren Anstiegs vermehrt Intervallhalbierungen durchgeführt. Die besondere Konstruktion dieser Methode erlaubt es zudem, alle Zwischenergebnisse zu verwenden, da die Funktionswerte an den Rändern des Intervalls $g(x_j)$ und $g(x_{j+1})$ ebenfalls bei den Teilintervallen $g(x_j), g((x_{j+1} + x_j))/2$ und $g((x_{j+1} + x_j))/2, g(x_{j+1})$ benötigt werden.
In einem Pseudo-Code läßt sich das rekursive Verfahren folgendermaßen darstellen, wobei hier eine Syntax benutzt wird, die sich stark an die statistische Programmiersprache S-PLUS anlehnt:

```
AdapSimpson(a,b,fa,fb)
{ fmit <- f((a+b)/2)
  if (abs(fa - fb) > epsilon)
    wert <- AdapSimpson(a,(a+b)/2,fa,fmit) +
            AdapSimpson((a+b)/2,b,fmit,fb)
  else wert <- (fa + 4*fmit + fb)*(b-a)/6
  return wert
}
```

Dabei wird in obiger Darstellung davon ausgegangen, daß eine externe Routine f vorliegt, mit der die Funktionswerte bestimmt werden können. Der Aufruf der Funktion erfolgt dann in der Form

```
AdapSimpson(a,b,f(a),f(b)).
```

Für den globalen Parameter ϵ, mit dem der Grad der Intervallschachtelung gesteuert wird, erwies sich ein Wert von $\epsilon = 1/2$ als zufriedenstellend.
Da die Quantilfunktion eine wesentliche Rolle bei der späteren Schätzung empirischer fQ-Parameter in Kapitel 7 spielt und sie in diesem Zusammenhang stets für einen Vektor $\mathbf{p} = (p_1, \ldots, p_n)$ mit $0 < p_1 < \ldots < p_n < 1$ berechnet wird, ist es sinnvoll, diese spezielle Aufgabenstellung bei der algorithmischen Darstellung zu berücksichtigen und eine Aufsplittung des Integrals durchzuführen:
Es gilt für $p_1 < p_2 < \frac{1}{2}$

$$
\begin{aligned}
Q(p_1) &= \mu - \frac{1}{\sigma} \int_{p_1}^{\frac{1}{2}} x^{-\alpha}(1-x)^{-\beta}\, dx \\
&= \mu - \frac{1}{\sigma} \int_{p_1}^{p_2} x^{-\alpha}(1-x)^{-\beta}\, dx - \frac{1}{\sigma} \int_{p_2}^{\frac{1}{2}} x^{-\alpha}(1-x)^{-\beta}\, dx \\
&= Q(p_2) - \frac{1}{\sigma} \int_{p_1}^{p_2} x^{-\alpha}(1-x)^{-\beta}\, dx.
\end{aligned}
$$

Durch diese Aufteilung wird die numerische Genauigkeit des Verfahrens verbessert, da bei der separaten Berechnung von $Q(p_1)$ und $Q(p_2)$ die Funktion über Intervalle mit der Gesamtlänge

$$I_1 = (\frac{1}{2} - p_1) + (\frac{1}{2} - p_2) = 1 - (p_1 + p_2)$$

integriert wird, während nach der Aufteilung ein Intervall der Länge

$$I_2 = (\frac{1}{2} - p_2) + (p_2 - p_1) = \frac{1}{2} - p_1$$

betrachtet wird. Weiterhin werden dadurch auch überflüssige Berechnungen vermieden.
Um diese Methode konkret nutzen zu können, ist es zusätzlich notwendig, den Vektor $\mathbf{p}$ an der Stelle $p = 1/2$ aufzuteilen, da die numerische Integration von diesem Punkt ausgehend in Richtung der Ränder $p = 0$ und $p = 1$ verläuft.
Zusammengefaßt läßt sich das Verfahren folgendermaßen darstellen:

Algorithmus 4.1
Die Quantilfunktion einer fQ-Verteilung $F = \mathcal{F}(\mu, \sigma, \alpha, \beta) \in \mathcal{FQ}$ für einen Vektor $\mathbf{p} = (p_1, \ldots, p_n)$ mit $0 < p_1 < \ldots p_n < 1$ wird in folgenden Schritten berechnet:

1) Aufteilung des Vektors $\mathbf{p}$ an der Stelle $p = 1/2$:
Es seien

$$p_{n_1} := \max_{1 \le i \le n} \{p_i < 1/2\}$$

und

$$\begin{aligned} p^u &:= (p_1, \ldots, p_{n_1}, 1/2) =: (p_1^u, \ldots, p_N^u), \\ p^o &:= (1/2, p_{n_1+1}, \ldots, p_n) =: (p_1^o, \ldots, p_M^o). \end{aligned}$$

2) Berechnung der Teilintegrale:

$$\begin{aligned} I_i^u &:= \int_{p_i^u}^{p_{i+1}^u} g(x)\, dx, \quad i = 1, \ldots, N-1, \\ I_i^o &:= \int_{p_i^o}^{p_{i+1}^o} g(x)\, dx, \quad i = 1, \ldots, M-1, \end{aligned}$$

mit

$$g(x) := x^\alpha (1-x)^\beta$$

mit Hilfe der adaptiven Simpson-Methode.

3) Sukzessive Berechnung der standardisierten Quantile:

$$\begin{aligned} q^u_{N-i} &:= \sum_{j=N-i}^{N-1} I^u_j \quad , i = 1, \ldots, N-1, \\ q^o_i &:= \sum_{j=i}^{M-1} I^o_j \quad , i = 1, \ldots, M-1. \end{aligned}$$

4) Umskalierung und Zusammenführung der Teilergebnisse:

$$Q(\mathbf{p}) = \mu + \frac{1}{\sigma}(-q^u_1, \ldots, -q^u_{N-1}, q^o_1, \ldots, q^o_{M-1}).$$

4.3 Berechnung der Verteilungsfunktion

Bei der Berechnung der Verteilungsfunktion wird von der Parametrisierung (A) der fQ-Verteilung Gebrauch gemacht:

$$F'(x) = \sigma F(x)^\alpha (1 - F(x))^\beta.$$

In Verbindung mit dem Lokationsparameter $\mu = Q(1/2)$ kann die Funktion F als Lösung $y(t)$ eines allgemeinen Anfangswertproblems (AWP)

$$y' = f(t, y), \quad t \in [a, b], \tag{4.4}$$

$$y(a) = y_0 \tag{4.5}$$

mit der speziellen Wahl

$$\begin{aligned} f(t, y) &:= \sigma y(t)^\alpha (1 - y(t))^\beta, \\ y(\mu) &:= 1/2 \end{aligned}$$

aufgefaßt werden.
Obwohl es sich bei der Gleichung (4.4) um eine elementare Differentialgleichung erster Ordnung mit getrennten Variablen handelt, existiert im allgemeinen Fall keine explizite Lösung der Gleichung. Ähnlich wie bei der Berechnung der Quantilfunktion müssen auch hier numerische Verfahren eingesetzt werden.
Aufgrund der stetigen Differenzierbarkeit der Funktion $f(t, y)$ bietet sich dazu ein spezielles Verfahren an, welches zunächst für das allgemeine AWP beschrieben wird (Burden und Faires [1989, S. 235ff]).

Das Taylorverfahren

Zur Berechnung von $y(b)$ mit $a < b$ seien zunächst N äquidistante Stützstellen definiert

$$t_i = a + ih, \quad h := \frac{b-a}{N},$$

mit deren Hilfe das Intervall $[a, b]$ in regelmäßige Teilintervalle zerlegt wird. Ein größeres N bewirkt eine Verbesserung der Lösung, führt aber auch zu einem vermehrten Rechenaufwand.

Durch eine Taylorentwicklung der unbekannten Lösung $y(t)$ an einer Stützstelle t_i erhält man:

$$\begin{aligned} y(t_{i+1}) &= y(t_i) + hy'(t_i) + \frac{h^2}{2} y''(t_i) + \ldots \\ &+ \frac{h^n}{n!} y^{(n)}(t_i) + \frac{h^{n+1}}{(n+1)!} y^{(n+1)}(\xi_i) \end{aligned} \tag{4.6}$$

mit einem $\xi \in (t_i, t_{i+1})$.

Aus der Differentialgleichung selbst folgt durch sukzessives Ableiten:

$$\begin{aligned} y'(t) &= f(t, y(t)), \\ y''(t) &= f'(t, y(t)) \end{aligned}$$

und somit

$$y^{(k)}(t) = f^{(k-1)}(t, y(t)).$$

Eingesetzt in (4.6) ergibt sich folgende Beziehung

$$\begin{aligned} y(t_{i+1}) &= y(t_i) + hf(t_i, y(t_i)) + \frac{h^2}{2} f'(t_i, y(t_i)) + \ldots \\ &+ \frac{h^n}{n!} f^{(n-1)}(t_i, y(t_i)) + \frac{h^{n+1}}{(n+1)!} f^{(n)}(\xi_i, y(\xi_i)), \end{aligned}$$

die in Verbindung mit der Randbedingung

$$y(t_0) = y(a) = y_0$$

eine rekursive Berechnung von $y(b), b > a$ ermöglicht.

Umsetzung im fQ-System

Da für den speziellen Fall der Verteilungsfunktion F der fQ-Verteilung

$$\begin{aligned} f'(t, y(t)) &= \frac{d}{dt}(\sigma F(t)^\alpha (1 - F(t))^\beta) \\ &= \sigma f(t) \left(\alpha F(t)^{\alpha-1}(1 - F(t))^\beta - \beta F(t)^\alpha (1 - F(t))^{\beta-1}\right), \end{aligned}$$

gilt, läßt sich die Rekursionsvorschrift (4.7) zur Berechnung des Funktionswertes an der Stelle $F(t_{i+1})$ durch eine Taylorapproximation bis zum zweiten Glied folgendermaßen darstellen:

$$\begin{aligned} F(t_{i+1}) &\approx F(t_i) + h\sigma F(t_i)^\alpha (1 - F(t_i))^\beta \\ &\quad \left(1 + \frac{h}{2}\sigma F(t_i)^{\alpha-1} (1 - F(t_i))^{\beta-1} (\alpha(1 - F(t_i)) - \beta F(t_i))\right). \end{aligned} \tag{4.7}$$

Erneut muß bei der algorithmischen Beschreibung des Verfahrens berücksichtigt werden, daß aufgrund der Anfangsbedingung $F(\mu) = 1/2$ die Fälle $x > \mu$ und $x < \mu$ unterschieden werden müssen.

In Analogie zur adaptiven Verfahrensweise bei der Berechnung der Quantilfunktion wird auch hier eine datenorientierte Bestimmung der Stützstellenanzahl für ein Intervall (x_i, x_{i+1}) durch die Wahl

$$N(x_i, x_{i+1}) := \max(N_{max}, \lfloor \frac{x_{i+1} - x_i}{\epsilon} \rfloor), \tag{4.8}$$

vorgeschlagen mit dem gleichen globalen Parameter ϵ wie beim adaptiven Simpsonverfahren und einer Konstanten N_{max}, mit der die maximale Anzahl von Teilschritten begrenzt werden kann. In der Praxis wurde hier der Wert $N_{max} = 5$ gewählt.

Die folgende Darstellung orientiert sich erneut an einer vektororientierten Problemstellung.

Algorithmus 4.2

Zur Berechnung der Verteilungsfunktion einer Verteilung $F = \mathcal{F}(\mu, \sigma, \alpha, \beta) \in \mathcal{FQ}$ *für einen Vektor* $\mathbf{x} = (x_1, ..., x_n)$ *mit* $x_1 < \ldots x_n$ *seien*

$$x_{n_1} := \max_{1 \le i \le n} \{x_i < \mu\}$$

und

$$\begin{aligned} x^u &:= (x_1, \ldots, x_{n_1}, \mu) =: (x_1^u, \ldots, x_N^u), \\ x^o &:= (\mu, x_{n_1+1}, \ldots, x_n) =: (x_1^o, \ldots, x_M^o). \end{aligned}$$

Ausgehend von x_N^u *mit* $F(x_N^u) = 1/2$ *werden für die Werte* x_{N-i}^u, $i = 1, \ldots, N-1$ *mit den Hilfsgrößen*

$$\begin{aligned} N_i &:= N(x_{N-i}^u, x_{N-i+1}^u) \quad \text{gemäß (4.8)} \\ h_i &:= (x_{N-i+1}^u - x_{N-i}^u)/N_i, \\ t_j &:= x_{N-i+1}^u - jh_i, \quad j = 0, \ldots, N_i \end{aligned}$$

durch sukzessive Anwendung der Rechenvorschrift (4.7) die Werte $F(x_{N-i}^u)$, $i = 1, \ldots, N-1$, *berechnet.*

Für x_i^o ,$i = 1, \ldots, M-1$, *führt eine analoge Vorgehensweise mit den Hilfsgrößen*

$$\begin{aligned} N_i &:= N(x_i^o, x_{i+1}^o) \\ h_i &:= (x_{i+1}^o - x_i^o)/N_i, \\ t_j &:= x_{i+1}^o + jh_i, \quad j = 1, \ldots, N_i \end{aligned}$$

zur Berechnung von $F(x_i^o), i = 1, \ldots, M-1$.

Der Lösungsvektor ist dann gegeben durch

$$F(\mathbf{x}) = (x_1^u, \ldots, x_{N-1}^u, x_1^o, \ldots, x_M^o).$$

4.4 Berechnung der Dichtefunktion

Die Bestimmung der Dichtefunktion an einer Stelle x erfolgt gemäß der Parametrisierung (A)

$$f(x) = \sigma F(x)^{\alpha}(1 - F(x))^{\beta}$$

durch Berechnung der Verteilungsfunktion $F(x)$ und Einsetzen dieses Wertes in obige Gleichung. Dabei kann das vektorielle Verfahren nach Algorithmus 4.2 benutzt werden.

Kapitel 5

Ordnungen, Metriken und Kennmaße

5.1 Motivation und Übersicht

Wann immer im Zuge einer Analyse mehrere Zufallsvariablen X_i bzw. ihre zugrundeliegenden Verteilungen F_i miteinander verglichen werden sollen, müssen sinnvolle Vergleichsoperatoren auf der Menge der möglichen Verteilungen definiert werden.
Man kann derartige Operatoren zunächst in zwei Kategorien einteilen:

- *Quantitative Vergleiche* durch Ordnungsbegriffe,
- *Metrische Vergleiche* durch Abstandsmaße.

Die quantitativen Vergleiche dienen dazu, eine Relation zwischen Verteilungen bezüglich eines globalen Merkmals zu schaffen. Im Rahmen dieser Arbeit wird dabei für folgende Merkmale jeweils ein spezieller Ordnungsbegriff auf $\mathcal{FQ}$ definiert:

$$\begin{aligned} \text{Streuung} &\longrightarrow \text{Dispersionsordnung,} \\ \text{Schiefe} &\longrightarrow \text{Konvexordnung,} \\ \text{Randverhalten} &\longrightarrow \text{Tailordnung.} \end{aligned}$$

Ein Problem bei den quantitativen Aussagen besteht darin, daß die resultierenden Bedingungen an die Verteilungen (genauer: an die charakterisierenden Parameter) in der Regel sehr restriktiv sind und für gewisse Parameterkonstellationen weder

in der einen noch in der anderen Vergleichsrichtung zu einem Ergebnis führen. Die entsprechenden Relationen definieren somit nur eine Halbordnung auf $\mathcal{FQ}$. Demgegenüber ist bei Verwendung einer *Metrik* per definitionem gewährleistet, daß zwei beliebige Verteilungen miteinander vergleichen werden können und zudem zu einer qualitativ feineren Aussage, einem Abstandsbegriff, führen. Im Gegensatz zu den quantitativen Aussagen ist dieser Abstand jedoch nur als Maßzahl zu verstehen und kann daher nicht direkt im Sinne eines globalen Merkmals interpretiert werden.
Um dies zu realisieren, werden im dritten Teil dieses Kapitels *Kennmaße* eingeführt, mit denen die Charakteristika einzelner Verteilungen beschrieben werden können. Eine Untersuchung der Kennmaße verschiedener Verteilungen liefert dann eine dritte Möglichkeit, um Verteilungen miteinander zu vergleichen.

5.2 Ordnungsrelationen

Allgemein gibt es für Klassen von Verteilungen eine Vielzahl von möglichen Definitionen für stochastische Ordnungsrelationen (siehe z.B. Shaked [1994]). Da die fQ-Verteilungen durch ihre Quantilfunktion beziehungsweise durch deren Ableitung charakterisiert werden, liegt es nahe, gerade solche Ordnungsrelationen auf $\mathcal{FQ}$ zu definieren, die durch Eigenschaften der Quantilfunktion beschrieben werden.
Als ein wichtiges analytisches Instrumentarium wird sich die folgende Hilfsfunktion erweisen, mit der man in der Lage ist, die Beziehung zweier Verteilungsfunktionen in analytischer Form zu beschreiben.

Lemma 5.1
Seien $F, G \in \mathcal{F}_S$ stetig differenzierbare Verteilungsfunktionen und X eine Zufallsvariable mit $X \sim F$.
Dann existiert eine Funktion $\Psi_{F,G} : D_f \to \mathbb{R}$, für die gilt

$$\Psi_{F,G}(X) \sim G.$$

$\Psi_{F,G}$ ist eindeutig bestimmt und hat die Form

$$\Psi_{F,G}(x) := G^{-1}(F(x)). \tag{5.1}$$

Diese Transferfunktion $\Psi_{F,G}$ *ist streng monoton steigend und stetig differenzierbar und es gilt*

$$\Psi'_{F,G}(x) = \frac{fQ_F(x)}{fQ_G(x)}. \tag{5.2}$$

Beweis:
Der Beweis von (5.1) beruht auf elementaren Umformungen unter Ausnutzung der bekannten Aussage für Verteilungsfunktionen

$$F_X(x) \sim U_{[0,1]}$$

(siehe z.B. David [1980, S. 72]).
Die Aussage (5.2) folgt direkt durch Differentiation der Funktion $\Psi_{F,G}$. □

Wenn aus dem Zusammenhang erkennbar ist, um welche Transferfunktion es sich handelt, wird im folgenden der Index F, G weggelassen.

5.2.1 Die Dispersionsordnung

Der Begriff der *Dispersion* geht zurück auf Brown und Tukey (1948) und beschreibt das unterschiedliche Streuungsverhalten zweier Verteilungen anhand der Abstände der Quantilfunktionen:

Definition 5.2
Seien $F, G \in \mathcal{F}_S$, dann hat G eine größere Dispersion *als F,*

$$F \overset{d}{<} G,$$

wenn gilt

$$G^{-1}(b) - G^{-1}(a) \geq F^{-1}(b) - F^{-1}(a), \quad \forall\, 0 < a < b < 1. \tag{5.3}$$

F und G heißen dispersionsgleich,

$$F \overset{d}{\sim} G,$$

wenn gilt

$$F \overset{d}{<} G \quad \text{und} \quad G \overset{d}{<} F.$$

Für $F \overset{d}{<} G$ kann äquivalent $G \overset{d}{>} F$ geschrieben werden.

Mit der folgenden Definition kann eine vergleichbare Aussage auch für Zufallsvariablen formuliert werden:

Definition 5.3
Sind X und Y zwei Zufallsvariablen mit $X \sim F \in \mathcal{F}_S$ und $Y \sim G \in \mathcal{F}_S$, dann hat Y eine größere Dispersion als X (ist dispersionsgleich zu X),

$$X \overset{d}{<} Y \quad (X \overset{d}{\sim} Y),$$

wenn diese Aussage für die Verteilungsfunktionen F und G gilt.

Offensichtlich ist durch diese Definition in der Tat ein Vergleich der Streuungen von F und G gegeben, da die Beziehung (5.3) bedeutet, daß die Differenz zweier beliebiger Quantile von G mindestens so groß ist wie die entsprechende Differenz der Quantile von F.

Bemerkung 5.4
Aus der Bedingung (5.3) folgt unmittelbar:

$$\begin{aligned} F &\overset{d}{<} F \qquad \text{und} \\ F \overset{d}{<} G, G \overset{d}{<} H \quad &\Rightarrow \quad F \overset{d}{<} H. \end{aligned}$$

Somit definiert die Relation $\overset{d}{<}$ zumindest eine Halbordnung auf $\mathcal{F}_S$. Ein Vergleich der Quantilfunktionen der Normalverteilung und der χ^2-Verteilung zeigt, daß $\overset{d}{<}$ keine Ordnung definiert, da gilt:

$$\begin{aligned} Q_{NV}(0.9) - Q_{NV}(0.5) \approx 1.28 \quad &< \quad Q_{\chi^2}(0.9) - Q_{\chi^2}(0.5) \approx 2.25 \\ Q_{NV}(0.5) - Q_{NV}(0.1) \approx 1.28 \quad &> \quad Q_{\chi^2}(0.5) - Q_{\chi^2}(0.1) \approx 0.44. \end{aligned}$$

Offensichtlich ist es mit dieser Relation also nicht möglich, Verteilungen miteinander zu vergleichen, die eine unterschiedliche Schiefe haben.

Für die definierende Bedingung (5.3) gibt es mehrere äquivalente Formulierungen, von denen für das fQ-System die folgende relevant ist:

Satz 5.5
Seien $F, G \in \mathcal{F}_S$ mit zugehöriger Transferfunktion $\Psi_{F,G}$, dann gilt:

$$F \overset{d}{<} G \Leftrightarrow \Psi'_{F,G}(x) \geq 1, \quad \forall x \in D_f.$$

Beweis:
a) Sei $F \overset{d}{<} G$. Dann gilt für $x_1 < x_2$

$$\begin{aligned} x_2 - x_1 &\leq G^{-1}(F(x_2)) - G^{-1}(F(x_1)) \\ \Leftrightarrow 1 &\leq \frac{\Psi(x_2) - \Psi(x_1)}{x_2 - x_1}, \end{aligned}$$

woraus durch die Grenzwertbildung $x_2 \to x_1$ die Behauptung folgt.

b) Sei $a < b$ und $x_1 = F^{-1}(a)$, $x_2 = F^{-1}(b)$. Dann gilt nach Definition von Ψ

$$G^{-1}(b) - G^{-1}(a) = \Psi(x_2) - \Psi(x_1).$$

Da nach Voraussetzung $\Psi'(x) \geq 1$ gilt, ergibt sich

$$\Psi(x_2) - \Psi(x_1) \geq x_2 - x_1,$$

und somit

$$G^{-1}(b) - G^{-1}(a) \geq x_2 - x_1 \geq F^{-1}(b) - F^{-1}(a),$$

also $F \overset{d}{<} G$. □

Da nach (5.1) die Ableitung der Transferfunktion Ψ mit Hilfe der entsprechenden fQ-Funktionen dargestellt werden kann, ergibt sich folgende Charakterisierung der Dispersionsordnung im fQ-System:

Satz 5.6
Seien $F, G \in \mathcal{FQ}$ mit $F = \mathcal{F}(\mu_1, \sigma_1, \alpha_1, \beta_1)$, $G = \mathcal{F}(\mu_2, \sigma_2, \alpha_2, \beta_2)$. Mit den Bezeichnungen

$$a^* := \alpha_1 - \alpha_2,\ b^* := \beta_1 - \beta_2,\ \sigma^* := \frac{\sigma_1}{\sigma_2}$$

gilt $F \overset{d}{<} G$ genau dann, wenn folgende Bedingungen erfüllt sind

$$\begin{aligned} \alpha_1 &\leq \alpha_2, \\ \beta_1 &\leq \beta_2, \\ \sigma^* &\geq \frac{(a^* + b^*)^{a^* + b^*}}{{a^*}^{a^*} {b^*}^{b^*}}. \end{aligned}$$

Beweis:

Aufgrund der Darstellung (5.1) der Funktion Ψ gilt

$$F \overset{d}{<} G \Leftrightarrow \frac{fQ_F(x)}{fQ_G(x)} \geq 1.$$

Mit den Bezeichnungen des Satzes ist dies äquivalent ist zu

$$h(x) := \sigma^* x^{a^*}(1-x)^{b^*} \geq 1. \tag{5.4}$$

Für diese Funktion h gelten folgende Aussagen:

$$\lim_{x\to 0} h(x) = \begin{cases} \infty & : \quad a^* < 0 \\ \sigma^* & : \quad a^* = 0 \\ 0 & : \quad a^* > 0 \end{cases}$$

und analog

$$\lim_{x\to 1} h(x) = \begin{cases} \infty & : \quad b^* < 0 \\ \sigma^* & : \quad b^* = 0 \\ 0 & : \quad b^* > 0 \end{cases}.$$

Aus diesen Beziehungen ergeben sich direkt die folgenden beiden notwendigen Bedingungen für (5.4)

$$a^* \leq 0, \quad b^* \leq 0,$$

da andernfalls

$$\lim_{x\to 0} h(x) = 0 \quad \text{oder} \quad \lim_{x\to 1} h(x) = 0$$

gelten würde.

Da weiterhin im Falle einer Parameterkonstellation der Form $a^*b^* = 0$ die Funktion h streng monoton fallend ($a^* = 0, b^* \neq 0$), streng monoton steigend ($b^* = 0, a^* \neq 0$) oder konstant ($a^* = 0, b^* = 0$) und somit beschränkt ist, ergibt sich hier die Bedingung $\sigma^* \geq 1$, da in jedem der drei obigen Fälle gilt

$$\max_{x\in[0,1]} h(x) = \sigma^*.$$

Somit bleibt nur noch der Fall $a^*, b^* < 0$ zu untersuchen. In diesem Fall kann zur Abschätzung von σ^* eine Aussage über die Gestalt einer entsprechenden Betadichte herangezogen werden (Johnson, Kotz und Balakrishnan [1995, S. 219]), wonach die Funktion h einen U-förmigen Verlauf hat und ein eindeutiges Minimum im Punkt $x_{min} = a^*/(a^* + b^*)$ besitzt. Damit gilt

$$h(x_{min}) \geq 1 \Leftrightarrow \sigma^* \frac{a^{*\,a^*} b^{*\,b^*}}{(a^* + b^*)^{a^* + b^*}} \geq 1. \tag{5.5}$$

Da im Fall $a^* b^* = 0$

$$\frac{a^{*\,a^*} b^{*\,b^*}}{(a^* + b^*)^{a^* + b^*}} = 1$$

gilt, wird die Bedingung an σ^* für diesen Fall ebenfalls durch die Aussage (5.5) abgedeckt, woraus die Behauptung folgt. □

Bemerkung 5.7

Der Vorteil der Dispersionsordnung liegt darin, daß sich die Beziehung $F \overset{d}{<} G$ *direkt interpretieren läßt, da die äquivalente Bedingung*

$$\Psi'(x) = \frac{fQ_F(x)}{fQ_G(x)} \geq 1$$

bedeutet, daß der Graph von fQ_F *nicht unterhalb des Graphen von* fQ_G *verläuft. Man beachte, daß diese Aussage allgemein für* $F, G \in \mathcal{F}_S$ *gilt - dieser Umstand wird später in Kapitel 5.5 ausgenutzt.*

Da weiterhin nach Lemma 3.4 für eine fQ-verteilte Zufallsvariable $X \sim F$ und $Y := aX + b$ $(a > 0)$

$$fQ_Y(x) = \frac{fQ_X(x)}{a}$$

gilt, folgt für die Transferfunktion

$$\Psi'(x) = a,$$

und somit

$$X \begin{cases} \overset{d}{<} & aX + b \quad : a > 1 \\ \overset{d}{\sim} & aX + b \quad : a = 1 \\ \overset{d}{>} & aX + b \quad : a < 1 \end{cases}.$$

Dies steht im Einklang mit den prinizipiellen Forderungen an eine Dispersionsordnung, nach denen die Relation einerseits lokationsinvariant sein sollte und andererseits die Multiplikation mit einem Skalenfaktor > 1 (< 1) zu einer Vergrößerung (Verkleinerung) der Streuung führen sollte.

5.2.2 Die Konvexordnung

Die Konvexordnung benutzt ebenso wie die Dispersionsordnung die Transferfunktion Ψ, wobei jedoch hier eine schwächere Bedingung formuliert wird. Der Begriff der Konvexordnung geht zurück auf van Zwet (1964).

Definition 5.8
Seien $F, G \in \mathcal{F}_S$, dann heißt F konvexkleiner *als G*

$$F \overset{k}{<} G,$$

falls die Transferfunktion $\Psi_{F,G}$ schwach konvex ist. F ist konvexäquivalent *zu G*

$$F \overset{k}{\sim} G,$$

falls gilt

$$F \overset{k}{<} G \quad \text{und} \quad G \overset{k}{<} F.$$

Sind X und Y zwei Zufallsvariablen mit $X \sim F \in \mathcal{F}_S$ und $Y \sim G \in \mathcal{F}_S$, dann heißt X konvexkleiner als Y (konvexäquivalent zu Y),

$$X \overset{k}{<} Y \quad (X \overset{k}{\sim} Y),$$

wenn diese Aussage für die Verteilungsfunktionen F und G gilt.

Im folgenden wird im Zusammenhang mit der Konvexordnung auf den Zusatz *schwach* bei der Konvexität verzichtet. In Anschluß an Lemma 5.13 wird ausgeführt, daß der unhandliche Ausdruck *X ist konvexkleiner als Y* durch den anschaulicheren Begriff *Y ist rechtsschiefer als X* ersetzt werden kann.

Bemerkung 5.9
Da die Identität konvex ist und eine monotone konvexe Transformation einer konvexen Transformation wieder konvex ist, folgt für $F, G, H \in \mathcal{F}_S$

$$\begin{aligned} F \quad &\overset{k}{<} \quad F \qquad \text{und} \\ F \overset{k}{<} G, G \overset{k}{<} H \quad &\Rightarrow \quad F \overset{k}{<} H. \end{aligned}$$

Somit definiert die Relation $\overset{k}{<}$ eine Halbordnung auf $\mathcal{F}_S$.

Das folgende Lemma liefert eine inhaltlich klarere Interpretation der Konvexäquivalenz.

Lemma 5.10
Seien $F, G \in \mathcal{F}_S$, dann gilt

$$F \overset{k}{\sim} G \Leftrightarrow \exists\, a > 0, b \in \mathbb{R} \quad mit \quad F(x) = G(ax + b).$$

Beweis:
Die Aussage folgt durch zweimalige Anwendung der Konvexitätsbedingung (siehe dazu David [1980, S. 73]). □

Dies impliziert also, daß die Äquivalenzklassen der Konvexordnung gerade den Lokations- und Skalenfamilien der Form

$$\left\{ F : F(x) = F_0 \left(\frac{x - \mu}{\sigma} \right), \sigma > 0, \mu \in \mathbb{R} \right\}$$

mit einer stetigen Verteilung $F_0 \in \mathcal{F}_S$ entsprechen.
Für eine leichtere algebraische Umsetzung der Konvexitätsbedingung wird sich später das folgende Lemma als nützlich erweisen:

Lemma 5.11
Seien $F, G \in \mathcal{F}_S$, dann gilt

$$F \overset{k}{<} G \Leftrightarrow \left(\frac{Q_G'(x)}{Q_F'(x)} \right)' \geq 0, \qquad x \in (0, 1), \tag{5.6}$$

wobei Q_F und Q_G die Quantilfunktionen von F und G sind.

Beweis:
Da nach Definition der Konvexität

$$\begin{aligned} F \overset{k}{<} G \quad &\Leftrightarrow \quad \Psi(x) \quad \text{konvex auf } D_f \\ &\Leftrightarrow \quad \Psi''(x) \geq 0, \quad x \in D_f \end{aligned}$$

gilt, folgt die Behauptung aus den Umformungen:

$$\begin{aligned}
\Psi'(x) &= (Q_G(F(x)))' \\
&= (Q'_G(F(x))F'(x) \\
&= \frac{Q'_G(F(x))}{Q'_F(F(x))} \\
&= \frac{Q'_G(y)}{Q'_F(y)}, \quad \text{mit } y := F(x) \in (0,1).
\end{aligned}$$

□

Wie auch schon bei der Dispersionsordnung läßt sich die Bedingung (5.6) im Falle des fQ-Systems durch einfache Ungleichungen an die Parameter der fQ-Funktionen beschreiben.

Satz 5.12

Seien $F, G \in \mathcal{FQ}$ mit $F = \mathcal{F}(\mu_1, \sigma_1, \alpha_1, \beta_1)$ und $G = \mathcal{F}(\mu_2, \sigma_2, \alpha_2, \beta_2)$, dann gilt

$$F \overset{k}{<} G \Leftrightarrow \alpha_1 \geq \alpha_2 \ \wedge \ \beta_1 \leq \beta_2. \tag{5.7}$$

Beweis:

Aufgrund von Lemma 5.10 folgt zunächst, daß beim Vergleich zweier fQ-Verteilungen die Parameter σ (als Skalenparameter) und μ (als Lokationsparameter) keine Rolle spielen, und somit o.B.d.A. $\sigma_i = 1$ und $\mu_i = 0, i = 1, 2$ gesetzt werden können.

Aus

$$Q'_F(x) = x^{-\alpha_1}(1-x)^{-\beta_1}$$

und der analogen Darstellung für Q'_G folgt zunächst

$$\frac{Q'_G(x)}{Q'_F(x)} = x^{\alpha_1 - \alpha_2}(1-x)^{\beta_1 - \beta_2}.$$

Mit der vereinfachenden Schreibweise $\alpha^* := \alpha_1 - \alpha_2, \beta^* := \beta_1 - \beta_2$ folgt dann mit Lemma 5.11:

$$\begin{array}{lll} & F \overset{k}{<} G & \\ \Leftrightarrow & \left(\frac{Q'_G(x)}{Q'_F(x)}\right)' \geq 0 & x \in (0,1) \\ \Leftrightarrow & \left(x^{\alpha^*}(1-x)^{\beta^*}\right)' \geq 0 & x \in (0,1) \\ \Leftrightarrow & \alpha^* x^{\alpha^*-1}(1-x)^{\beta^*} \geq \beta^* x^{\alpha^*}(1-x)^{\beta^*-1} & x \in (0,1) \\ \Leftrightarrow & \alpha^*(1-x) \geq \beta^* x & x \in (0,1). \end{array}$$

Mit der Substitution

$$x \in (0,1) \rightarrow y := \frac{1-x}{x} > 0$$

ist dies äquivalent zu der Bedingung

$$\alpha^* y \geq \beta^*, \quad y > 0,$$

die für alle $y > 0$ nur dann erfüllt sein kann, wenn α^* nichtnegativ und β^* nichtpositiv ist, also

$$\alpha^* \geq 0 \quad \text{und} \quad \beta^* \leq 0,$$

woraus die Behauptung folgt. □

Das folgende Lemma wird zur Interpretation der Konvexordnung für die entsprechenden Dichtefunktionen benötigt:

Lemma 5.13
Seien $F \overset{k}{<} G \in \mathcal{FQ}$ zwei fQ-Verteilungen mit $\mu_F = \mu_G = 0$. Weiter seien $\sigma_F = 1$ und σ_G so gewählt, daß gilt

$$f(0) = g(0).$$

Dann gilt

$$F(x) \geq G(x), \quad \forall x \in D_f \cap D_g.$$

Beweis:
Nach Groeneveld und Meeden (1984) gilt unter den Voraussetzungen des Lemmas

$$F \overset{k}{<} G \Leftrightarrow h(x) := \begin{cases} \frac{\Psi(x)}{x} & , \quad x \neq 0 \\ \frac{f(0)}{g(0)} & , \quad x = 0 \end{cases}$$

ist monoton wachsend in $D_f \cap D_g$. Somit folgt

$$h(x) \begin{cases} \geq & 1, \quad x \geq 0 \\ \leq & 1, \quad x \leq 0 \end{cases},$$

und daraus unmittelbar die Behauptung. $\square$

Die Bedingungen an die Parameter in diesem Lemma stellen keine Einschränkung dar, da Lemma 5.10 die Invarianz der Konvexordnung unter affinen Transformationen sichert.
Mit diesem Satz folgt also aus $F \overset{k}{<} G$

$$Q_F(1/2) - Q_F(p) \quad \geq \quad Q_G(1/2) - Q_G(p), \quad p < 1/2 \tag{5.8}$$
$$Q_F(p) - Q_F(1/2) \quad \leq \quad Q_G(p) - Q_G(1/2), \quad p > 1/2. \tag{5.9}$$

Somit ist das u-Quantil für $p < 1/2$ ($p > 1/2$) von G nicht weiter (näher) vom Median entfernt als das entsprechende u-Quantil von F zum Median von F. Dies entspricht der anschaulichen Definition des Begriffs *Rechtsschiefe* (oder *Linkssteilheit*), wie anhand des folgenden Beispiels gezeigt wird. In der Abbildung sind die Dichten der symmetrischen Normalverteilung und der rechtsschiefen $F(4, 10)$-Verteilung dargestellt, wobei die Verteilungen linear transformiert wurden, um die Voraussetzungen von Lemma 5.13 zu erfüllen.

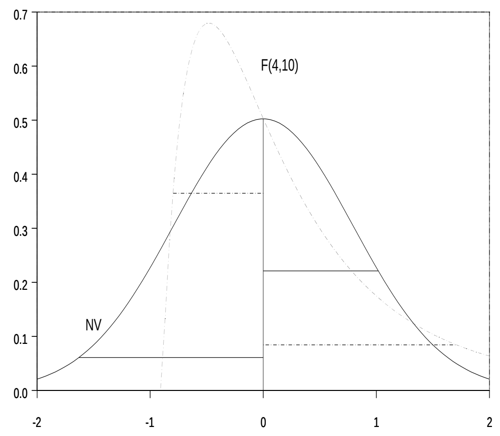

Abbildung 5.1: Vergleich der Normalverteilung und der $F(4, 10)$-Verteilung

Anhand der eingezeichneten Abstände der Quantile zum Median $\mu = 0$ für die Beipiele $p = 0.02$ und $p = 0.9$ kann man sich davon überzeugen, daß die Beziehungen (5.8) und (5.9) erfüllt sind.

5.3 Tailordnung des fQ-Systems

Während die Dispersionsordnung und die Konvexordnung dazu dienen, Verteilungen anhand von globalen Merkmalen zu vergleichen, die für den ganzen Verlauf der Dichten oder Verteilungsfunktionen Gültigkeit haben, werden die *Tailordnungen* herangezogen, um das unterschiedliche Randverhalten der Dichten bezüglich *eines* Randes zu untersuchen.
Eine der Anwendungen, die Parzen (1979b) in seiner Arbeit über die fQ-Funktion untersucht, ist die Definition einer derartigen Tailordnung mit Hilfe einer asymptotischen Aussage über das Randverhalten der fQ-Funktion. Im Zusammenhang mit dem fQ-System wird hier eine Verallgemeinerung dieses Ansatzes verwendet, die von Alzaid und Al-Osh (1989) eingeführt wurde:

Definition 5.14
Sei $F \in \mathcal{F}_S$, dann heißt die Funktion $\gamma_F(p) : [\,0,1] \to \mathbb{R}$ mit

$$\gamma_F(p) = \frac{pJ_F(1-p)}{fQ_F(1-p)} \tag{5.10}$$

die Exponentfunktion *von F. Dabei bezeichnet J_F die Scorefunktion von F.*

Bemerkung 5.15
Alzaid und Al-Osh zeigen in ihrer Arbeit, daß zwischen der Exponentfunktion und der Konvexordnung eine enge Beziehung besteht, da für zwei stetige Verteilungen $F, G, \in \mathcal{F}_S$ gilt:

$$F \overset{k}{<} G \Leftrightarrow \gamma_F(p) \leq \gamma_G(p), \quad 0 \leq p \leq 1.$$

Da jedoch eine äquivalente Darstellung dieser Bedingung im fQ-System analytisch aufwendiger ist, wurde bei der Definition der Konvexordnung die Ψ-Funktion verwendet.

Mit Hilfe der Exponentfunktion kann nun folgende Relation definiert werden:

Definition 5.16
Seien $F, G \in \mathcal{F}_S$, dann ist F tailkleiner *als G,*

$$F \overset{p}{<} G,$$

wenn gilt

$$\lim_{p \to 0+} \gamma_F(p) =: \gamma_F(0) < \gamma_G(0).$$

F ist tailäquivalent *zu G,*

$$F \overset{p}{\sim} G,$$

wenn gilt

$$\gamma_F(0) = \gamma_G(0).$$

Beide Begriffe werden analog auf Zufallsvariablen X, Y mit $X \sim F$ und $Y \sim G$ übertragen.

Wie in der Literatur üblich, wird die Tailordnung für den rechten Rand der Verteilung $Q(1)$ definiert. Dies liegt zum Teil daran, daß in den Anwendungen dieser Begriff häufig für das Verhalten von Extremwertstatistiken $X_{(n)}$ herangezogen wird. Analog kann die Tailordnung durch Betrachtung von

$$\gamma_F(1) := \lim_{p \to 1-} \gamma_F(p)$$

auf den linken Rand $Q(0)$ übertragen werden. In Würdigung der Arbeit von Parzen hat sich der Begriff *Parzen-Ordnung* für die Relation $\overset{p}{<}$ eingebürgert.
Aus Bemerkung 5.15 und den entsprechenden Aussagen für die Konvexordnung folgt unmittelbar

Lemma 5.17
a) Die Parzen-Ordnung ist eine Halbordnung.
b) Die Parzen-Ordnung ist lokations- und skaleninvariant.

Beweis:
Übertragung der Aussage von Lemma 5.10. □

Die Umsetzung der charakterisierenden Bedingung innerhalb des fQ-Systems ist aufgrund der Struktur der Funktion γ_F trivial:

Satz 5.18
Seien $F, G \in \mathcal{FQ}$ mit $F = \mathcal{F}(\mu_1, \sigma_1, \gamma_1, \beta_1)$ und $G = \mathcal{F}(\mu_2, \sigma_2, \gamma_2, \beta_2)$, dann gilt

$$F \overset{p}{<} G \Leftrightarrow \beta_1 < \beta_2. \tag{5.11}$$

Beweis:
Aufgrund der Beziehung (3.8) aus Lemma 3.5 gilt

$$\begin{aligned} \gamma_F(p) &= p\frac{J_F(1-p)}{fQ_F(1-p)} \\ &= p\left(\frac{\beta_1}{p} - \frac{\gamma_1}{1-p}\right) \\ &= \beta_1 - \gamma_1\frac{p}{1-p}, \end{aligned}$$

und somit

$$\gamma_F(0) = \lim_{p \to 0+} \gamma_F(p) = \beta_1,$$

woraus unmittelbar die Behauptung folgt. □

Dieses Ergebnis entspricht der intuitiven Erwartung, da die Parzen-Ordnung ein Spezialfall der Konvexordnung mit der Einschränkung auf den rechten Rand der Verteilung ist.

5.4 Metriken

Während sich die Motivation und Definition der Ordnungsbegriffe im vorangegangenen Kapitel an der direkten *statistischen* Interpretation der eingeführten Begriffe orientiert, werden nun die *funktionalanalytischen* Eigenschaften des fQ-Systems herangezogen, um zwei Metriken und ein Ähnlichkeitsmaß zu definieren.
Allen vorgestellten Begriffen ist gemein, daß die bei der Definition benutzte fQ-Funktion beziehungsweise ihre Ableitung, die Scorefunktion, als stetige Funktionen auf dem Intervall $[0,1]$ aufgefaßt werden. Somit können die bekannten Begriffe und Ergebnisse für diesen Raum direkt übertragen werden.

Definition und Satz 5.19
Der Raum der auf dem Intervall $[0,1]$ *stetigen, quadratintegrierbaren Funktionen wird mit der Abbildung*

$$d(g,h) := \left(\int_0^1 (g(t) - h(t))^2 \, dt \right)^{1/2} \tag{5.12}$$

zu einem metrischen Raum, der mit dem Symbol

$$\mathbf{C}_{L_2}([0,1])$$

bezeichnet wird. Mit der punktweisen Addition und Multiplikation, dem Skalarprodukt

$$< g, h > := \int_0^1 g(t)h(t) \, dt$$

und der Norm

$$\|g\| := < g, g >$$

wird $\mathbf{C}_{L_2}([0,1])$ *zu einem* Prähilbertraum.

Beweis:
Bronstein, Semendjajew [1984b, S. 1ff]. □

5.4.1 Die Metrik nach Hájek und Šidák

Die erste vorgestelle Metrik benutzt als Grundinformation eine normierte Variante der Scorefunktion $J = -fQ'$, wobei nur die Verteilungen mit endlicher Fisher-Informationen $F \in \mathcal{F}_I$ betrachtet werden können. Mit dieser Einschränkung läßt sich die Scorefunktion folgendermaßen standardisieren:

Definition 5.20
Sei $F \in \mathcal{F}_I$, dann heißt die Funktion

$$J^*(p) := \frac{fQ'(p)}{\sqrt{I(F)}} = \frac{-J(p)}{\sqrt{\|J\|}}, \quad p \in (0,1)$$

die normierte Scorefunktion *von F.*

Der Begriff *normierte* Scorefunktion wird dabei durch folgenden Satz motiviert:

Satz 5.21
Für die normierte Scorefunktion einer Verteilung $F \in \mathcal{F}_I$ gelten die Aussagen:

$$\int_0^1 J^*(p)\, dp = 0 \quad \textit{und} \quad \int_0^1 (J^*(p))^2 \; dp = 1. \tag{5.13}$$

Beweis:
Hájek und Šidák [1967, S. 19]. □

Da nach (5.13) somit $J^* \in \mathbf{C}_{L_2}([\,0,1])$ gilt, kann durch Einschränkung der allgemeinen Metrik d (5.12) auf die Untermenge $\mathcal{F}_I$ folgender Begriff definiert werden:

Definition und Satz 5.22
Auf dem Raum $\mathcal{F}_I$ definiert die Funktion

$$d_J(F,G) := \left(\int_0^1 (J_F^*(p) - J_G^*(p))^2 \; dp \right)^{1/2} \tag{5.14}$$

eine Metrik, die mit dem Begriff Score-Metrik *bezeichnet sei.*

Beweis:

Anwendung von Satz 5.19 auf $\mathcal{F}_I$. □

Die unmittelbare Bedeutung dieses Abstandsbegriffs liegt darin, daß er als Indikator für die Zugehörigkeit zu einer Lokations- und Skalenfamilie benutzt werden kann:

Lemma 5.23

Seien $F, G \in \mathcal{F}_I$ Elemente einer Lokations- und Skalenfamilie, dann gilt

$$d_J(F, G) = 0.$$

Beweis:

Nach Lemma 3.3 gilt für die Scorefunktionen

$$J_F(p) = \frac{J_G(p)}{\sigma},$$

und somit für die Fisher-Informationen

$$I(F) = \frac{I(G)}{\sigma^2},$$

woraus mit der Definition von J^* die Behauptung folgt. □

Die besondere Bedeutung der Funktion d_J wird erkennbar, wenn man aus dieser Metrik ein *Ähnlichkeitsmaß* ableitet:

Für $F, G \in \mathcal{F}_I$ gilt

$$\begin{aligned} d_J^2(F, G) &= \int_0^1 \left(J_F^*(p) - J_G^*(p)\right)^2 \, dp \\ &= \int_0^1 \left(J_F^*(p)\right)^2 \, dp - 2\int_0^1 J_F^*(p) J_G^*(p) \, dp + \int_0^1 \left(J_G^*(p)\right)^2 \, dp \\ &= 2\left(1 - \int_0^1 J_F^*(p) J_G^*(p) \, dp\right), \end{aligned}$$

wobei die letzte Gleichung aus (5.13) folgt.

Mit der Bezeichnung

$$\rho(F, G) := \int_0^1 J_F^*(p) J_G^*(p) \, dp \tag{5.15}$$

gilt somit

$$d_J^2(F,G) = 2(1 - \rho(F,G)),$$

und aufgrund der Ungleichung

$$\begin{aligned} & (1 - \rho(F,G))^2 && \geq 0 \\ \Leftrightarrow \quad & \rho^2(F,G) - 1 && \geq 2\rho(F,G) - 2 \end{aligned}$$

auch die Abschätzung

$$d_J^2(F,G) \geq 1 - \rho^2(F,G).$$

Da sich die Funktion $\rho(F,G)$ mit den Bezeichnungen aus den Definitionen 5.19 und 5.20 auch in der Form

$$\begin{aligned} \rho(F,G) &= \frac{\int_0^1 J_F(p) J_G(p)\, dp}{\sqrt{\int_0^1 J_F^2(p)\, dp \int_0^1 J_G^2(p)\, dp}} \\ &= \frac{< J_F, J_G >}{\sqrt{\|J_F\|\, \|J_G\|}} \end{aligned}$$

schreiben läßt, definiert sie in der Tat ein Ähnlichkeitsmaß, da sie die Korrelation zweier Elemente F,G aus $\mathcal{F}_I \subset \mathbf{C}_{L_2}([\,0,1])$ mißt.
Im Rahmen der nichtparametrischen statistischen Verfahren kann die Funktion $\rho(F,G)$ direkt genutzt werden, da folgende Aussage gilt:

Lemma 5.24
Für $F,G \in \mathcal{F}_I$ seien T_F, T_G die lokal optimalen Rangtests, die mit Hilfe der Scorefunktionen J_F und J_G konstruiert werden. Dann gilt für die asymptotische relative Effizienz von T_F bezüglich T_G, wenn G die tatsächliche Verteilung der Grundgesamtheit ist:

$$ARE(T_F, T_G|G) = \rho(F,G).$$

Beweis:
Siehe Hájek und Šidák [1967, S. 268ff]. □

Da sich im Falle des fQ-Systems sowohl die Integrale über die fQ-Funktion als auch über die Scorefunktion durch die vollständige Betafunktion ausdrücken lassen, kann die Funktion ρ ebenfalls durch einen geschlossenen Term dargestellt werden:

Satz 5.25

Seien $F, G \in \mathcal{FQ}_I$ mit $F = \mathcal{F}(\mu_1, \sigma_1, \alpha_1, \beta_1)$ und $G = \mathcal{F}(\mu_2, \sigma_2, \alpha_2, \beta_2)$. Dann gilt für die Korrelation von F und G

$$\begin{aligned}\rho(F,G) &= \sqrt{\frac{(2\alpha_1-1)(2\beta_1-1)(2\alpha_2-1)(2\beta_2-1)}{4\alpha_1\beta_1\alpha_2\beta_2(\alpha_1+\beta_1-1)(\alpha_2+\beta_2-1)}} \\ &\quad \cdot \frac{\alpha_1\beta_2(\beta_2-1)(\alpha_1-1)+\alpha_2\beta_1(\beta_1-1)(\alpha_2-1)}{(\beta_1+\beta_2-1)(\alpha_1+\alpha_2-1)} \\ &\quad \cdot \frac{B(\alpha_1+\alpha_2, \beta_1+\beta_2)}{\sqrt{B(2\alpha_1, 2\beta_1)B(2\alpha_2, 2\beta_2)}}. \qquad (5.16)\end{aligned}$$

Beweis:

Aufgrund der Invarianzaussage in Lemma 5.23 kann o.B.d.A. $\sigma_1 = \sigma_2 = 1$ angenommen werden.

Die Darstellung für die Fisher-Information für fQ-Verteilungen wurde bereits in Satz 3.7 hergeleitet und wurde in obiger Darstellung direkt übernommen. Mit der Darstellung (3.7) aus Lemma 3.5 für die Scorefunktion

$$J_F(p) = \beta_1 p^{\alpha_1}(1-p)^{\beta_1-1} - \alpha_1 p^{\alpha_1-1}(1-p)^{\beta_1}$$

und der entsprechenden Formel für $J_G(p)$ kann der Ausdruck

$$< J_F, J_G > = \int_0^1 J_F(p) J_G(p)\, dp$$

durch Ausmultiplizieren der Terme unter Berücksichtigung der Beziehung

$$\int_0^1 p^a(1-p)^b\, dp = B(a+1, b+1)$$

in folgender Weise dargestellt werden:

$$\begin{aligned}< J_F, J_G > &= \alpha_1\alpha_2 B(\alpha_1+\alpha_2-1, \beta_1+\beta_2+1) \\ &\quad -(\alpha_1\beta_2+\alpha_2\beta_1)B(\alpha_1+\alpha_2, \beta_1+\beta_2) \\ &\quad +\beta_1\beta_2 B(\alpha_1+\alpha_2+1, \beta_1+\beta_2-1).\end{aligned}$$

Durch Anwendung der Beziehung

$$B(a,b) = \frac{b-1}{a} B(a+1, b-1)$$

können dann der erste und der letzte Summand ebenfalls auf die Form

$$T_i(\alpha_1, \alpha_2, \beta_1, \beta_2) \cdot B(\alpha_1 + \alpha_2, \beta_1 + \beta_2), \quad i = 1, 2$$

mit zwei algebraischen Termen T_i gebracht werden. Durch weitere elementare Umformungen und Zusammenfassungen folgt die Aussage des Satzes. □

5.4.2 Die fQ-Metrik

Wie im letzten Abschnitt betont wurde, liegt die Bedeutung der Score-Metrik darin, daß sie ein Ähnlichkeitsmaß definiert, welches im *analytischen* Sinne in der nichtparametrischen Statistik genutzt werden kann. Demgegenüber orientiert sich die nun vorgestellte fQ-Metrik an der Grundintention des fQ-Systems und definiert ein Abstandsmaß, welches sich *graphisch* interpretieren läßt. Unter Berücksichtigung der späteren Anwendungen wird dabei nur der Raum der unimodalen Verteilungen betrachtet.

Definition und Satz 5.26
Seien $F, G \in \mathcal{F}_U$ mit fQ-Funktionen fQ_F und fQ_G. Dann wird durch die Funktion

$$d_{fQ}(F, G) := \left(\int_0^1 (fQ_F(p) - fQ_G(p))^2 \; dp \right)^{1/2} \tag{5.17}$$

eine Metrik definiert, die mit dem Begriff fQ-Metrik *bezeichnet sei.*

Beweis:
Aufgrund der Unimodalität gibt es zunächst für $F \in \mathcal{F}_U$ ein $K \in \mathbb{R}$ mit

$$f(x) \leq K, \quad \forall x \in D_f.$$

Somit folgt:

$$\begin{aligned} fQ_F(p) &\leq K, \quad p \in [0,1] \\ \Rightarrow \int_0^1 fQ_F^2(p)\, dp &\leq K^2, \end{aligned}$$

also $fQ_F \in \mathbf{C}_{L_2}([0,1])$. Nach Satz 5.19 definiert somit d_{fQ} als Einschränkung von $d(\cdot,\cdot)$ (5.12) auf dem Unterraum $\mathcal{F}_U$ eine Metrik. □

Im Vergleich zur Score-Metrik wird also hier der quadratische Abstand der nicht-standardisierten fQ-Funktionen gemessen. Dies entspricht der Konstruktionsidee des fQ-Systems, nach der Verteilungen, deren Gestalt ähnlich ist, auch durch ähnliche Parameter repräsentiert werden.
Analog zu Satz 5.25 kann d_{fQ} im Falle von $F, G \in \mathcal{FQ}_U$ geschlossen dargestellt werden, wobei die resultierende Formel aufgrund der fehlenden Standardisierung wesentlich einfacher ist.

Satz 5.27
Seien $F, G \in \mathcal{FQ}_U$ mit $F = \mathcal{F}(\mu_1, \sigma_1, \alpha_1, \beta_1)$ und $G = \mathcal{F}(\mu_2, \sigma_2, \alpha_2, \beta_2)$. Dann gilt

$$\begin{aligned} d_{fQ}^2(F,G) &= \sigma_1^2 B(2\alpha_1+1, 2\beta_1+1) - 2\sigma_1\sigma_2 B(\alpha_1+\alpha_2+1, \beta_1+\beta_2+1) \\ &\quad +\sigma_2^2 B(2\alpha_2+1, 2\beta_2+1). \end{aligned} \tag{5.18}$$

Beweis:
Die Behauptung folgt wie in Satz 5.25 durch Einsetzen der fQ-Darstellungen und Integration der ausquadrierten Ausdrücke. □

Man beachte, daß bei der fQ-Metrik im Gegensatz zur Score-Metrik auch der Skalenparameter σ berücksichtigt wird und diese somit nicht mehr skaleninvariant ist. In Kapitel 6 wird d_{fQ} herangezogen, um ein Gütemaß bei der Approximation von stetigen Verteilungen durch fQ-Verteilungen zu definieren.

5.5 Klassifikationsmaße für Verteilungen

Mittels der in den vorangehenden Abschnitten beschriebenen Konzepte der Ordnungsbegriffe und der Metriken werden analytische Werkzeuge bereitgestellt, mit denen quantitative und qualitative Vergleiche zwischen fQ-Verteilungen durchgeführt werden können. Neben dieser Fragestellung ist häufig die Beschreibung von einzelnen Verteilungen im Hinblick auf globale Charakteristika von Bedeutung. Büning (1991) unterscheidet dabei zwischen den *Primärmaßen* (Lage- und Skalenmaßen) und den *Sekundärmaßen*, mit denen die Gestaltparameter der Verteilung wie zum Beispiel die Schiefe beschrieben werden.

Der Sinn dieser Maßbegriffe liegt einerseits darin, mit ihnen verschiedene theoretische Verteilungstypen vergleichbar zu machen, und andererseits darin, strukturelle Aussagen über die zugrundeliegende Verteilung einer gegebenen Datenmenge anhand dieser Maße zu treffen. Eine Anwendung dieser Methode liegt in der Analyse von Zeitreihendaten, bei denen a priori keine Verteilungsinformationen bekannt sind und eine dynamische Änderung des Verteilungstyps nicht ausgeschlossen werden kann.
Da die Konstruktion des fQ-Systems gerade auf einer globalen Eigenschaft von Verteilungen (der approximativen Darstellung der fQ-Funktion) basiert, werden im folgenden Sekundärmaße untersucht.

Sekundärmaße

Nach Büning [1991, S. 224] und (1985) können diese Maße grundsätzlich nach zwei Gesichtpunkten eingeteilt werden: nach dem Konstruktionsprinzip des Maßes und der Art des Gestaltparameters. Folgende Maße mögen als Beispiele für diese Einteilung herangezogen werden:

- Die klassischen Momentenschätzer für die *Schiefe* und die *Kurtosis* (Wölbung)
$$\beta_1 := \frac{\mu_3}{\sigma^3}, \quad \beta_2 := \frac{\mu_4}{\sigma^4}.$$
- Der Schätzer H_2 für die *Peakedness*, der in Analogie zu einem entsprechenden empirischen Maß von Hogg (1974, 1982) die Gestalt
$$H_2 := \frac{\mu(0.8; 0.95) - \mu(0.05; 0.2)}{\mu(0.5; 0.8) - \mu(0.2; 0.5)}$$
hat, wobei
$$\mu(a; b) := \frac{1}{b-a} \int_a^b Q(p)\, dp, \quad 0 < a < b,$$
ein verallgemeinerter Mittelwert ist.
- Schätzer für die *Tailstärke*, die mit Hilfe von Quantilen konstruiert werden:
$$T(a, b) := \frac{Q(1-a) - Q(a)}{Q(1-b) - Q(b)}, \quad 0 < a < b < 0.5.$$

Während der Begriff der *Schiefe* zumindest inhaltlich wohldefiniert ist, gibt es in der Literatur vollkommen unterschiedliche Auffassungen darüber, was mit den Begriffen *Tailstärke* und *Peakedness* bezeichnet wird bzw. was mit diesen Maßen gemessen werden soll (vgl. dazu die Ausführungen von Büning [1991, S. 230ff]). Die Peakedness - so Büning - beschreibt die Größe der Konzentration der Verteilung um den Median, während die Tailstärke das Randverhalten der Verteilung beschreibt, wobei diese Aussagen nur durch heuristische Beobachtungen gestützt werden. Neben dieser prinzipiellen inhaltlichen Problematik listet Büning noch weitere analytische Nachteile auf: (Büning [1991, S. 227ff])

- Die Kurtosis - als Vertreter des ersten Konstruktionsprinzips - ist als Klassifikationswerkzeug gänzlich ungeeignet, da Verteilungen mit vollkommen unterschiedlichen Dichten eine gleiche Kurtosis aufweisen können.

- Einige der bekannten Maßbegriffe setzen die Existenz der ersten beiden Momente voraus. Damit ist die Untersuchung vieler Verteilungsklassen grundsätzlich ausgeschlossen.

- Die Maße, die mit Hilfe der verallgemeinerten Mittelwerte konstruiert werden, sind ausnahmslos von subjektiv gewählten Parametern a und b abhängig, die nicht kanonisch festgelgt werden.

Der letzte Punkt dieser Auflistung bedeutet insbesondere, daß unterschiedliche Maßbegriffe μ_1, μ_2 verschiedene Ordnungen der Verteilungen induzieren können:

$$\mu_1(F) < \mu_1(G) \not\Rightarrow \mu_2(F) < \mu_2(G).$$

Als Ausweg schlägt Büning vor, nur solche Maße μ zu betrachten, die sich aus einer entsprechenden Halbordnung $<$ für Verteilungen konstruieren lassen, also mit

$$F < G \Rightarrow \mu(F) < \mu(G),$$

wobei nicht notwendigerweise auch die Rückrichtung ($\Leftarrow$) gelten muß. Dieses Konstruktionsprinzip wird bei den folgenden Begriffen berücksichtigt.

5.6 Maßzahlen für unimodale Verteilungen

Da die Konstruktionsidee des fQ-Systems darauf zielt, stetige unimodale Verteilungen gut zu approximieren, werden im folgenden zwei Maßzahlen definiert, die von der Eigenschaft der Unimodalität Gebrauch machen.
Wesentlich für die beiden in diesem Abschnitt vorgestellten Begriffe der Schiefe und der Dispersion ist sowohl der enge Bezug zu den Ordnungsbegriffen aus Kapitel 5.2 als auch die Tatsache, daß sich die beiden Maße direkt graphisch interpretieren lassen.

Ein Schiefemaß

Ein anschaulich leicht nachvollziehbares Maß für die Schiefe einer unimodalen Verteilung F läßt sich aus der Überlegung ableiten, daß die Differenz zwischen dem Modus der Verteilung

$$x_{mod} := \arg\max_{x \in D_f} f(x)$$

und dem Median $\mu = Q(1/2)$ direkt interpretierbar ist, da gilt

$$d := \mu - x_{mod} \begin{cases} > & 0 \\ = & 0 \\ < & 0 \end{cases} \Leftrightarrow \quad F \text{ ist} \quad \begin{cases} \text{rechtsschief} \\ \text{symmetrisch} \\ \text{linksschief} \end{cases} . \tag{5.19}$$

Um eine sinnvolle Skalierung von d im Sinne eines Maßbegriffs zu erreichen, können die Größen x_{mod} und μ durch ihrer Quantilwerte ersetzt werden, da die Monotonie der Verteilungsfunktion nichts an den qualitativen Aussagen (5.19) ändert. Insgesamt wird schließlich folgende Definition motiviert:

Definition 5.28
Sei $F \in \mathcal{F}_U$ eine stetige unimodale Verteilung, dann sei mit

$$\mu_S(F) := 1 - 2F(x_{mod}) \tag{5.20}$$

die Schiefe *von F bezeichnet.*

Die Skalierung wurde dabei so gewählt, daß gilt:

$$\mu_S(F) \in [-1, 1].$$

Unmittelbar klar ist, daß diese Skalierung die Interpretationen (5.19)

$$\mu_S(F) \begin{cases} > & 0 \\ = & 0 \\ < & 0 \end{cases} \Leftrightarrow \quad F \text{ ist } \begin{cases} \text{rechtsschief} \\ \text{symmetrisch} \\ \text{linksschief} \end{cases}$$

unverändert läßt.
Ebenso wie das klassische Schiefemaß $\beta_1 := \frac{\mu_3}{\sigma^3}$ ist $\mu_S(F)$ eine normierte Größe, die es ermöglicht, Verteilungen mit unterschiedlicher Streuung zu vergleichen. Der Vorteil des Maßes $\mu_S(F)$ besteht darin, daß es auch dann anwendbar ist, wenn die Momente der Verteilung nicht existieren.

Das Dispersionsmaß

Mit dem zweiten Klassifikationsmaß soll hier ein neuer, allgemeiner Begriff eingeführt werden, mit dem das Streuungsverhalten einer unimodalen Verteilung gemessen werden kann, wobei bei der Definition unmittelbar die Dispersionsordnung aus Kapitel 5.2 benutzt wird.
Nach Satz 5.5 und Bemerkung 5.7 gilt für $F, G \in \mathcal{F}_U$

$$F \overset{d}{<} G \Leftrightarrow fQ_F(p) \geq fQ_G(p), \quad \forall p \in (0,1).$$

Anschaulich bedeutet dies, daß die Dispersionordnung direkt mit dem Flächeninhalt der fQ-Funktionen korrespondiert, wodurch folgender Begriff motiviert ist:

Definition 5.29
Sei $F \in \mathcal{F}_U$. Dann heißt die Funktion $\mu_D : \mathcal{F}_U \to \mathbb{R}_+$ mit

$$\mu_D(F) := \left(\int_0^1 fQ_F(p)\, dp \right)^{-1} \tag{5.21}$$

die Dispersion *von F.*

Da für zwei Verteilungen $F, G \in \mathcal{F}_U$ mit

$$F(x) = G\left(\frac{x-\mu}{\sigma}\right),\ \sigma > 0, \mu \in \mathbb{R}$$

$$fQ_G(p) = \frac{1}{\sigma} fQ_F(p)$$

und somit auch

$$\mu_D(G) = \sigma \cdot \mu_D(F)$$

gilt, ist das Dispersionsmaß zwar lokations- aber *nicht* skaleninvariant. Dies unterscheidet das Dispersionsmaß von allen anderen skaleninvarianten Maßen, die Büning [1991, S. 224ff] in seiner Auflistung vorstellt.

Bemerkung 5.30
An dieser Stelle sei die prinzipielle Frage gestellt, ob ein sinnvolles Maß für die Streuung einer Verteilung überhaupt skaleninvariant sein kann. *Hat sich etwa als Ergebnis einer parametrischen oder nichtparametrischen Dichteschätzung anhand empirischer Daten eine Dichte ergeben, deren Struktur darauf hindeutet, daß die zugrundeliegende Verteilung viel Masse an den Rändern aufweist, so ist aufgrund dieser Feststellung allein zunächst vollkommen unklar, ob dies auf eine Verteilung mit kleineren Tails, zum Beispiel die Normalverteilung, und großer Varianz oder auf eine Verteilung mit großen Tails, zum Beispiel die Cauchyverteilung, und kleinem Skalenparameter zurückzuführen ist. Mit Hilfe eines skaleninvarianten Streumaßes ist diese Fragestellung prinzipiell nicht zu beantworten, da der skalare Anteil an der Streuung nicht geschätzt werden kann.*

Das vorgeschlagene Dispersionsmaß erlaubt eine weitere Unterscheidung in zwei multiplikative Anteile, die eine einfache inhaltliche Interpretation zulassen. Zur Motivation sei hier zunächst eine Abbildung der fQ-Funktionen der Normalverteilung mit Varianz $\sigma = (4/3)^2$ und der Cauchyverteilung mit annähernd gleicher Dispersion gegeben:

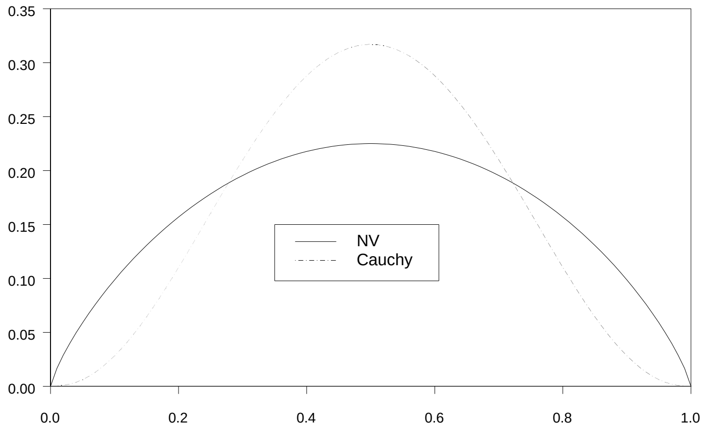

Abbildung 5.2: Zwei fQ-Funktionen mit gleichem Flächeninhalt

Beide Kurven besitzen zwar den gleichen Flächeninhalt, haben aber hinsichtlich der Konzentration ihrer *Wahrscheinlichkeitsmasse* offensichtlich eine unterschiedliche Struktur. Diese kann durch die Division der Funktionswerte durch den jeweiligen maximalen Funktionswert der Kurven und die damit verbundene Normierung der Kurven sichtbar gemacht werden:

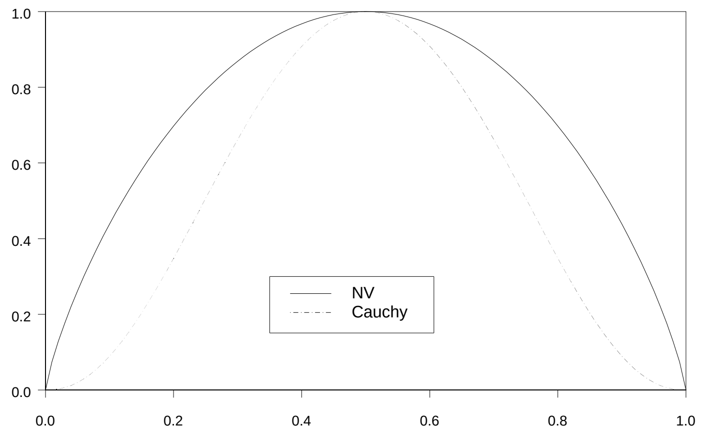

Abbildung 5.3: Normierte fQ-Funktionen

Die Betrachtung der Flächeninhalte dieser normierten Kurven sowie der Normierungsfaktor liefern nun Aussagen über die strukturelle Konzentration der Verteilung und auch eine Schätzung für den multiplikativen Skalierungsfaktor. Dies wird in folgender Definition zusammengefaßt:

Definition 5.31
Sei $F \in \mathcal{F}_U$ mit Modus x_{mod}, dann heißt die Funktionen $\mu_D^S : \mathcal{F}_U \to \mathbb{R}_+$ mit

$$\mu_D^S(F) := f(x_{mod})^{-1} \tag{5.22}$$

die Skalendispersion *von F und die Funktion $\mu_D^F : \mathcal{F}_U \to \mathbb{R}_+$ mit*

$$\mu_D^F(F) := \frac{\mu_D(F)}{\mu_D^S(F)} = \frac{f(x_{mod})}{\int_0^1 fQ(p)\,dp} \tag{5.23}$$

die Familiendispersion *von F.*

Die Wahl der Namen für diese Begriffe wird aufgrund der folgenden Überlegung gerechtfertigt:

Lemma 5.32
Seien $F, G \in \mathcal{F}_U$ mit

$$F(x) = G\left(\frac{x-\mu}{\sigma}\right),\ \sigma > 0, \mu \in \mathbb{R}.$$

Dann gelten folgende Aussagen:

$$\begin{aligned} \mu_D^S(F) &= \sigma^{-1}\,\mu_D^S(G) \\ \mu_D^F(F) &= \mu_D^F(G). \end{aligned} \tag{5.24}$$

Beweis:
Beide Aussagen folgen direkt aus der entsprechenden Beziehung für die Dichtefunktionen

$$f(x) = \frac{1}{\sigma}\,g\left(\frac{x-\mu}{\sigma}\right),$$

beziehungsweise für die der fQ-Funktionen. □

Die Beziehung (5.24) bedeutet somit, daß μ_D^F in der Tat invariant für Elemente einer Lokations- und Skalenfamilie ist. Die Skalendispersion, die dem Inversen des maximalen Funktionswertes der Dichtefunktion entspricht, ist für sich genommen natürlich nur ein grobes Maß für die Dispersion einer Verteilung. In Verbindung mit der Familiendispersion stellt sie jedoch ein nützliches Werkzeug dar, um im explorativen Sinne strukturelle Unterschiede zwischen Verteilungen aufzuzeigen. Für das obige Beispiel der Normalverteilung (NV) mit $\sigma = (4/3)^2$ und der Cauchyverteilung (C) ergeben sich die Werte

$$\begin{array}{rclrcl} \mu_D^F(NV) & \approx & 1.4106, \qquad & \mu_D^S(NV) & \approx & 4.454, \\ \mu_D^F(C) & = & 2, & \mu_D^S(C) & = & \pi, \end{array}$$

womit der optische Eindruck der größeren Varianz bei kleinerer Konzentration der Normalverteilung gegenüber der Cauchyverteilung bestätigt wird.

Darstellung im fQ-System

Der folgende Satz liefert die parametrische Darstellung der betrachteten Maße innerhalb des fQ-Systems, wobei wieder ausgenutzt wird, daß alle erforderlichen Hilfsgrößen in geschlossener funktionaler Form dargestellt werden können.

Satz 5.33

Sei $F = \mathcal{F}(\mu, \sigma, \alpha, \beta) \in \mathcal{FQ}_U$, dann gelten folgende Darstellungen

$$\mu_S(F) = \frac{\beta - \alpha}{\alpha + \beta}, \tag{5.25}$$

$$\mu_D(F) = \frac{1}{\sigma\, B(\alpha+1, \beta+1)}, \tag{5.26}$$

$$\mu_D^S(F) = \frac{(\alpha+\beta)^{\alpha+\beta}}{\sigma\, \alpha^\alpha \beta^\beta}, \tag{5.27}$$

$$\mu_D^F(F) = \frac{\alpha^\alpha \beta^\beta}{(\alpha+\beta)^{\alpha+\beta} B(\alpha+1, \beta+1)}. \tag{5.28}$$

Beweis:

Zunächst kann die fQ-Funktion von F

$$fQ(p) = \sigma p^{\alpha}(1-p)^{\beta} \tag{5.29}$$

als eine umskalierte Dichte einer Betaverteilung der Form

$$\tilde{f}(x) = B(\alpha+1, \beta+1)^{-1} x^{\alpha}(1-x)^{\beta} \tag{5.30}$$

aufgefaßt werden.

Nach Johnson, Kotz und Balakrishnan [(1995), S. 219] besitzt dann $\tilde{f}$ für $\alpha, \beta > 0$ ein eindeutiges Maximum im Punkt

$$\tilde{x}_{mod} = \frac{\alpha}{\alpha+\beta}. \tag{5.31}$$

Somit wird f, die Dichte von F, maximal an der Stelle $x_{mod} := Q(\tilde{x}_{mod})$ und es folgt

$$\mu_S(F) = 1 - 2F(x_{mod}) = 1 - 2\,\frac{\alpha}{\alpha+\beta},$$

also (5.25).

Die Behauptungen (5.26) - (5.28) ergeben sich durch Einsetzen von $\tilde{x}_{mod}$ in (5.29), beziehungsweise aus der Definition der Betafunktion.

$\square$

Mit dem folgenden Lemma wird der von Büning geforderte Zusammenhang (siehe Abschnitt 5.5) zwischen den entsprechenden Halbordnungen, der Konvexordnung und der Dispersionsordnung, und den beiden Maßbegriffen hergestellt:

Lemma 5.34

Seien $F, G \in \mathcal{FQ}_U$, dann gelten die Beziehungen

$$F \overset{d}{<} G \;\Rightarrow\; \mu_D(F) < \mu_D(G) \tag{5.32}$$

$$F \overset{k}{<} G \;\Leftrightarrow\; \mu_S(F) < \mu_S(G). \tag{5.33}$$

Beweis:

Während die erste Aussage unmittelbar aus der Definition folgt, wird zum Nachweis von (5.33) folgende Überlegung benötigt:

Nach Satz 5.12 gilt für $F = \mathcal{F}(\mu_1, \sigma_1, \alpha_1, \beta_1)$ und $G = \mathcal{F}(\mu_2, \sigma_2, \alpha_2, \beta_2)$

$$\begin{aligned}
F \overset{k}{<} G \quad &\Leftrightarrow \quad \alpha_1 \geq \alpha_2 \ \wedge \ \beta_1 \leq \beta_2 \\
&\Leftrightarrow \quad \frac{\alpha_1}{\alpha_2} \geq 1 \geq \frac{\beta_1}{\beta_2} \\
&\Leftrightarrow \quad \alpha_1\beta_2 + \beta_1\beta_2 \geq \alpha_2\beta_1 + \beta_1\beta_2 \\
&\Leftrightarrow \quad \frac{\beta_2}{\alpha_2 + \beta_2} \geq \frac{\beta_1}{\alpha_1 + \beta_1} \\
&\Leftrightarrow \quad \frac{2\beta_2}{\alpha_2 + \beta_2} - 1 \geq \frac{2\beta_1}{\alpha_1 + \beta_1} - 1 \\
&\Leftrightarrow \quad \frac{\beta_2 - \alpha_2}{\alpha_2 + \beta_2} \geq \frac{\beta_1 - \alpha_1}{\alpha_1 + \beta_1}.
\end{aligned}$$

□

Kapitel 6

Darstellung von Verteilungen

6.1 Die fQ-Approximation von unimodalen Verteilungen

Nachdem in den bisherigen Kapiteln die wesentlichen Eigenschaften des fQ-Systems als eigenständiges Verteilungssystem behandelt wurden, soll nun die eigentliche Grundidee der Darstellung von unimodalen Verteilungen wieder aufgegriffen werden. Mit der Definiton 5.26 der fQ-Metrik läßt sich dieses Problem formal folgendermaßen beschreiben:

Definition 6.1
Sei $F \in \mathcal{F}_U$ eine stetige unimodale Verteilung. Dann heißt eine Funktion $\hat{F} \in \mathcal{FQ}_U$ mit

$$\hat{F} := \arg\min_{G \in \mathcal{FQ}_U} d_{fQ}(F, G), \tag{6.1}$$

$$\hat{F}^{-1}(1/2) = F^{-1}(1/2) \tag{6.2}$$

die fQ-Approximation *von F.*

Da einige Verteilungen wie zum Beispiel die logistische Verteilung sich im fQ-System exakt darstellen lassen (siehe auch Tabelle 2.1), wird im folgenden für $\hat{F}$ synonym auch der Ausdruck *fQ-Darstellung* verwendet. Man beachte, daß die Bedingung (6.2) notwendig ist, da die fQ-Funktion unabhängig von einem Lokationsparameter μ ist.

Bemerkung 6.2

Faßt man $\mathcal{FQ}_U$ als Teilmenge des metrischen Raums $(\mathcal{F}_U, d_{fQ})$ auf, so ist für ein festes $F \in \mathcal{F}_U$ der Abstand zwischen F und $\mathcal{FQ}_U$ gegeben durch

$$d_{fQ}(F, \mathcal{FQ}_U) = \inf_{G \in \mathcal{FQ}_U} \{d_{fQ}(F, G)\}.$$

Aufgrund der Stetigkeit von d_{fQ} in $G \in \mathcal{FQ}_U$ und der damit verbundenen Abgeschlossenheit von $\mathcal{FQ}_U$ folgt, daß dieses Infimum in $\mathcal{FQ}_U$ angenommen wird, also die Existenz einer Lösung $\hat{F} \in \mathcal{FQ}_U$.
Man beachte, daß in dieser Definition nicht die Eindeutigkeit einer Lösung vorausgesetzt wird.

Zur numerischen Lösung dieses Problems kann das Verfahren von Gauß-Newton zur Minimierung von $d_{fQ}(F, G)$ bei festem $F \in \mathcal{F}_U$ bezüglich der Parameter σ, α und β aus $G = \mathcal{F}(\mu, \sigma, \alpha, \beta)$ herangezogen werden. Bei der Darstellung des fQ-Abstandes

$$d_{fQ}(F, G) := \left(\int_0^1 \left(fQ_F(p) - fQ_G(p) \right)^2 \, dp \right)^{1/2}$$

bieten sich dabei zwei Möglichkeiten an:

- Numerische Integration
- Approximation von d_{fQ} durch

  $$\tilde{d}_{fQ} := \frac{1}{\sqrt{n+1}} \left(\sum_{i=1}^{n} \left(fQ_F(p_i) - fQ_G(p_i) \right)^2 \right)^{1/2},$$

  mit $p_i = i/(n+1), i = 1, \ldots, n$.

Die folgende Tabelle zeigt einen Vergleich dieser Methoden für die Normalverteilung, wobei zusätzlich zu den Parameterschätzungen auch die benötigte Rechenzeit notiert wurde. Die Zeitangaben sind natürlich extrem abhängig von der eingesetzten Hardware und nur als relative Angaben zu verstehen. Die numerische Integration erfolgte mit einer adaptiven 15-Punkte Gauß-Quadraturformel, die in S-PLUS vordefiniert ist (siehe auch Dongarra und Grosse [1987]).

Verfahren	$\hat{\sigma}$	$\hat{\alpha}$	$\hat{\beta}$	Zeit (s)
Gauß-Q.	1.224078	0.8071276	0.8071276	8.9
n = 1000	1.224079	0.8071281	0.8071281	0.8
n = 100	1.223939	0.8070559	0.8070559	0.25

Da der Genauigkeitsverlust für den Fall $n = 1000$ im Vergleich zum Ergebnis bei der numerischen Integration durch den erheblichen Zeitgewinn aufgehoben wird, wird bei den folgenden Auswertungen stets diese Methode gewählt.

6.2 Umparametrisierung des fQ-Systems

Die definierenden Gleichungen (A) - (C) des fQ-Systems (siehe Kapitel 2.1) ergaben sich direkt aus der approximativen Darstellung der fQ-Funktion durch eine Betadichte. In dieser Form konnten einige Aussagen über globale Charakteristika der fQ-Verteilung direkt durch einfache Beziehungen für die Parameter (α, β, σ) gemacht werden. Der Vorteil dieser Parametrisierung liegt somit in der Möglichkeit einer unmittelbaren Interpretation der Parameter.
Ein gravierender Nachteil liegt jedoch in der Tatsache, daß die Parameter konstruktionsbedingt hochkorreliert sein können, da zum Beispiel bei einer Normalverteilung die approximierende vierparametrische fQ-Verteilung überparametrisiert ist. Um einen optischen Eindruck von diesem Phänomen zu bekommen, zeigt Abbildung 6.1 ein Konturdiagramm für den fQ-Abstand d_{fQ} der Normalverteilung.
In dieser Abbildung sind diejenigen Parameterkombinationen (α, σ) durch Höhenlinien verbunden, die zu identischen fQ-Abständen

$$d_{fQ}\left(fQ_{NV}, \mathcal{F}(0, \sigma, \alpha, \alpha)\right)$$

führen, wobei mit fQ_{NV} die fQ-Funktion der Normalverteilung bezeichnet sei. Durch die Einschränkung auf symmetrische fQ-Verteilungen ($\beta = \alpha$) wird diese zweidimensionale Darstellung ermöglicht.

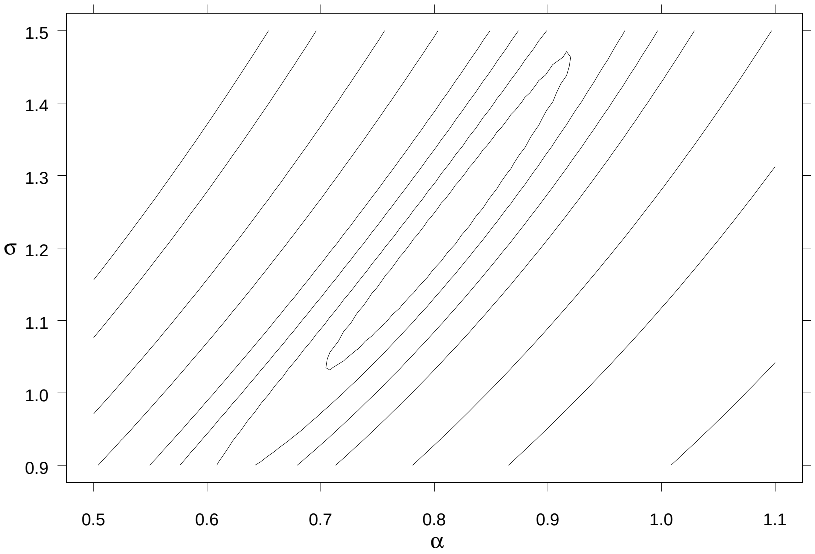

Abbildung 6.1: Konturdiagramm für $d_{fQ}(fQ_{NV}, \cdot)$ in (σ, α, β)-Parametrisierung

Die Lage der durch die Höhenlinien gebildeten Bereiche läßt auf eine hohe Korrelation der Parameter α und σ schließen. Für die Praxis bedeutet dies, daß eine Klassifikation von Verteilungen auf der Basis dieser Parametrisierung zu unbefriedigenden Ergebnissen führen muß.

Als Ausweg bietet es sich an, die Kennmaße für unimodale Verteilungen aus Abschnitt 5.6 zu verwenden, da durch die inhaltlich gegebene Unabhängigkeit der eingeführten Begriffe *Schiefe* und *Dispersion* - mit den Varianten *Familiendispersion* und *Skalendispersion* - eine bessere Trennung verschiedener Verteilungen zu erwarten ist. Um diese Kennmaße als alternative Darstellung nutzen zu können, muß jedoch der Nachweis einer eindeutigen Abbildung zwischen den verschiedenen Parametrisierungen erbracht werden.

Als Vorgriff und zur Motivation sei die folgende Abbildung mit dem entsprechenden Konturdiagramm mit der Familien- und Skalendispersion für die Normalverteilung angeführt:

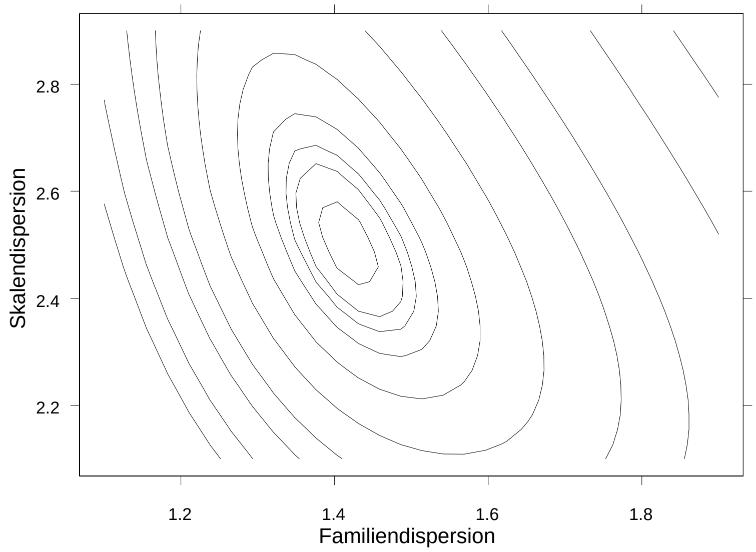

Abbildung 6.2: Konturdiagramm für $d_{fQ}(fQ_{NV}, \cdot)$ in (μ_D^F, μ_D^S)-Parametrisierung

Die ellipsoide Struktur der Höhenlinienkurven mit nahezu gleichen Achsenlängen deutet auf eine geringe Korrelation zwischen μ_D^F und μ_D^S hin und rechtfertigt somit diesen Ansatz.
Aufgrund der Einschränkung der Kennmaße auf unimodale Verteilungen wird im folgenden nur die Menge $\mathcal{FQ}_U$ betrachtet. Da weiterhin die fQ-Funktion unabhängig vom Lokationsparameter μ ist, erfolgt die Umparametrisierung nur für die Parameter (σ, α, β) der fQ-Verteilungen.

Satz 6.3
Sei $F = \mathcal{F}(\mu, \sigma, \alpha, \beta) \in \mathcal{FQ}_U$. Dann existiert eine bijektive Abbildung

$$\tau : \mathbb{R}_+^3 \to \mathbb{R}_+^2 \times [-1, 1], \tag{6.3}$$

mit

$$\tau(\sigma, \alpha, \beta) = \left(\mu_D(F), \mu_D^F(F), \mu_S(F)\right). \tag{6.4}$$

Beweis:

Im folgenden konstruktiven Beweis werden schrittweise die fQ-Darstellungen der obigen Kennmaße gemäß Satz 5.33 nach den Parametern (σ, α, β) aufgelöst. Zunächst folgt aus der Darstellung für das Schiefemaß:

$$\mu_S(F) = \frac{\beta - \alpha}{\alpha + \beta} \Leftrightarrow \beta = \alpha \frac{1 + \mu_S(F)}{1 - \mu_S(F)} =: \alpha c.$$

Damit folgt für die Familiendispersion

$$\begin{aligned} \mu_D^F(F) &= \frac{\alpha^\alpha \beta^\beta}{(\alpha+\beta)^{\alpha+\beta} B(\alpha+1, \beta+1)} \\ \Leftrightarrow \mu_D^F(F) &= \frac{\alpha^\alpha (\alpha c)^{\alpha c}}{(\alpha + \alpha c)^{\alpha + \alpha c} B(\alpha+1, \alpha c+1)}. \end{aligned} \tag{6.5}$$

Da

$$\frac{\alpha^\alpha (\alpha c)^{\alpha c}}{(\alpha + \alpha c)^{\alpha + \alpha c}} = \frac{\alpha^{\alpha + \alpha c} c^{\alpha c}}{(\alpha + \alpha c)^{\alpha + \alpha c}} = \left(\left(\frac{1}{1+c} \right)^{1+c} c^c \right)^\alpha$$

stetig und monoton steigend in α ist und

$$\begin{aligned} B(\alpha+1, \alpha c+1) &= \int_0^1 x^\alpha (1-x)^{\alpha c} \, dx \\ &= \int_0^1 (x(1-x)^c)^\alpha \, dx \end{aligned}$$

ebenfalls stetig und monoton fallend in α ist, folgt insgesamt, daß der Ausdruck auf der rechten Seite von (6.5) stetig und monoton steigend in α ist. Dadurch ist die Existenz einer eindeutigen Lösung der Gleichung (6.5) für α gesichert.
Mit diesen somit bestimmten Werten für α und β kann aus der Beziehung

$$\mu_D(F) = \frac{1}{\sigma \, B(\alpha+1, \beta+1)}$$

schließlich der Parameter σ eindeutig bestimmt werden. □

Bemerkung 6.4

In der Praxis kann die Lösung von (6.5) durch ein Standardverfahren zur numerischen Bestimmung von Nullstellen bestimmt werden. Dazu bietet sich beispielsweise die Sekantenmethode an, da bei dieser Methode keine Ableitungen benötigt werden (siehe Björck and Dahlquist [1972, S. 168ff]).

6.3 Erwartungswerte von Ordnungsstatistiken

Als erste Anwendung der fQ-Darstellung sei hier die Bestimmung der Erwartungswerte von Ordnungsstatistiken für unimodale Verteilungen betrachtet. Diese Aufgabenstellung ist auch im Zeitalter der schnellen Computer ein praktisches Problem, da die dazu benötigten numerischen Verfahren zur exakten Berechnung von

$$E(X_{(i)}) = n\binom{n-1}{i-1}\int_0^1 Q(p)p^{i-1}(1-p)^{n-i}\,dp \tag{6.6}$$

anspruchsvoll und nicht in der gängigen Standardsoftware zu finden sind. Tabellenwerke haben den Nachteil, daß sie sehr umfangreich sein müssen, da für jeden Stichprobenumfang n eine eigene Tabelle notwendig ist. Es gibt aus diesem Grunde viele Verfahren, mit denen für spezielle Verteilungen gute Approximationen berechnet werden können (siehe etwa Johnson, Kotz und Balakrishnan [1994, 1995]).

In diesem Abschnitt soll die Menge dieser Approximationsformeln durch ein Verfahren, welches für beliebige unimodale Verteilungen die fQ-Approximation benutzt, erweitert werden, um eine allgemeine Näherungsformel für die Erwartungswerte zu konstruieren. Dabei wird das von Blom (1958) vorgeschlagene Verfahren der α,β-*Korrektur* benutzt, das von ihm in einem allgemeineren Zusammenhang untersucht wurde (siehe auch Kapitel 2.3.1). Um Verwechslungen mit den Parametern der fQ-Verteilungen zu vermeiden, werden abweichend von der Notation von Blom im folgenden die Bezeichnungen a_i und b_i verwendet. Blom geht bei seinem Ansatz von folgender Darstellung aus:

$$\begin{aligned} E(X_{(i)}) &= Q(\pi_i) + R(\pi_i, a_i, b_i), && (6.7)\\ \text{mit} \quad \pi_i &:= \frac{i-a_i}{n-a_i-b_i+1}, && (6.8)\end{aligned}$$

wobei Q die Quantilfunktion der zugrundeliegenden Verteilung und R einen Restterm bezeichnet, der ebenso wie π_i von zunächst unbekannten Werten a_i, b_i abhängt.

Damit wird die Formel von David und Johnson (1954)

$$\begin{aligned} E(X_{(i)}) &\approx Q(p_i) + \frac{p_i(1-p_i)}{2(n+2)} Q''(p_i), \\ p_i &= \frac{i}{n+1}, \end{aligned}$$

die auf einer Taylorentwicklung der Funktion $Q(U_{(i)})$ beruht, dadurch modifiziert, daß die zweite Ableitung der Quantilfunktion durch eine geeignete Wahl von Korrekturwerten a_i und b_i im Argument von Q ersetzt wird. Für spezielle Verteilungen wie zum Beispiel die Normalverteilung oder die Extremwertverteilung erhält Blom aufgrund von theroretischen und empirischen Überlegungen Darstellungen, in denen die Korrekturwerte a_i und b_i unabhängig von n gewählt werden können. Im Falle des fQ-Systems gilt folgende Aussage:

Satz 6.5
Für den Erwartungswert der i-ten Ordnungsstatistik einer fQ-verteilten Zufallsvariable $X \sim \mathcal{F}(\mu, \sigma, \alpha, \beta)$ mit $i \in [\alpha/2, n+1-\beta/2]$ gilt:

$$\begin{aligned} E(X_{(i)}) &= Q(\pi_i) + O(n^{-3/2}), && (6.9) \\ \textit{mit} \quad \pi_i &:= \frac{i - \frac{\alpha}{2}}{n - \frac{\alpha+\beta}{2} + 1}. && (6.10) \end{aligned}$$

Beweis:
Nach Blom [1958, S. 64] kann der allgemeine Ansatz (6.7) durch die Definition

$$\tilde{R}(\pi_i, a_i, b_i) := \frac{1}{2}\pi_i(1-\pi_i)Q''(\pi_i) + (a_i(1-\pi_i) - b_i\pi_i)Q'(\pi_i), \qquad (6.11)$$

präzisiert werden zu:

$$E(X_{(i)}) = Q(\pi_i) + \frac{1}{n}\tilde{R}(\pi_i, a_i, b_i) + O(n^{-3/2}). \qquad (6.12)$$

Da mit den Ableitungsformeln für die Quantilfunktion aus Lemma 3.5

$$\frac{Q''(p)}{Q'(p)} = \frac{fQ'(p)}{fQ(p)} = \frac{\beta}{1-p} - \frac{\alpha}{p}, \quad p \in (0,1)$$

gilt, folgt durch Einsetzen dieser Beziehung in (6.11)

$$\begin{aligned} \tilde{R}(\pi_i, a_i, b_i) &= \frac{1}{2}\pi_i(1-\pi_i)Q'(\pi_i)\left(\frac{\beta}{1-\pi_i} - \frac{\alpha}{\pi_i}\right) + (a_i(1-\pi_i) - b_i\pi_i)Q'(\pi_i) \\ &= Q'(\pi_i)\left((1-\pi_i)(a_i - \frac{\alpha}{2}) + \pi_i(\frac{\beta}{2} - b_i)\right). \end{aligned}$$

Somit gilt für die Koeffizienten $a_i := \alpha/2$ und $b_i := \beta/2$

$$\tilde{R}(\pi_i, a_i, b_i) = 0, \quad \forall i, n.$$

Die Einschränkungen für den Index i in der Voraussetzung folgen aus der notwendigen Bedingung $\pi_i \in [0, 1]$. □

Die Bedeutung dieses Satzes für praktische Anwendungen liegt darin, daß die Koeffizienten a_i und b_i unabhängig von i und n direkt aus den Parametern der fQ-Approximation abgeleitet werden können.
Durch eine analoge Vorgehensweise wie in Satz 6.5 ergibt sich für die Kovarianzen der Ordnungsstatistiken folgende Darstellung:

Lemma 6.6
Für fQ-verteilte Zufallsvariablen $X_i, X_j \sim \mathcal{F}(\mu, \sigma, \alpha, \beta)$ gilt:

$$Cov(X_{(i)}, X_{(j)}) = \frac{\pi_i(1-\pi_j)}{n - \frac{\alpha+\beta}{2} + 2} Q'(\pi_i)Q'(\pi_j) + O(n^{-2}), \tag{6.13}$$

für $\alpha/2 < i < j < n + 1 - \beta/2$ und π_i wie in Satz 6.5.

Beweis:
Mit der Darstellung in Blom [1958, S. 62ff] verläuft der Beweis analog zu dem von Satz 6.5. □

Durch Anwendung der fQ-Approximation einer bekannten unimodalen Verteilung F gemäß der Vorgehensweise aus Abschnitt 6 können diese Formeln nun genutzt werden, um Näherungsformeln für die ersten Momente der Ordnungsstatistiken von F zu bestimmen. Die Qualität dieser Approximationen wird im Vergleich mit anderen Näherungsformeln und dem exakten Ergebnis in Kapitel 10 näher untersucht.
Daneben bietet dieses Verfahren die Möglichkeit, mittels Anpassung einer fQ-Verteilung an empirische Daten (siehe Kapitel 7) Schätzungen für die Momente der Ordnungsstatistiken für eine *unbekannte* zugrundeliegende Verteilung zu erhalten.

Kapitel 7

Dichteschätzung

7.1 Übersicht

Da die Definition der fQ-Verteilung keine explizite Darstellung der Dichtefunktion und der Momente erlaubt, erweisen sich die Maximum-Likelihood-Methode und die Momentenmethode zur Schätzung der Parameter als nicht praktikabel. Aus diesem Grund kommt hier die KQ-Methode zur Anwendung, wobei die Anpassung im Hinblick auf die Quantilfunktion erfolgt.

Bei der methodischen Umsetzung dieses Ansatzes wäre es grundsätzlich möglich, das entstehende KQ-Minimierungsproblem innerhalb eines kompletten, entsprechend der Definiton der fQ-Verteilung vierparametrischen Modells zu formulieren und mittels der bekannten Verfahren zur Lösung nichtlinearer Regressionsprobleme zu behandeln. Diese allgemeine Methode würde aber der speziellen Situation des vorliegenden Problems nicht gerecht werden, da sich das Modell in einen linearen und einen nichtlinearen Anteil zerlegen läßt, und für beide Teile jeweils optimale Verfahren existieren.

Neben diesem eher verfahrenstechnischen Grund für die Aufteilung des Minimierungsproblems zeigt es sich, daß damit zugleich die bekannte Theorie der optimalen linearen Schätzer, die auf Ordnungsstatistiken beruhen (L-Schätzer), erweitert wird. Demzufolge beginnt die Darstellung des eigentlichen Schätzverfahrens für die fQ-Parameter mit diesem Teil, dem eine eigenständige Bedeutung zukommt.

7.2 L-Schätzer (BLUE)

Ausgangsbasis für die Untersuchungen in diesem Abschnitt ist ein parametrisches Lokations- und Skalenmodell der Form

$$\mathcal{F} = \{F_{\mu,\sigma}; F_{\mu,\sigma}(x) = F_0(\mu + \sigma x), \sigma \in \mathbb{R}^+, \mu \in \mathbb{R}\} \tag{7.1}$$

mit einer stetigen unimodalen Verteilung $F_0 \in \mathcal{F}_U$. Dann ist zu einer Stichprobe unabhängiger, identisch verteilter Zufallsvariablen $X_1, .., X_n \sim F_{\mu,\sigma} \in \mathcal{F}$ ein optimaler Schätzer für die unbekannten Parameter μ und σ gesucht.

Zur Abkürzung wird diese Situation im folgenden mit dem Begriff *Schätzung im LS-Modell mit Grundverteilung* F_0 bezeichnet. Man beachte, daß in diesem Modell die Grundverteilung F_0 als fest und bekannt vorausgesetzt wird.

Zur Lösung dieser Aufgabe leitete Loyd (1952) den folgenden KQ-Schätzer her, der auf den Ordnungsstatistiken der Grundverteilung F_0 beruht.
Zunächst gilt für die standardisierte Variable

$$Y := \frac{X - \mu}{\sigma} \sim F_0$$

aufgrund der Monotonie der Transformation die Beziehung

$$Y_{(i)} = \frac{X_{(i)} - \mu}{\sigma}, \quad i = 1, .., n.$$

Seien nun $\mathbf{a} := E(\mathbf{Y})$ der Vektor der Erwartungswerte und $\mathbf{B} := Cov(\mathbf{Y})$ die Kovarianzmatrix von $\mathbf{Y} = (Y_{(1)}, .., Y_{(n)})$, also

$$a_i = E(Y_{(i)}), \tag{7.2}$$

$$b_{i,j} = Cov(Y_{(i)}, Y_{(j)}). \tag{7.3}$$

Dann folgt

$$\begin{aligned} E(X_{(i)}) &= \mu + \sigma a_i, \\ Cov(X_{(i)}, X_{(j)}) &= \sigma^2 b_{i,j}. \end{aligned}$$

Die Parameter μ und σ können demnach mit Hilfe eines verallgemeinerten linearen Modells der Form

$$\mathbf{X} = \mathbf{A}\boldsymbol{\gamma} + \mathbf{e} \tag{7.4}$$

geschätzt werden. Dabei sind

$$\begin{aligned}
\mathbf{X} &:= (X_{(1)}, .., X_{(n)})', \\
\mathbf{A} &:= (\mathbf{1_n}, \mathbf{a}), \\
\boldsymbol{\gamma} &:= (\mu, \sigma)', \\
\mathbf{e} &\sim (0, \sigma^2\mathbf{B}).
\end{aligned}$$

Der beste lineare, unverzerrte Schätzer (BLUE) $\hat{\gamma}$ ist nach dem Satz von Gauß-Markov in diesem Modell gegeben durch den Aitken-Schätzer

$$\hat{\boldsymbol{\gamma}} = (\mathbf{A}'\mathbf{B}^{-1}\mathbf{A})^{-1}\mathbf{A}'\mathbf{B}^{-1}\mathbf{X}. \tag{7.5}$$

Mit den weiteren Bezeichnungen

$$\begin{aligned}
\boldsymbol{\Delta} &:= \frac{\mathbf{B}^{-1}(\mathbf{1_n}\mathbf{a}' - \mathbf{a}\mathbf{1_n'})\mathbf{B}^{-1}}{D}, && (7.6) \\
D &:= (\mathbf{a}'\mathbf{B}^{-1}\mathbf{a})(\mathbf{1_n'}\mathbf{B}^{-1}\mathbf{1_n}) - (\mathbf{a}'\mathbf{B}^{-1}\mathbf{1_n})^2 && (7.7)
\end{aligned}$$

kann $\hat{\gamma}$ dann in seine Komponenten $\hat{\mu}$ und $\hat{\sigma}$ zerlegt werden und man erhält schließlich folgende Darstellung:

Satz 7.1
Die BLU-Schätzer im LS-Modell mit Grundverteilung F_0 sind gegeben durch

$$\begin{aligned}
\hat{\mu}_{opt} &:= -\mathbf{a}'\boldsymbol{\Delta}\mathbf{X}, && (7.8) \\
\hat{\sigma}_{opt} &:= \mathbf{1_n}\boldsymbol{\Delta}\mathbf{X}. && (7.9)
\end{aligned}$$

Beweis:
Die Details der Matrixumformungen zur Zerlegung von $\hat{\gamma}$ sind in Balakrishnan und Cohen [1991, S. 80ff] zu finden. $\square$

Bemerkung 7.2

Die Bezeichnung L-Schätzer *für $\hat{\mu}_{opt}$ und $\hat{\sigma}_{opt}$ wird motiviert durch die alternative Darstellung der Schätzer in der Form*

$$\hat{\mu}_{opt} = \sum_{i=1}^{n} c_i\, X_{(i)} \quad \text{und} \quad \hat{\sigma}_{opt} = \sum_{i=1}^{n} d_i\, X_{(i)}, \tag{7.10}$$

wobei die Koeffizienten aus den Beziehungen in Satz 7.1 gewonnen werden können. In dieser Form sind die Schätzer offensichtlich Linearkombinationen *der Ordnungsstatistiken $X_{(i)}$. Da die resultierenden Formeln für die Größen c_i und d_i im allgemeinen Fall sehr komplex sind, wird hier zunächst die kompaktere Schreibweise benutzt.*

Für spätere Untersuchungen ist der folgende Satz wichtig, mit dem Abschätzungen für die Effizienz der Schätzer gemacht werden können:

Satz 7.3

Für die BLU-Schätzer im LS-Modell gilt:

$$\begin{aligned} Var(\hat{\mu}_{opt}) &= \sigma^2 \frac{\mathbf{a}'\mathbf{B}^{-1}\mathbf{a}}{D}, && (7.11)\\ Var(\hat{\sigma}_{opt}) &= \sigma^2 \frac{\mathbf{1_n}'\mathbf{B}^{-1}\mathbf{1_n}}{D}. && (7.12) \end{aligned}$$

Beweis:

Die Aussage folgt durch Einsetzen der Beziehungen (7.6) und (7.7) in die Varianzformeln für das verallgemeinerte lineare Modell 7.4 (siehe Balakrishnan und Cohen [1991, S. 81]). □

Im Falle einer symmetrischen Grundverteilung ergeben sich wesentliche Vereinfachungen, die folgendermaßen zusammengefaßt werden können:

Lemma 7.4

Die BLU-Schätzer für ein LS-Modell mit symmetrischer Grundverteilung F_0 sind gegeben durch:

$$\begin{aligned} \hat{\mu}_{opt} &= \frac{\mathbf{1_n}'\mathbf{B}^{-1}}{\mathbf{1_n}'\mathbf{B}^{-1}\mathbf{1_n}}\mathbf{X}, && (7.13)\\ \hat{\sigma}_{opt} &= \frac{\mathbf{a}'\mathbf{B}^{-1}}{\mathbf{a}'\mathbf{B}^{-1}\mathbf{a}}\mathbf{X}. && (7.14) \end{aligned}$$

Beweis:
Aufgrund der Symmetriebedingung gilt

$$B_{i,j} = B_{n-j+1,n-i+1}, \quad i,j = 1,\dots,n,$$

woraus sich durch weitere Umformungen (siehe Balakrishnan und Cohen [1991, S. 82]) die Beziehung

$$\mathbf{a}'\mathbf{B}^{-1}\mathbf{1_n} = 0$$

ergibt. Daraus folgen unmittelbar die vereinfachten Formeln durch Einsetzen dieser Aussage in Satz 7.1. □

Bemerkung 7.5
Da speziell für die Normalverteilung zusätzlich

$$\mathbf{1_n'}\mathbf{B}^{-1} = \mathbf{1_n'}$$

gilt (siehe Balakrishnan und Cohen [1991, S. 59], folgt

$$\begin{aligned} \hat{\mu}_{opt} &= \bar{\mathbf{X}}, \\ Var(\hat{\mu}_{opt}) &= \frac{\sigma^2}{n}, \end{aligned}$$

im Einklang mit den bekannten Resultaten für die Normalverteilung.

Für den praktischen Gebrauch haben die Schätzer $\hat{\mu}_{opt}$ und $\hat{\sigma}_{opt}$ im allgemeinen zwei wesentliche Nachteile:

- Die Berechnung des Erwartungswertvektors **a** und der Kovarianzmatrix **B** ist nur für sehr wenige Verteilungen explizit möglich. In den übrigen Fällen sind nichttriviale numerische Probleme zu lösen (siehe z.B. Arnold, Balakrishnan und Nagaraja [1992]).

- Die Invertierung der $n \times n$ Matrix **B** ist mit erheblichem Rechenaufwand verbunden. Zudem sind die Größen **a** und **B** vom Stichprobenumfang n abhängig und müssen also für verschiedene Datensätze erneut berechnet oder tabelliert werden.

Im folgenden Abschnitt wird beschrieben, mit welchen Modifikationen diese Schätzer dennoch nutzbar gemacht werden können.

7.3 Approximative L-Schätzer (ABLUE)

Zur Lösung der technischen Hindernisse bei der Berechnung der L-Schätzer wurden Ansätze verfolgt, in denen zum einen die exakten Werte für $\mathbf{a}$ und $\mathbf{B}$ durch Approximationen ersetzt werden und in denen zum anderen nur Teilmengen $\tilde{\mathbf{X}} \subset \mathbf{X}$ der Ausgangsdaten betrachtet werden. Da die letztere Methode gleichzeitig auch die Betrachtung von zensierten Daten und eine Robustifizierung der Schätzung ermöglicht, kommt ihr eine besondere Bedeutung zu.
Eine sehr einfache Variante der L-Schätzung basiert auf der Tatsache, daß nach Bemerkung (7.5) für die Normalverteilung gilt:

$$\mathbf{1}_\mathbf{n}'\mathbf{B}^{-1} = \mathbf{1}_\mathbf{n}'.$$

Gupta (1952) schlägt vor, die Invertierung der Kovarianzmatrix $\mathbf{B}$ durch die drastische Annahme $\mathbf{B} = \mathbf{I}$ zu umgehen, wodurch sich die Darstellungen der Schätzer erheblich vereinfachen:

Satz 7.6
Die approximativen L-Schätzer im LS-Modell sind nach Gupta gegeben durch die Formeln

$$\hat{\mu}_G = \sum_{i=1}^{n} c_i^G X_{(i)} \quad \textit{und} \quad \hat{\sigma}_G = \sum_{i=1}^{n} d_i^G X_{(i)}, \tag{7.15}$$

mit

$$c_i^G := \frac{1}{n} - \frac{\bar{a}(a_i - \bar{a})}{\sum_{j=1}^{n}(a_j - \bar{a})^2} \quad \textit{und} \quad d_i^G := \frac{a_i - \bar{a}}{\sum_{j=1}^{n}(a_j - \bar{a})^2}. \tag{7.16}$$

Beweis:
Die Herleitung der vereinfachten Formeln sind in Balakrishnan und Cohen [1991, S. 94ff] zu finden. □

Für kleine Stichproben verfügt dieser Schätzer im Falle der Normalverteilung über eine hohe Effizienz (Sarhan und Greenberg [1962]), wie auch die Simulationsstudie in Kapitel 11 zeigt. Zu beachten ist bei diesem Schätzer, daß zu seiner Berechnung die Kenntnis der Werte $\mathbf{a} = E(\mathbf{Y})$ notwendig ist.

Wesentlich genauer als die Methode von Gupta ist das allgemeine Verfahren von Blom (1958), bei dem eine Approximation $\hat{\mathbf{B}}$ der Kovarianzmatrix benutzt wird. Zur Darstellung dieses ABLU-Schätzers werden folgende Bezeichnungen benötigt: Seien

$$\tilde{\mathbf{X}} := (X_{(n_1)}, \ldots, X_{(n_k)}), \quad 1 \leq n_1 < \ldots < n_k \leq n$$

und

$$\hat{\mathbf{a}} := \hat{E}(\tilde{\mathbf{Y}}), \quad \hat{\mathbf{B}} := \widehat{Cov}(\tilde{\mathbf{Y}})$$

Schätzungen für die Erwartungswerte der Ordnungsstatistiken und die Kovarianz der standardisierten Variable $\tilde{\mathbf{Y}} = (\tilde{\mathbf{X}} - \mu)/\sigma$ im LS-Modell.
Die Elemente von $\hat{\mathbf{B}}$ seien dabei durch folgende Beziehung gegeben:

$$\hat{\mathbf{B}}_{(i,j)} =: \hat{b}_{i,j} = \frac{p_i(1-p_j)}{\hat{n} f_i f_j}, \quad 1 \leq i \leq j \leq k \tag{7.17}$$

mit $0 < p_1 < \ldots < p_k < 1$, $f_i > 0$ und $\hat{n} > 0$.
Diese Form der $\hat{b}_{i,j}$ wird durch die Näherungsformel von David und Johnson (1954) motiviert:

$$Cov(Y_{(i)}, Y_{(j)}) \approx \frac{p_i(1-p_j)}{(n+2) fQ(p_i) fQ(p_j)} \quad i < j. \tag{7.18}$$

Mit der allgemeinen Schreibweise (7.17) ist es möglich, die Herleitung verschiedener Varianten der Schätzer geschlossen darzustellen.

Werden nun die Näherungen $\tilde{\mathbf{X}}, \hat{\mathbf{a}}$ und $\hat{\mathbf{B}}$ anstelle der exakten Größen $\mathbf{X}, \mathbf{a}$ und $\mathbf{B}$ in der Darstellung des optimalen L-Schätzers in Satz 7.1 eingesetzt, so erhält man vollkommen analog die approximativen Schätzer

$$\hat{\mu}_{appr} := -\hat{\mathbf{a}}' \hat{\mathbf{\Delta}} \tilde{\mathbf{X}}, \tag{7.19}$$

$$\hat{\sigma}_{appr} := \mathbf{1_k} \hat{\mathbf{\Delta}} \tilde{\mathbf{X}}, \tag{7.20}$$

mit

$$\hat{\mathbf{\Delta}} = \frac{\hat{\mathbf{B}}^{-1}(\mathbf{1_k}\hat{\mathbf{a}}' - \hat{\mathbf{a}}\mathbf{1}'_\mathbf{k})\hat{\mathbf{B}}^{-1}}{(\hat{\mathbf{a}}'\hat{\mathbf{B}}^{-1}\hat{\mathbf{a}})(\mathbf{1}'_\mathbf{k}\hat{\mathbf{B}}^{-1}\mathbf{1_k}) - (\hat{\mathbf{a}}'\hat{\mathbf{B}}^{-1}\mathbf{1_k})^2}. \tag{7.21}$$

Der Vorteil in der Darstellung der Matrix $\hat{\mathbf{B}}$ liegt darin, daß ihre Inverse sich direkt berechnen läßt. Mit der Bezeichung

$$\delta_i^p := p_i - p_{i-1}, \quad i = 1, \dots, k+1,$$

und $p_0 = 0$, $p_{k+1} = 1$ erhält man folgende Darstellung:

$$\hat{\mathbf{B}}^{-1}_{(i,j)} = \begin{cases} 2\hat{n} f_i^2 \left(\frac{1}{\delta_{i+1}^p} - \frac{1}{\delta_i^p} \right) & , \quad j = i \\ -\frac{\hat{n} f_i f_{i-1}}{\delta_i^p} & , \quad j = i+1 \text{ oder } j = i-1 \\ 0 & , \quad \text{sonst} \end{cases} .$$

Durch die einfache, tridiagonale Struktur von $\hat{\mathbf{B}}$ können nun alle Terme in der Formel der Matrix $\hat{\boldsymbol{\Delta}}$ (7.21) mit elementaren Ausdrücken dargestellt werden. Dazu seien zunächst folgende Abkürzungen eingeführt:

$$\begin{aligned} \delta_i^f &:= f_i - f_{i-1}, \quad i = 1, \dots, k+1, \\ \delta_i^{af} &:= \hat{a}_i f_i - \hat{a}_{i-1} f_{i-1}, \quad i = 1, \dots, k+1, \end{aligned}$$

mit $f_0 = f_{k+1} = \hat{a}_0 f_0 = \hat{a}_{k+1} f_{k+1} = 0$.
Damit gelten folgende Beziehungen (siehe auch Balakrishnan und Cohen [1991, S. 269]):

$$K_1 := \mathbf{1}_\mathbf{k}' \hat{\mathbf{B}}^{-1} \mathbf{1}_\mathbf{k} = \hat{n} \sum_{i=1}^{k+1} \frac{(\delta_i^f)^2}{\delta_i^p}, \tag{7.22}$$

$$K_2 := \hat{\mathbf{a}}' \hat{\mathbf{B}}^{-1} \hat{\mathbf{a}} = \hat{n} \sum_{i=1}^{k+1} \frac{(\delta_i^{af})^2}{\delta_i^p}, \tag{7.23}$$

$$K_3 := \hat{\mathbf{a}}' \hat{\mathbf{B}}^{-1} \mathbf{1}_\mathbf{k} = \hat{n} \sum_{i=1}^{k+1} \frac{\delta_i^{af} \delta_i^f}{\delta_i^p}. \tag{7.24}$$

Mit den weiteren Abkürzungen

$$\delta_i^{fp} := \frac{\delta_i^f}{\delta_i^p} - \frac{\delta_{i+1}^f}{\delta_{i+1}^p}, \tag{7.25}$$

$$\delta_i^{afp} := \frac{\delta_i^{af}}{\delta_i^p} - \frac{\delta_{i+1}^{af}}{\delta_{i+1}^p} \tag{7.26}$$

können dann die Schätzer schließlich in folgender Gestalt formuliert werden:

Satz 7.7
Die approximativen L-Schätzer (ABLUE) im LS-Modell bei fester Grundverteilung F_0 sind mit den Näherungen $\hat{\mathbf{a}} = \hat{E}(\tilde{\mathbf{Y}})$ und $\hat{\mathbf{B}} = \widehat{Cov}(\tilde{\mathbf{Y}})$ gemäß (7.17) gegeben durch die Formeln

$$\hat{\mu}_{appr} = \sum_{i=1}^{k} c_i^* x_{(n_i)}, \quad \hat{\sigma}_{appr} = \sum_{i=1}^{k} d_i^* x_{(n_i)}, \tag{7.27}$$

mit

$$c_i^* \;=\; \hat{n} f_i \left(\frac{K_2}{K_4} \delta_i^{fp} - \frac{K_3}{K_4} \delta_i^{afp} \right), \tag{7.28}$$

$$d_i^* \;=\; \hat{n} f_i \left(\frac{K_3}{K_4} \delta_i^{fp} - \frac{K_1}{K_4} \delta_i^{afp} \right). \tag{7.29}$$

Für die Varianzen dieser Schätzer gilt:

$$Var(\hat{\mu}_{appr}) = \sigma^2 \frac{K_1}{K_4} \quad \text{und} \quad Var(\hat{\sigma}_{appr}) = \sigma^2 \frac{K_2}{K_4}. \tag{7.30}$$

Beweis:
Die Darstellungen für die Varianzen können mit obigen Bezeichnungen direkt aus Satz 7.3 übernommen werden. □

Aufgrund der allgemeinen Darstellungsweise der Schätzer müssen bei der folgenden Beschreibung verschiedener Varianten lediglich die Formeln für die Approximationen $\hat{\mathbf{a}}$ und $\hat{\mathbf{B}}$ angegeben werden.

Die Schätzer von Blom

Bei der ursprünglichen Wahl von Blom (1958), auf den das Konstruktionsprinzip der ABLU-Schätzung zurückgeht, werden die bekannten Formeln von David und Johnson (1954) herangezogen, d.h.

$$\hat{a}_i^{B_1} \;:=\; Q_i + \frac{p_i(1-p_i)}{2(n+2)} Q_i'', \tag{7.31}$$

$$\hat{b}_{i,j}^{B_1} \;:=\; \frac{p_i(1-p_j)}{(n+2)} Q_i' Q_j', \quad i < j, \tag{7.32}$$

mit $p_i := i/(n+1)$ und $Q_i := Q(p_i)$.

Die gemäß Satz 7.7 daraus entstehenden Schätzer seien mit $\hat{\mu}^B_{appr}$ und $\hat{\sigma}^B_{appr}$ bezeichnet und werden im Rahmen der Vergleichsstudie in Kapitel 11 untersucht.

Da in der Praxis für viele Verteilungen die Darstellung der ersten beiden Ableitungen der Quantilfunktion kompliziert ist, schlägt Blom im zweiten Schritt vor, alternativ zu obigen Formeln die Näherungen zu wählen, die er mit der Methode der α-β-Korrektur gewonnen hat. Diese Formeln wurden bereits in Kapitel 6.3 erwähnt und haben die Gestalt

$$\begin{aligned} \hat{a}_i^{B_2} &= Q(\pi_{i,i}), && (7.33)\\ \hat{b}_{i,j}^{B_2} &= \frac{\pi_{i,j}(1-\pi_{j,i})}{n-c_{i,j}-d_{i,j}+2} Q'(\pi_{i,j})Q'(\pi_{j,i}), && (7.34)\\ \pi_{i,j} &:= \frac{i-c_{i,j}}{n-c_{i,j}-d_{i,j}+1}. && (7.35) \end{aligned}$$

Das Problem bei dieser Darstellung liegt darin, daß die Wahl der optimalen Konstanten $c_{i,j}$ und $d_{i,j}$ stark von der zugrundeliegenden Verteilung F_0 abhängt und es im allgemeinen Fall keine kanonische Lösung gibt. Deswegen kann diese Variante der ABLU-Schätzung nur selten angewendet werden.

Der fQ-ABLU-Schätzer

Die praktische Schwierigkeit bei den Schätzern von Blom ist die Bestimmung der Konstanten $c_{i,j}$ und $c_{i,j}$ zur Approximation der Momente der Ordnungsstatistiken für beliebige Verteilungen. Daher liegt es nahe, die Grundverteilung F_0 des LS-Modells durch deren fQ-Approximation $\hat{F}_0 = \mathcal{F}(\mu_0, \sigma_0, \alpha_0, \beta_0)$ gemäß Definition 6.1 zu ersetzen, um die Näherungsformeln aus Kapitel 6.3 heranziehen zu können.

Man beachte dabei, daß der Lokationsparameter μ_0 von $\hat{F}_0$ aufgrund der Standardisierung der fixen Grundverteilung F_0 des LS-Modells als 0 angenommen werden kann. Dagegen muß der Gestaltparameter σ_0 der fQ-Darstellung von $\hat{F}_0$ von dem unbekannten Skalenparameter σ des LS-Modells unterschieden werden. Da die fQ-Darstellung auf unimodale Verteilungen beschränkt ist, muß muß bei dieser Methode grundsätzlich die Einschränkung $F_0 \in \mathcal{F}_U$ gemacht werden.

Seien also im folgenden $\hat{Q}_0$ und $\widehat{fQ}_0$ die Quantil- bzw. fQ-Funktion der fQ-Verteilung mit Parametern $(0, \sigma_0, \alpha_0, \beta_0)$. Nach Satz 6.5 und Satz 6.6 können dann folgende Approximationen benutzt werden:

$$\hat{a}_i^{fQ} := \hat{Q}_0(\pi_i), \tag{7.36}$$

$$\hat{b}_{i,j}^{fQ} := \frac{\pi_i(1-\pi_j)}{n - \frac{\alpha+\beta}{2} + 2} \hat{Q}_0'(\pi_i)\hat{Q}_0'(\pi_j), \tag{7.37}$$

mit

$$\pi_i := \frac{i - \alpha/2}{n - (\alpha+\beta)/2 + 1} \quad , \alpha/2 < i < n + 1 - \beta/2. \tag{7.38}$$

Die notwendige Einschränkung der π_i durch die Parameterbedingungen hat zur Folge, daß nicht der komplette Datensatz bei der Schätzung berücksichtigt werden kann, sondern nur die $X_{(i)}$, deren Index die obige Bedingung erfüllt. Damit gelangt man zu folgendem Satz

Satz 7.8
Sei $F \in \mathcal{F}_U$ mit fQ-Approximation $\hat{F}_0 \in \mathcal{FQ}_U$. Dann werden die approximativen fQ-Schätzer *(fQ-ABLU-Schätzer) $\hat{\mu}_{fQ}$, $\hat{\sigma}_{fQ}$ im LS-Modell mit*

$$\tilde{\mathbf{X}} := (X_{(\lfloor \alpha/2 \rfloor + 1)}, \ldots, X_{(\lfloor n+1-\beta/2 \rfloor)})$$

und $\hat{a}_i^{fQ}$, $\hat{b}_{i,j}^{fQ}$ gemäß (7.36) und (7.37) aus Satz 7.7 berechnet.

Bemerkung 7.9
Man beachte, daß bei der Konstruktuktion des fQ-Schätzers ein wesentlicher Schritt in der Approximation der Grundverteilung F_0 durch eine fQ-Verteilung $\hat{F}_0$ besteht. Da diese Näherung für alle Verteilungen $F_0 \notin \mathcal{FQ}$ zu einem Approximationsfehler $d_{fQ}(F_0, \hat{F}_0)$ führt, folgt unmittelbar, daß die fQ-Schätzer für μ und σ im LS-Modell mit der exakten Grundverteilung F_0 verzerrt sein können. Dieser Tatbestand trifft in ähnlicher Form auf die Schätzer von Blom zu, die in der Literatur aus diesem Grund häufig auch als nearly unbiased nearly best estimates *Blom [1958, S. 125] bezeichnet werden.*

Die Rechtfertigung dafür, den fQ-ABLU-Schätzer dennoch einzusetzen, liegt darin, daß die Approximationsfehler für viele Verteilungen nicht gravierend sind und vor allem die grundsätzliche Idee der L-Schätzer erst durch Näherungsformeln nutzbar wird, da die technischen Probleme bei der Berechnung der exakten Werte für $\mathbf{a} := E(\mathbf{Y})$ *und* $\mathbf{B} := Cov(\mathbf{Y})$ *erheblich sind.*

Der eingeschränkte fQ-ABLU-Schätzer

Die Einschränkung des Datensatzes auf $\tilde{\mathbf{X}}$ in Definiton 7.8 begründet sich durch die notwendige Bedingung für die Terme π_i. Neben diesem technischen Grund für die Restriktion von $\mathbf{X}$ geben in der Praxis häufig die Daten selbst Anlaß, von der Betrachtung eines eingeschränkten Datensatzes Gebrauch zu machen, und zwar im Falle von *zensierten* Daten.
Sind etwa für einen Datensatz $\mathbf{X} = (X_1, \ldots, X_n)$ aus einem LS-Modell mit Grundverteilung F_0 nur die Daten $\tilde{\mathbf{X}} = (X_{r+1}, \ldots, X_{n-s})$ mit $0 \leq r, s$ bekannt (Typ-II-Zensierung), so können die fQ-ABLU-Schätzer gemäß Satz 7.7 ohne Modifikationen angewendet werden, sofern $r > \alpha/2$ und $s < \beta/2 - 1$ gilt.
Diese Technik, die in dieser Form von Benett (1952) hergeleitet wurde, wird in Kapitel 11 genutzt, um den fQ-ABLU-Schätzer nur für den Teil der Datenmenge zu definieren, für den die Näherungsformeln $\hat{\mathbf{a}}$ zu akzeptablen Approximationen führt.

7.4 fQ-Dichteschätzung

Mit dem fQ-ABLU-Schätzer aus dem letzten Abschnitt wurde die Schätzung des Lokations- und Skalenparameters für eine *feste* Grundverteilung F_0 behandelt. Um dem übergeordneten Ziel, der Dichteschätzung von unimodalen Verteilungen, näher zu kommen, wird nun die Grundverteilung F_0 ebenfalls als variabel angenommen. Dabei werden im Gegensatz zur Situation beim fQ-ABLU-Schätzer nur noch fQ-Verteilungen betrachtet, die in standardisierter Form lediglich von den Gestaltparametern α und β abhängen. Das Problem der Schätzung aller fQ-Parameter wird somit in einen linearen Teil, der Schätzung von μ und σ

durch den fQ-ABLU-Schätzer, und in einen nichtlinearen Teil, der Schätzung der Gestaltparameter α und β, die im folgenden beschrieben wird, aufgeteilt.
Seien dazu in einem LS-Modell mit standardisierter fQ-Verteilung

$$F(\alpha, \beta) := \mathcal{F}(0, 1, \alpha, \beta)$$

die gemäß Abschnitt 7.3 bestimmten fQ-ABLU-Schätzer mit $\hat{\mu}^{(\alpha,\beta)}$ und $\hat{\sigma}^{(\alpha,\beta)}$ bezeichnet.
Um die Anpassungsgüte der Verteilung $\mathcal{F}(\hat{\mu}^{(\alpha,\beta)}, \hat{\sigma}^{(\alpha,\beta)}, \alpha, \beta)$ an die Daten zu messen und somit Paare (α, β) und $(\tilde{\alpha}, \tilde{\beta})$ vergleichen zu können, werden die Residuen aus der L-Schätzung

$$\hat{e}_j^{(\alpha,\beta)} := x_{(j)} - \hat{\mu}_j^{(\alpha,\beta)} - \hat{\sigma}_j^{(\alpha,\beta)} \hat{a}_j^{fQ}$$

herangezogen, wobei die Information über die Kovarianzstruktur des zugehörigen LS-Modells berücksichtigt wird, wenn die gewichtete Residuenquadratsumme

$$S^{(\alpha,\beta)} := \hat{\mathbf{e}}^{(\alpha,\beta)\prime} (\hat{\mathbf{B}}^{fQ})^{-1} \hat{\mathbf{e}}^{(\alpha,\beta)} \tag{7.39}$$

als Gütekriterium benutzt wird. Dabei wird die Matrix

$$\hat{\mathbf{B}}_{(i,j)}^{fQ} := b_{i,j}^{fQ}$$

im folgenden kurz $\hat{\mathbf{B}}$ genannt.

Wird die Größe S bei festem (α, β) als Funktion von $\mathbb{R}^2$ nach $\mathbb{R}$ aufgefaßt, so kann die Schätzung der fQ-Parameter formal durch die Definition

$$\begin{aligned} (\hat{\alpha}_{opt}, \hat{\beta}_{opt}) \quad &:= \quad \arg\min_{(\alpha,\beta)} S^{(\alpha,\beta)} && (7.40) \\ (\hat{\mu}_{opt}, \hat{\sigma}_{opt}) \quad & \quad\quad \text{gemäß Satz 7.8} && (7.41) \end{aligned}$$

beschrieben werden. Die rechnerische Lösung dieses Problems kann in dieser Form mit dem Verfahren von Gauß-Newton erbracht werden.

Das Problem bei diesem Ansatz liegt darin, daß die Matrix $\hat{\mathbf{B}}$ ebenfalls von den unbekannten Parametern α und β abhängt und somit nicht bekannt ist. Die naheliegende Idee, $\hat{\mathbf{B}}$ durch $\hat{\mathbf{B}}^{(\alpha,\beta)}$ zu schätzen, führt in der Praxis zu divergierenden Schätzern $(\hat{\alpha}, \hat{\beta})$. Dies ist dadurch zu erklären, daß für wachsende Werte von α

und β zwar die Modellanpassung immer schlechter und die Residuen $\hat{\mathbf{e}}^{(\alpha,\beta)}$ immer größer werden, diese Fehler aber andererseits durch die gleichzeitige, stärker wirkende Verringerung der Terme $f_i := fQ^{(\alpha,\beta)}(\pi_i)$ in der Darstellung von $\hat{\mathbf{B}}^{(\alpha,\beta)}$ wieder aufgefangen und sogar annulliert werden.

Als Ausweg aus dieser Situation bietet es sich an, eine analoge Vorgehensweise wie beim IRLS-Verfahren (*Iteratively Reweighted Least Squares*) zu verwenden. Damit läßt sich die iterative Methode zur Anpassung einer fQ-Verteilung an gegebene Daten $\mathbf{x} = (x_1, \ldots, x_n)$ folgendermaßen formulieren:

Algorithmus 7.10

Sei $F(\mu_0, \sigma_0, \alpha_0, \beta_0)$ eine näherungsweise fQ-Dichteschätzung und $\hat{\mathbf{B}}_0 := \hat{\mathbf{B}}^{(\alpha_0,\beta_0)}$ die zugehörige Kovarianzmatrix gemäß (7.37).
Mittels einer Gauß-Newton-Minimierung wird

$$(\alpha_1, \beta_1) := \arg\min_{(\alpha,\beta)} S_0^{(\alpha,\beta)}, \tag{7.42}$$

mit

$$S_0^{(\alpha,\beta)} := \hat{\mathbf{e}}^{(\alpha,\beta)\prime}\, \hat{\mathbf{B}}_\mathbf{0}^{-\mathbf{1}}\, \hat{\mathbf{e}}^{(\alpha,\beta)}$$

bestimmt. Das Verfahren wird mit $F(\hat{\mu}_1, \hat{\sigma}_1, \alpha_1, \beta_1)$ und $\hat{\mathbf{B}}_\mathbf{1} := \hat{\mathbf{B}}^{(\alpha_1,\beta_1)}$ fortgesetzt, wobei die Parameter $\hat{\mu}_1$ und $\hat{\sigma}_1$ mittels einer fQ-ABLU-Schätzung bestimmt werden.
Die Iteration wird abgebrochen, wenn der fQ-Abstand zweier Näherungsschritte hinreichend klein ist:

$$d_{fQ}\left(F(\hat{\mu}_{i-1}, \hat{\sigma}_{i-1}, \alpha_{i-1}, \beta_{i-1}), F(\hat{\mu}_i, \hat{\sigma}_i, \alpha_i, \beta_i)\right) < \epsilon.$$

Das Problem der Schätzung einer Startverteilung $F(\mu_0, \sigma_0, \alpha_0, \beta_0)$ wird dabei im nächsten Abschnitt behandelt. Für die Konvergenz dieses Verfahrens gelten ähnliche Aussagen wie für die IRLS-Methode im Fall der einfachen linearen Regression (siehe auch Ryan [1997, S. 65]):

- Es gibt im allgemeinen keine Garantie für die Konvergenz des Verfahrens.

- Gerade zu Beginn einer Iteration ist zu beobachten, daß die Zielgröße $S_i^{(\alpha,\beta)}$ nicht immer monoton fallend ist. Dies tritt üblicherweise dann ein, wenn die geschätzte Startverteilung nicht gut gewählt ist und die Gauß-Newton-Näherung die optimale Suchrichtung noch nicht gefunden hat.
- In den durchgeführten Untersuchungen reichen typischerweise 3-4 Iterationsschritte aus, um eine hinreichende Stabilität des Verfahrens zu erreichen.

7.5 Ein Kernschätzer für die fQ-Parameter

Zur Initialisierung des iterativen Schätzverfahrens 7.10 ist es notwendig, eine näherungsweise Schätzung $F(\mu_0, \sigma_0, \alpha_0, \beta_0)$ für die zugrundeliegende Verteilung anzugeben. Da empirische Untersuchungen mit intuitiv gewählten Startparametern zeigten, daß das Iterationsverfahren sehr anfällig gegenüber schlecht gewählten Startwerten ist, werden die Startparameter datenorientiert gewählt.
Da dieser Schätzer, der auf einer Kernschätzung der empirischen fQ-Funktion beruht, kein Iterationsverfahren benötigt und somit mit vergleichbar geringem Aufwand bestimmt werden kann, liefert diese Methode einen alternativen Schnellschätzer zum fQ-Dichteschätzer.
Zur Anpassung einer fQ-Verteilung $\mathcal{F}(\mu, \sigma, \alpha, \beta)$ an einen gegebenen Datenvektor $\mathbf{x} = (x_1, \ldots, x_n)$ wird zunächst der Lokationsparameter durch den Median der Daten

$$\hat{\mu} := x_{med}$$

geschätzt. Im folgenden werden nur noch die transformierten Daten $\mathbf{y} := \mathbf{x} - x_{med}$ betrachtet.
Sei weiterhin $\mathbf{p} = (p_1, \ldots, p_k)$ mit $p_i = i/(k+1)$ ein Vektor von äquidistanten Stützstellen, an denen die Dichte-Quantilfunktion geschätzt werden soll. Bei der Wahl der Anzahl k dieser Stützstellen hat sich

$$k := min(n, 20)$$

in der Praxis bewährt, um verläßliche Parameterschätzungen zu erhalten. Seien weiter

$$q_i := Q_n(p_i) = y_{(\lfloor np_i \rfloor + 1)}$$

die Werte der empirische Quantilfunktion der standardisierten Werte $\mathbf{y}$ an den Stellen p_i. Durch Anwendung eines Kernschätzers $\hat{f}$ für die Dichtefunktion erhält man mit

$$\widehat{fQ}_i := \hat{f}(q_i) = \frac{1}{b}\sum_{j=1}^{n} K\left(\frac{q_i - y_j}{b}\right), \quad i = 1, \ldots, k,$$

eine Schätzung für die Dichte-Quantilfunktion fQ an den Stellen q_i, wobei für K der Gaußkern

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{\frac{-u^2}{2}}$$

gewählt wird. Die Bandbreite b wird dabei gemäß der Formel von Silverman (1986) gewählt (siehe (12.2) in Kapitel 12).

Um aus der resultierenden Punktemenge $(p_i, \widehat{fQ}_i))$ eine Schätzung der tatsächlichen fQ-Funktion in der Form

$$fQ(p) = \sigma p^{\alpha}(1-p)^{\beta}, \quad p \in [0,1]$$

zu bestimmen, werden die Werte $\widehat{fQ}_i$ als Dichteschätzung einer diskretisierten Beta-Verteilung $B(a,b)$ an den Stellen p_i aufgefaßt und mittels der Momentenmethode die Parameter dieser Verteilung geschätzt. Dabei ist zu beachten, daß dazu die Werte $\widehat{fQ}_i$ durch die normierten Werte

$$fQ_i^* := \frac{\widehat{fQ}_i}{\sum_{i=1}^{k} \widehat{fQ}_i}$$

ersetzt werden müssen, um eine Dichte zu erhalten. Für die empirischen Momente ergibt sich dann

$$\begin{aligned} m_1 &= \sum_{i=1}^{k} fQ_i^* p_i, \\ m_2 &= \sum_{i=1}^{k} fQ_i^* (p_i - m_1)^2. \end{aligned}$$

Nach Johnson, Kotz und Balakrishnan [1995, S. 228] ergeben sich daraus die Momentenschätzer für die Beta-Verteilung als

$$\begin{aligned} \hat{a} &= \frac{m_1^2(1-m_1)}{m_2} - m_1 \\ \hat{b} &= \frac{\hat{a}(1-m_1)}{m_1}, \end{aligned}$$

wodurch sich dann schließlich aufgrund der unterschiedlichen Parametrisierung der fQ-Funktion die endgültigen Schätzer

$$\hat{\alpha} = \hat{a} - 1 \quad \text{und} \quad \hat{\beta} = \hat{b} - 1 \tag{7.43}$$

ergeben.

Die im letzten Schritt notwendige Schätzung des Skalenparameters σ erfolgt durch eine KQ-Schätzung der Form

$$\hat{\sigma} := \arg\min_{\sigma} \sum_{i=1}^{k} \left(\widehat{fQ}_i - \sigma p_i^{\hat{\alpha}} (1 - p_i)^{\hat{\beta}} \right)^2$$

mit der Lösung

$$\hat{\sigma} = \frac{\sum_{i=1}^{k} \widehat{fQ}_i \, p_i^{\hat{\alpha}} (1 - p_i)^{\hat{\beta}}}{\sum_{i=1}^{k} \widehat{fQ}_i^2}.$$

Teil II

Anwendungen

Kapitel 8

Übersicht

Alle folgenden Kapitel des praktischen Teils orientieren sich an der chronologischen Reihenfolge der Themen des theoretischen Teils. Durch diese Vorgehensweise können die Ergebnisse aus den Vergleichsstudien in den ersten Kapiteln in den späteren, darauf aufbauenden Abschnitten unmittelbar genutzt werden.

Die Darstellung beginnt in Kapitel 9 mit den Ergebnissen für das eigentliche Ziel des fQ-Systems, der Approximation von unimodalen Verteilungen.
Als erste Anwendung werden in Kapitel 10 die Methoden aus Abschnitt 6.3 zur Berechnung der Erwartungswerte von Ordnungsstatistiken für einige Verteilungen mit den exakten Ergebnissen und anderen Approximationsformeln verglichen.
In Kapitel 11 werden anschließend die für die Praxis relevanten fQ-ABLU-Schätzer einerseits den ABLU-Schätzern von Gupta und Blom und andererseits weiteren Schätzverfahren gegenübergestellt, wobei die Vergleiche durch Effizienzaussagen und Simulationen quantifiziert werden können.
Im abschließenden Kapitel 12 wird die fQ-Dichteschätzung anhand systematisch ausgewählter Beispiele mit der Methode der Kernschätzung verglichen.

Um ein möglichst vollständiges Bild der untersuchten Verteilungen zu liefern, konzentrieren sich die Vergleiche auf drei Verteilungen, die an dieser Stelle kurz vorgestellt und motiviert werden sollen:

1) Die Normalverteilung (NV)

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad \mu \in \mathbb{R}, \sigma > 0.$$

2) Die Cauchyverteilung (CV)

$$f(x) = \frac{1}{\pi\sigma}\left(1 + \left(\frac{x-\mu}{\sigma}\right)^2\right)^{-1}, \quad \mu \in \mathbb{R}, \sigma > 0.$$

3) Die Extremwertverteilung, Typ 1 (EV)

$$f(x) = \frac{e^{-\frac{x-\mu}{\sigma}}}{\sigma} \exp\left(-e^{-\frac{x-\mu}{\sigma}}\right), \quad \mu \in \mathbb{R}, \sigma > 0.$$

Die Normalverteilung wird aufgrund ihrer zentralen Stellung in der gesamten Statistik herangezogen. Die mit dem fQ-System erzielten Ergebnisse können direkt mit den entsprechenden Resultaten aus der reichhaltigen Literatur verglichen werden.
Die Cauchyverteilung wird untersucht, da bei dieser Verteilung aufgrund der fehlenden Momente die Schätzung der Lokations- und Skalenparameter im allgemeinen technische Schwierigkeiten bereitet.
Die Extremwertverteilung, oder auch Gumbelverteilung, schließlich wird aufgrund ihrer Rolle bei der Untersuchung von Extremwertstatistiken und ihrer asymmetrischen Dichte in den Vergleich aufgenommen.
Allen drei Verteilungen ist neben ihrer Unimodalität gemeinsam, daß sich ihre Parameter μ und σ und ihre Verteilungsfunktionen in der Form eines LS-Modells (siehe auch Abschnitt 7.2)

$$\mathcal{F} = \{F_{\mu,\sigma}; F_{\mu,\sigma}(x) = F_0(\mu + \sigma x), \sigma \in \mathbb{R}^+, \mu \in \mathbb{R}\}, \tag{8.1}$$

schreiben lassen, für die die fQ-ABLU-Schätzung direkt nutzbar ist.

Kapitel 9

Approximation unimodaler Verteilungen

In Kapitel 6.1 wurde die fQ-Approximation $\hat{F}$ einer unimodalen Verteilung F mit Hilfe einer Minimierungsaussage des fQ-Abstandes definiert:

$$\hat{F} := \arg \min_{G \in \mathcal{F}Q_U} d_{fQ}(F, G), \tag{9.1}$$

mit

$$d_{fQ}(F, G) := \left(\int_0^1 \left(fQ_F(p) - fQ_G(p) \right)^2 \, dp \right)^{1/2}.$$

Da dieser Abstandsbegriff ein globales Maß für die Gleichheit der fQ-Funktionen darstellt, ist zunächst unklar, welche lokale Anpassungsgüte mit $\hat{F}$ erreicht wird. Dies soll anhand der drei in Kapitel 8 erwähnten Verteilungen hauptsächlich mit graphischen Darstellungen untersucht werden.

In Abbildung 2.3 wurde bereits deutlich, daß die Anpassungsqualität nur durch eine Graphik mit den Approximationsfehlern visualisiert werden kann. Demzufolge werden im folgenden nur die Abstände zwischen den exakten und den approximierenden Kurven gezeigt. Im einzelnen werden dabei jeweils die fQ-Funktionen, die Dichtefunktionen, die Verteilungsfunktionen und die Quantilfunktionen der standardisierten Verteilungen ($\mu = 0, \sigma = 1$) mit ihren jeweiligen Approximationen verglichen.

9.1 Die Normalverteilung

Als Ergebnis des Anpassungsverfahrens (9.1) erhält man folgende Approximation der Normalverteilung durch eine fQ-Verteilung:

$$NV \approx \mathcal{F}(0, 1.224, 0.807, 0.807) =: \hat{F}_{NV}. \tag{9.2}$$

Gemäß Satz 3.16 existieren für $\hat{F}_{NV}$ alle Momente und durch numerische Integration (siehe Bemerkung 3.25) erhält man

$$Var(\hat{F}_{NV}) \approx 0.967.$$

Dies ist ein Indiz dafür, daß mittels der fQ-Approximation lediglich eine globale Näherung für die Gestalt der Dichte (bzw. der fQ-Funktion) möglich ist und keine Anpassung direkter analytischer Merkmale (Momente) vorgenommen wird.
In der folgenden Abbildung ist die Qualität der Anpassung von $\hat{F}_{NV}$ an die Normalverteilung dargestellt, wobei in jedem Diagramm immer die Differenz in der Form $fQ_{NV} - fQ_{\hat{F}_{NV}}$ etc. abgetragen ist.

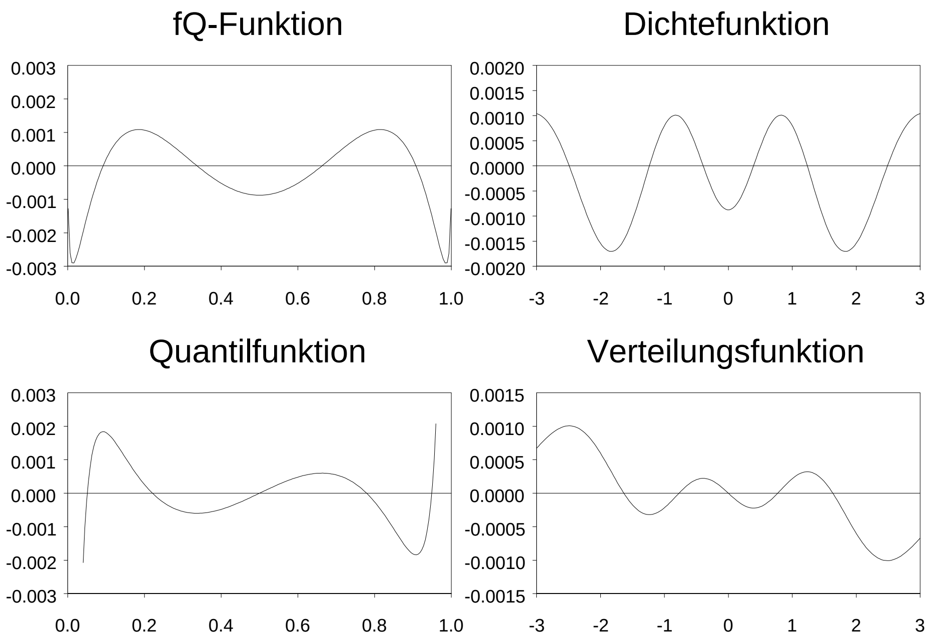

Abbildung 9.1: Approximationsfehler bei der Normalverteilung

Zunächst kann man der Gestalt der Kurven entnehmen, daß die Anpassung zu einem systematischen, nichtlinearen Fehler führt. Die Größenordnungen der absoluten Fehler sind vergleichbar mit denen von speziellen Näherungsmethoden für die Normalverteilung (siehe Johnson, Kotz und Balakrishnan [1995, S. 111ff]). Eine Sonderstellung nimmt die Graphik für die Approximation der Quantilfunktion ein, die in obiger Abbildung zur Verdeutlichung der lokalen Anpassungsgüte nur für den Bereich $p \in [0.04, 0.96]$ gezeichnet wurde. Eine genauere Darstellung für den Bereich $p \in [0.0001, 0.05]$ zeigt folgende Abbildung:

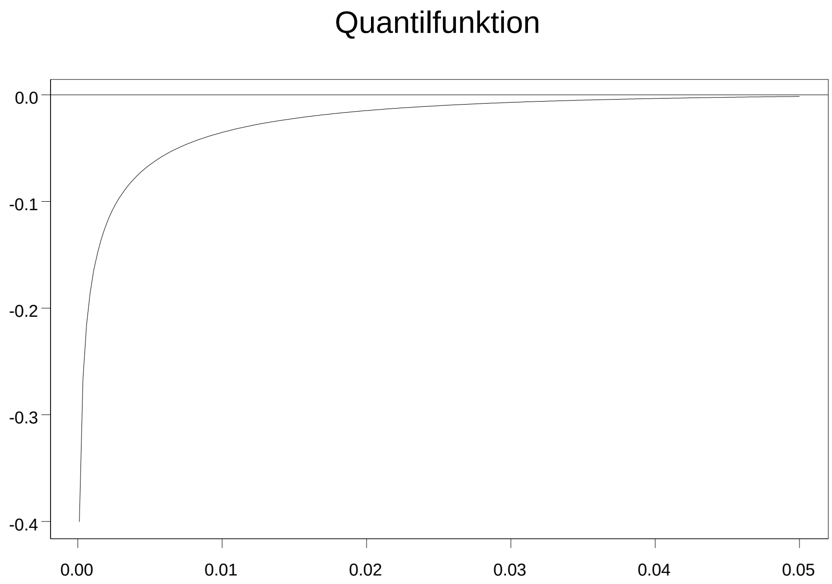

Abbildung 9.2: Approximationsfehler bei der Quantilfunktion der Normalverteilung

Die sinkende Approximationsgüte für $p \to 0$ erkärt sich aus Satz 3.14, da für den linken Rand des Definitionsbereichs von $\hat{F}_{NV}$ gilt:

$$\alpha(\hat{F}_{NV}) \approx -\frac{B(1-0.807, 1-0.807)}{2 \cdot 1.224} \approx -4.037 \approx F_{NV}(0.000027).$$

Die Approximation der Quantilfunktion durch eine fQ-Verteilung mit einem endlichen Träger muß somit zwangsläufig zu einem wachsenden Fehler jenseits dieser Grenze führen.

Um in weiteren Anwendungen der fQ-Approximation Bereiche angeben zu können, in denen der absolute Fehler der Quantilfunktion begrenzt ist, sei hier allgemein definiert:

Definition 9.1
Sei $\hat{F} \in \mathcal{F}\mathcal{Q}_U$ die fQ-Approximation einer Verteilung $F \in \mathcal{F}_U$. Dann heißt für $s > 0$ die Menge $D_{\hat{F}}^{(s)}$ mit

$$D_{\hat{F}}^{(s)} := \{p \in [\,0,1] : |Q_F(p) - Q_{\hat{F}}(p)| \leq s\}$$

die s-Approximationsmenge *von $\hat{F}$.*
Ist $D_{\hat{F}}^{(s)}$ ein zusammenhängendes Intervall mit den Grenzen

$$\alpha_{\hat{F}}^{(s)} := \inf_{p \in [\,0,1]} \left\{ \hat{D}_{\hat{F}}^{(s)} \right\}, \tag{9.3}$$

$$\omega_{\hat{F}}^{(s)} := \sup_{p \in [\,0,1]} \left\{ \hat{D}_{\hat{F}}^{(s)} \right\}, \tag{9.4}$$

so heißt $D_{\hat{F}}^{(s)}$ das s-Approximationsintervall *von $\hat{F}$.*

Bemerkung 9.2
Wie das Beispiel der Normalverteilung zeigt (siehe auch Abbildung 9.3), führt der nichtlineare Approximationsfehler dazu, daß mit $s \to 0$ die Menge $D_{\hat{F}}^{(s)}$ zwangsläufig unzusammenhängend ist. Da die konkrete Notation von $D_{\hat{F}}^{(s)}$ in diesen Fällen kompliziert sein kann, werden im folgenden lediglich Approximationsintevalle betrachtet.

Der Nutzen dieser Definition liegt darin, daß in den Anwendungen, in denen die fQ-Approximation spezieller Verteilungen genutzt wird, durch eine Einschränkung auf den *sicheren* Bereich $D_{\hat{F}}^{(s)}$ bessere Ergebnisse erzielt werden können.
Speziell für die Normalverteilung gilt:

$$D_{\hat{F}_{NV}}^{(0.01)} \approx [0.033, 0.967], \tag{9.5}$$

wie in folgender Abbildung veranschaulicht wird:

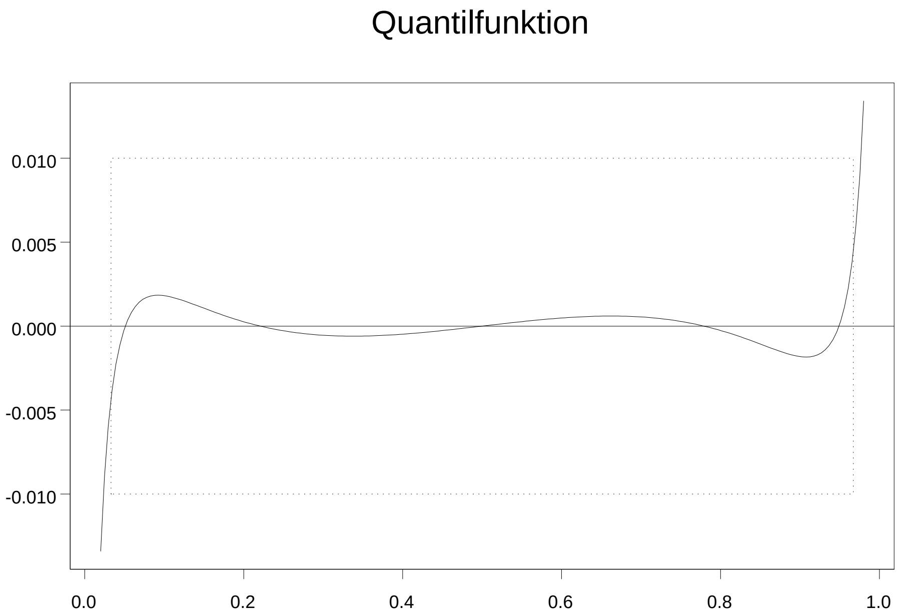

Abbildung 9.3: Das 0.01-Approximationsintervall der Normalverteilung

Bemerkung 9.3

Da die Parameter von $\hat{F}_{NV}$ *der Bedingung* $\alpha, \beta < 1$ *genügen, läßt sich durch Anwendung von Satz 3.11 die Quantilfunktion der Normalverteilung in der folgenden Form approximieren:*

$$Q_{NV}(p) \approx 8.0725 \cdot (F_{a,b}(p) - 1/2), \tag{9.6}$$

wobei $F_{a,b}$ *die Verteilungsfunktion einer Beta-Verteilung mit den Parametern* $a = b = 0.19295$ *ist.*

Der absolute Fehler dieser Näherung ist für $p \in [\,0.01, 0.99]$ *kleiner als* 0.0339 *und für* $p \in [\,0.1, 0.9]$ *kleiner als* 0.0018.

9.2 Die Cauchyverteilung

Für die Cauchyverteilung ergibt sich die fQ-Darstellung

$$CV \approx \mathcal{F}(0, 8.444, 2.368, 2.368) =: \hat{F}_{CV}, \tag{9.7}$$

woraus mit Satz 3.16 folgt, daß für $\hat{F}_{CV}$ kein Moment existiert. Die entsprechenden Approximationen führen zu folgenden Kurven:

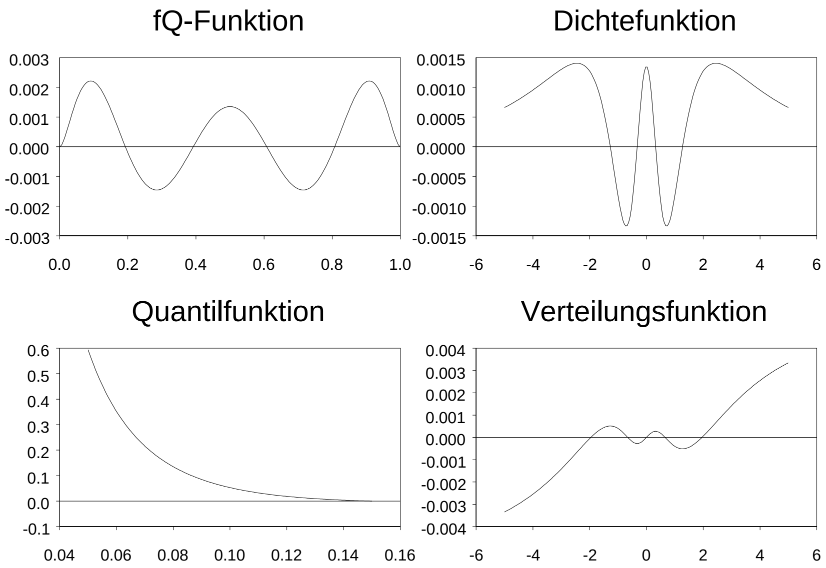

Abbildung 9.4: Approximationsfehler bei der Cauchyverteilung

Während sich bei der fQ-Funktion, der Dichtefunktion und der Verteilungsfunktion erneut eine zyklische, systematische Fehlerstruktur zeigt, kehrt sich das Vorzeichen bei der Anpassung der Quantilfunktion im Vergleich zur Normalverteilung gerade um, die Quantilfunktion von $\hat{F}_{CV}$ liegt also stets unterhalb der exakten Quantilfunktion. Darüber hinaus zeigt der gewählte Ausschnitt auf der x-Achse, daß die Qualität der Anpassung wesentlich schlechter ist als bei der Normalverteilung, was ebenfalls durch das kleinere 0.01-Approximationsintervall

$$D^{(0.01)}_{\hat{F}_{CV}} \approx [0.129, 0.871]. \tag{9.8}$$

ausgedrückt wird.

9.3 Die Extremwertverteilung

Die Extremwertverteilung wird approximiert durch die fQ-Verteilung mit den Parametern

$$EV \approx \mathcal{F}(0.367, 1.152, 0.650, 1.073) =: \hat{F}_{EV}. \tag{9.9}$$

Nach Satz 3.16 existieren die Momente von $\hat{F}_{EV}$ nur bis zur Ordnung 13, da

$$\max(0.650, 1.073) > 1 + 1/14 \approx 1.0714.$$

Im Gegensatz dazu existieren für die Extremwertverteilung alle Momente. Diese Erscheinung ist typisch für viele fQ-Approximationen, deren geschätzte Parameter häufig auf die fehlende Existenz von sehr hohen Momenten hindeuten. In der Graphik der Approximationen ist ebenfalls eine asymmetrische Anpassungsgüte zu finden.

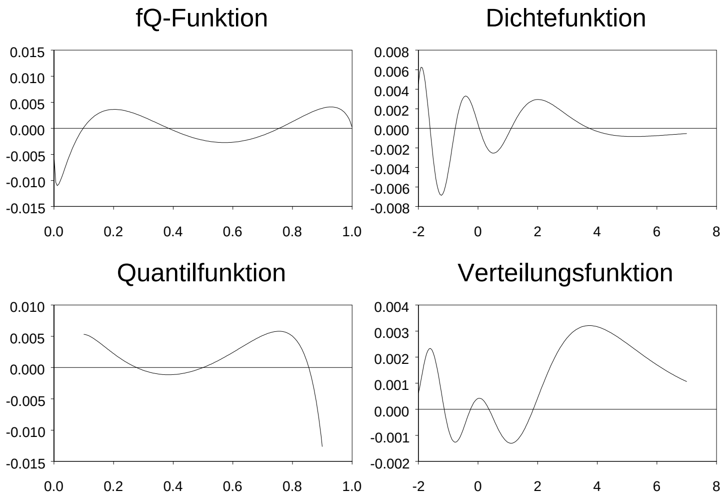

Abbildung 9.5: Approximationsfehler bei der Extremwertverteilung

Entsprechend asymmetrisch ist somit auch das 0.01-Approximationsintervall:

$$D_{\hat{F}_{EV}}^{(0.01)} \approx [0.030, 0.894]. \tag{9.10}$$

9.4 Weitere Verteilungen

In der folgenden tabellarischen Übersicht sind die fQ-Darstellungen einiger Verteilungen zu finden. Soweit nicht anders angegeben, haben die Verteilungen dabei ihre kanonische Standardform $\mu = 0$ und $\sigma = 1$. Für die Notationen der Dichten sei auf Tabelle 2.1 beziehungsweise auf Formelsammlungen verwiesen. Da der Lokationsparameter μ aus der Bedingung $\hat{F}(\mu) = F(\mu)$ gewonnen wird und somit unabhänigig von der eigentlichen fQ-Anpassung bestimmt werden kann, wird er nicht notiert.

Zunächst werden die Verteilungen aufgeführt, deren fQ-Funktionen im fQ-System exakt dargestellt werden können:

Name	σ	α	β
Rechteck	1	0	0
Exponential	1	0	1
Logistisch	1	1	1
Potenz	c	0	$1 - c^{-1}$
Pareto	c^{-1}	0	$1 + c$

Tabelle 9.1: Exakte fQ-Darstellungen

In den folgenden Tabellen sind neben den fQ-Parametern jeweils auch die Kennmaße μ_S, μ_D^F, μ_D^S sowie die Grenzen eines Approximationsintervalls angegeben. Die Daten können genutzt werden, um mit den Parametern Schätzungen für die Erwartungswerte der Ordnungstatistiken (Kapitel 10) oder fQ-ABLU-Schätzer (Kapitel 11) zu konstruieren. Zur Vervollständigung seien zunächst noch einmal die Darstellungen der bereits untersuchten Verteilungen notiert:

Name	σ	α	β	μ_S	μ_D^F	μ_D^S	$\alpha_{\hat{F}}^{(0.01)}$	$\omega_{\hat{F}}^{(0.01)}$
NV	1.224	0.807	0.807	0.000	1.417	2.501	0.033	0.967
CV	8.444	2.368	2.368	0.000	1.996	3.155	0.129	0.871
EV	1.152	0.650	1.073	-0.245	1.471	2.721	0.030	0.894

Tabelle 9.2: fQ-Approximation der untersuchten Verteilungen

df	σ	α	β	μ_D^F	μ_D^S	$\alpha_{\hat{F}}^{(0.01)}$	$\omega_{\hat{F}}^{(0.01)}$
1.00	8.444	2.368	2.368	1.996	3.155	0.129	0.871
2.00	2.828	1.500	1.500	1.698	2.828	0.054	0.946
3.00	2.077	1.248	1.248	1.601	2.718	0.054	0.946
4.00	1.799	1.130	1.130	1.554	2.663	0.050	0.950
5.00	1.657	1.062	1.062	1.526	2.630	0.047	0.953
6.00	1.570	1.017	1.017	1.507	2.608	0.044	0.956
7.00	1.513	0.986	0.986	1.494	2.592	0.043	0.957
8.00	1.471	0.962	0.962	1.484	2.581	0.042	0.958
9.00	1.440	0.945	0.945	1.477	2.572	0.041	0.959
10.00	1.416	0.930	0.930	1.470	2.565	0.040	0.960
20.00	1.315	0.868	0.868	1.443	2.533	0.036	0.964
50.00	1.259	0.831	0.831	1.427	2.514	0.034	0.966
100.00	1.241	0.819	0.819	1.422	2.507	0.033	0.967

Tabelle 9.3: fQ-Approximation der t-Verteilung

c	σ	α	β	μ_S	μ_D^F	μ_D^S	$\alpha_{\hat{F}}^{(0.01)}$	$\omega_{\hat{F}}^{(0.01)}$
1.50	1.530	0.340	0.867	−0.436	1.404	1.340	0.001	0.951
2.00	2.080	0.515	0.807	−0.221	1.371	1.164	0.005	0.961
2.50	2.640	0.623	0.773	−0.108	1.372	0.989	0.007	0.968
3.00	3.204	0.695	0.751	−0.039	1.379	0.850	0.007	0.973
3.50	3.771	0.747	0.736	0.008	1.387	0.741	0.007	0.976
4.00	4.341	0.787	0.725	0.041	1.394	0.656	0.007	0.979

Tabelle 9.4: fQ-Approximation der Weibull-Verteilung

df	σ	α	β	μ_S	μ_D^F	μ_D^S	$\alpha_{\hat{F}}^{(0.01)}$	$\omega_{\hat{F}}^{(0.01)}$
3.00	0.468	0.269	1.035	−0.587	1.514	4.154	0.007	0.842
4.00	0.420	0.386	1.024	−0.452	1.468	5.455	0.016	0.853
5.00	0.380	0.453	1.009	−0.380	1.451	6.500	0.021	0.671
6.00	0.350	0.497	0.997	−0.334	1.442	7.399	0.090	0.659
7.00	0.325	0.529	0.986	−0.302	1.437	8.200	0.118	0.653
8.00	0.305	0.553	0.976	−0.277	1.434	8.930	0.134	0.650
9.00	0.288	0.572	0.968	−0.258	1.431	9.605	0.146	0.647
10.00	0.273	0.587	0.961	−0.242	1.429	10.235	0.155	0.646

Tabelle 9.5: fQ-Approximation der χ^2-Verteilung

σ^*	σ	α	β	μ_S	μ_D^F	μ_D^S	$\alpha_{\hat{F}}^{(0.01)}$	$\omega_{\hat{F}}^{(0.01)}$
0.10	12.274	0.749	0.869	−0.074	1.420	0.249	0.009	0.990
0.20	6.188	0.694	0.934	−0.147	1.430	0.491	0.013	0.986
0.30	4.182	0.642	1.003	−0.220	1.447	0.719	0.015	0.988
0.40	3.198	0.593	1.077	−0.290	1.471	0.927	0.015	0.921
0.50	2.622	0.546	1.156	−0.358	1.502	1.109	0.014	0.887
0.60	2.252	0.502	1.240	−0.424	1.541	1.263	0.013	0.856
0.70	2.001	0.459	1.329	−0.486	1.589	1.384	0.012	0.825
0.80	1.825	0.419	1.425	−0.545	1.646	1.472	0.011	0.794
0.90	1.701	0.381	1.529	−0.601	1.713	1.527	0.010	0.763
1.00	1.614	0.345	1.640	−0.653	1.791	1.549	0.010	0.731

Tabelle 9.6: fQ-Approximation der Log-Normalverteilung

f_2	σ	α	β	μ_S	μ_D^F	μ_D^S	$\alpha_{\hat{F}}^{(0.02)}$	$\omega_{\hat{F}}^{(0.02)}$
3.00	1.379	0.242	1.863	−0.770	2.014	1.536	0.002	0.687
4.00	1.380	0.247	1.660	−0.741	1.888	1.510	0.002	0.728
5.00	1.382	0.250	1.538	−0.720	1.813	1.492	0.002	0.755
10.00	1.389	0.258	1.290	−0.666	1.663	1.447	0.001	0.814
20.00	1.395	0.263	1.163	−0.631	1.588	1.418	0.001	0.847
50.00	1.399	0.267	1.087	−0.606	1.543	1.399	0.001	0.868
100.00	1.401	0.268	1.061	−0.597	1.529	1.392	0.001	0.875

Tabelle 9.7: fQ-Approximation der $F(3, f_2)$-Verteilung

f_2	σ	α	β	μ_S	μ_D^F	μ_D^S	$\alpha_{\hat{F}}^{(0.02)}$	$\omega_{\hat{F}}^{(0.02)}$
3.00	1.550	0.329	1.920	−0.708	1.953	1.645	0.008	0.686
4.00	1.566	0.338	1.705	−0.669	1.831	1.598	0.007	0.729
5.00	1.578	0.345	1.575	−0.641	1.757	1.565	0.006	0.758
10.00	1.616	0.362	1.307	−0.566	1.612	1.481	0.006	0.822
20.00	1.643	0.373	1.168	−0.516	1.539	1.428	0.006	0.858
50.00	1.663	0.381	1.082	−0.479	1.496	1.391	0.006	0.881
100.00	1.670	0.383	1.053	−0.466	1.482	1.378	0.006	0.889

Tabelle 9.8: fQ-Approximation der $F(4, f_2)$-Verteilung

c	σ	α	β	μ_S	μ_D^F	μ_D^S	$\alpha_{\hat{F}}^{(0.01)}$	$\omega_{\hat{F}}^{(0.01)}$
1.50	0.935	0.269	1.035	−0.587	1.514	2.077	0.003	0.865
2.00	0.839	0.386	1.024	−0.452	1.468	2.728	0.011	0.873
2.50	0.761	0.453	1.009	−0.380	1.451	3.250	0.016	0.879
3.00	0.699	0.497	0.997	−0.334	1.442	3.699	0.020	0.884
3.50	0.650	0.529	0.986	−0.302	1.437	4.100	0.024	0.888
4.00	0.609	0.553	0.976	−0.277	1.434	4.465	0.026	0.891

Tabelle 9.9: fQ-Approximation der Γ-Verteilung

Kapitel 10

Erwartungswerte von Ordungsstatistiken

Als erste Anwendung der Darstellungen aus Kapitel 9 werden nun die Approximationen der Erwartungswerte von Ordnungsstatistiken gemäß den Formeln aus Abschnitt 6.3 betrachtet. Für die drei Verteilungen NV, CV und EV werden für verschiedene Stichprobenumfänge folgende Berechnungsmethoden für die Werte $E(X_{(i)})$ für eine Stichprobe $\mathbf{X} = (X_{(1)}, \ldots, X_{(n)})$ mit $X_i \sim \mathcal{F}_U$ miteinander verglichen:

- Die exakten Werte durch eine numerische Integration des Ausdrucks (E)

$$n\binom{n-1}{i-1}\int_0^1 Q(p)\, p^{i-1}(1-p)^{n-i}\, dp \tag{10.1}$$

- Die Approximation nach David und Johnson (1980) (DJ)

$$Q(p_i) + \frac{p_i(1-p_i)}{2(n+2)}Q''(p_i), \quad p_i = \frac{i}{n+1} \tag{10.2}$$

- Die Näherungsformeln nach Satz 6.5 durch die fQ-Approximationen

$$Q(\pi_i), \quad \pi_i = \frac{i-\frac{\alpha}{2}}{n-\frac{\alpha+\beta}{2}+1}. \tag{10.3}$$

In den jeweiligen Tabellen werden zur Beurteilung der Anpassungsqualität die maximale absolute Abweichung (MA) und die Summe der quadratischen

Abstände (SQ) angegeben, wobei die Notation SQ^a bedeuten soll, daß die Werte für SQ aufgrund ihrer Größenordnung zur Verdeutlichung mit einem Faktor 10^a multipliziert wurden.

10.1 Normalverteilung

Für die Normalverteilung werden neben den eingangs erwähnten Größen ebenfalls die Näherungsformeln von Blom [1958, S. 71] (BNV1 und BNV2)

$$\Phi^{-1}\left(\frac{i-3/8}{n+1/4}\right), \quad \Phi^{-1}\left(\frac{i-1/2}{n}\right). \tag{10.4}$$

aufgeführt. Da diese Formeln speziell für die Normalverteilung entwickelt wurden, ist hier eine hohe Anpassungsgüte zu erwarten. Die folgende Tabelle zeigt die Näherungswerte für den Fall $n = 21$:

i	E	DJ	BNV1	BNV2	fQ
20	1.4336	1.4248	1.4292	1.4652	1.4389
19	1.1605	1.1556	1.1575	1.1798	1.1640
18	0.9538	0.9506	0.9518	0.9674	0.9561
17	0.7815	0.7792	0.7802	0.7916	0.7828
16	0.6298	0.6282	0.6289	0.6375	0.6305
15	0.4915	0.4903	0.4909	0.4972	0.4917
14	0.3620	0.3612	0.3616	0.3661	0.3620
13	0.2384	0.2379	0.2382	0.2410	0.2383
12	0.1184	0.1181	0.1182	0.1196	0.1183
11	0.0000	0.0000	0.0000	0.0000	0.0000
MA		0.0984	0.0044	0.0316	0.0052
SQ^4		18.184	0.0352	1.7762	0.0475

Tabelle 10.1: Approximationen für die Normalverteilung ($n = 21$, $i = 11, (1), 20$)

Folgende Ergebnisse lassen sich festhalten:

- Die asymptotischen Formeln von David und Johnson liefern aufgrund des geringen Stichprobenumfangs nur unzureichende Ergebnisse.

- Die zweite Näherungsformel von Blom (BNV2) ist deutlich schlechter als die erste Variante (BNV1)

- Die fQ-Näherung ist insgesamt qualitativ mit der Formel (BNV1) vergleichbar, führt jedoch an den Rändern zu größeren Abweichungen.

Für einen größeren Stichprobenumfang ($n = 101$) ergibt sich folgende Tabelle:

i	E	DJ	BNV1	BNV2	fQ
101	2.5111	2.4940	2.5021	2.5793	2.4617
96	1.5962	1.5953	1.5932	1.6031	1.5960
91	1.2552	1.2549	1.2539	1.2593	1.2569
86	1.0188	1.0186	1.0181	1.0217	1.0199
81	0.8289	0.8288	0.8285	0.8311	0.8293
76	0.6651	0.6650	0.6648	0.6667	0.6649
71	0.5175	0.5175	0.5173	0.5187	0.5171
66	0.3805	0.3804	0.3803	0.3813	0.3800
61	0.2502	0.2502	0.2502	0.2508	0.2498
56	0.1241	0.1241	0.1241	0.1244	0.1239
51	0.0000	0.0000	0.0000	0.0000	0.0000
MA		0.0171	0.0090	0.0681	0.0494
SQ^4		0.2935	0.0931	4.7231	2.4483
$\widetilde{SQ}^4$		0.0009	0.0113	0.0829	0.0049

Tabelle 10.2: Approximation für die Normalverteilung ($n = 101$,$i = 51, (5), 101$)

Dabei wurde bei den Werten in der Zeile $\widetilde{SQ}^4$ die Abweichung für $i = 101$ von der Berechnung ausgeschlossen.

Hier gilt:

- Alle Verfahren sind ungenau bei der Schätzung des Maximums $E(X_{(n)})$.
- Das approximative Verfahren (DJ) ist erwartungsgemäß deutlich besser als im vorherigen Fall.
- Das fQ-Verfahren und die Näherung von Blom (BNV1) liefern vergleichbare Resultate.
- Wird die fQ-Approximation nur für die Fälle $i/n \in D^{(0.01)}_{\hat{F}_{NV}}$, also für $i < 97$ betrachtet, so ist die fQ-Näherung den Formeln von Blom überlegen.

10.2 Cauchyverteilung

Aufgrund der größeren Tailstärke der Cauchyverteilung ist zu erwarten, daß die Güte der Approximationen hier wesentlich schlechter ist als bei der Normalverteilung. Die folgende Tabelle für den Fall $n = 101$ bestätigt diesen Eindruck:

i	E	DJ	BCVO	BCVU	fQ
91	3.101	3.111	3.140	3.078	3.184
86	1.976	1.985	1.988	1.963	1.983
81	1.385	1.395	1.389	1.376	1.382
76	1.006	1.016	1.007	1.000	1.003
71	0.730	0.741	0.731	0.727	0.730
66	0.512	0.523	0.512	0.510	0.512
61	0.326	0.339	0.327	0.325	0.327
56	0.159	0.173	0.159	0.158	0.160
51	0.000	0.016	0.000	0.000	0.000
MA		0.0101	0.0391	0.0229	0.0831
SQ^4		1.200	1.660	0.8358	6.952
$\widetilde{SQ}^4$		1.106	0.167	0.2971	0.0675

Tabelle 10.3: Approximationen für die Cauchyverteilung $(n = 101, i = 51(5), 91)$

Dabei wurden die von Blom [1958, S. 73] vorgeschlagenen Abschätzungen (BCVU) und (BCVO)

$$\tan\left(\pi\left(\frac{i-1.23}{n-1.46}-\frac{1}{2}\right)\right) \leq E(X_{(i)}) \leq \tan\left(\pi\left(\frac{i-1}{n-1}-\frac{1}{2}\right)\right), \tag{10.5}$$

sowie die Näherungsformel von David und Johnson (1980) (DJ) in der von Johnson, Kotz und Balakrishnan [1994, S. 305] angegebenen Form

$$\tan(\pi(i/n-1/2)) \tag{10.6}$$

verwendet. Mittels der folgenden Graphik können die Zahlen veranschaulicht werden:

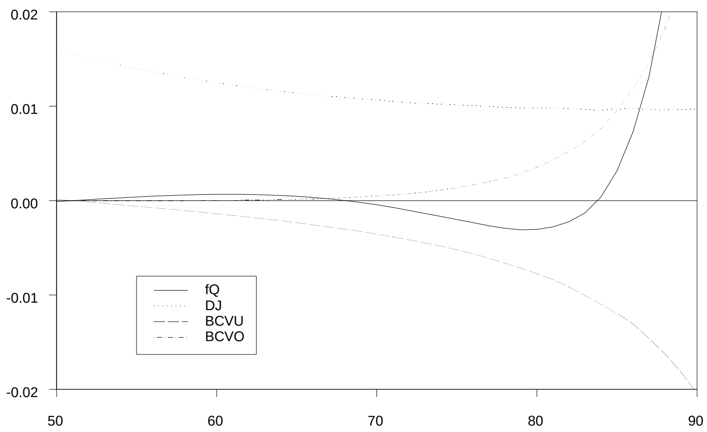

Abbildung 10.1: Güte der Approximation von $E(X_{(i)})$ bei der Cauchyverteilung

Folgende qualitativen Aussagen lassen sich festhalten:

- Die Näherung von Johnson und David (DJ) ist in der gewählten Form offensichtlich verzerrt. Dies ist darauf zurückzuführen, daß in der Formel (10.6) der Anteil der vernachlässigten weiteren Terme zu groß ist.
- Die Formeln von Blom liefern ein gleichförmiges Approximationsintervall, wobei die obere Abschätzung (BNVO) für kleine i bessere Werte liefert.
- Unter Berücksichtigung des Approximationsintervalls $D^{(0.01)}_{\hat{F}_{CV}}$ ist die fQ-Näherung das genaueste Verfahren.

10.3 Extremwertverteilung

Zur Berechung der Formel von David und Johnson (10.2) wird in diesem Fall auch die zweite Ableitung der Quantilfunktion $Q_{EV}(p)$

$$Q''_{EV}(p) = \frac{1}{p^2 \log(p)} + \frac{1}{p^2 \log(p)^2},$$

verwendet. Zum Vergleich werden die von Blom [1958, S. 75] (BEV1 und BEV2) vorgeschlagenen Abschätzungen

$$Q_{EV}\left(\frac{i-1/2}{n}\right) \leq E(X_{(i)}) \leq Q_{EV}\left(\frac{i-1/4}{n+1/4}\right). \tag{10.7}$$

herangezogen.

Für das Beispiel ($n = 101$) wurden die Approximationsgrenzen gemäß der Beziehung (9.10) mit $D^{(0.01)}_{\hat{F}_{EV}} \approx [0.030, 0.894]$ berücksichtigt und demnach nur die Approximationen für $E(X_{(i)})$ mit $i = 4, \ldots, 89$ berechnet.

Als Ergebnis läßt sich folgende Graphik festhalten:

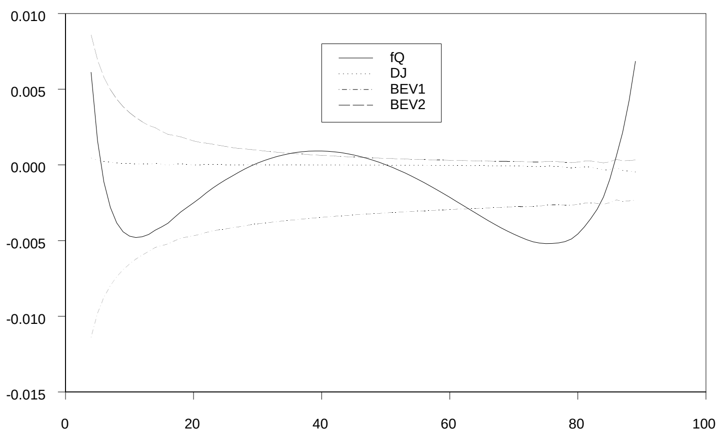

Abbildung 10.2: Güte der Approximation von $E(X_{(i)})$ bei der Extremwertverteilung

In diesem Fall bringt die Approximation von David und Johnson eine nahezu optimale Anpassung. Auffallend ist hier die Tatsache, daß die fQ-Anpassung die Grenzen der Näherung von Blom mehrfach schneidet.

Zusammenfassung

Der Theorie entsprechend liefert das allgemeine Approximationsverfahren von David und Johnson unabhängig von der Verteilung bei großem Stichprobenumfang n die beste Näherung. Diese Optimalität geht jedoch bei geringen Werten von n verloren. In diesen Fällen liefern die speziellen Formeln von Blom und die fQ-Näherung eine deutlich bessere Anpassung.
Gegenüber den anderen Verfahren besitzt die fQ-Methode folgende Vorteile:

- Es gibt nur für sehr wenige Verteilungen optimierte Formeln wie die von Blom.

- Die Näherungen von David und Johnson erfordern für akzeptable Resultate die Berechnung der zweiten Ableitung der Quantilfunktion, wie das Beispiel der Cauchyverteilung zeigt. Dies ist für einige Verteilungen keine triviale Aufgabe.

- Die fQ-Näherung ist ein allgemeines Konstruktionsprinzip, zu dessen Anwendung lediglich die fQ-Parameter der Verteilung bekannt sein müssen.

- Werden die fQ-Parameter anhand empirischer Daten geschätzt, so liefert diese Methode Schätzungen für die Erwartungswerte der Ordnungsstatistiken der *unbekannten* Grundverteilung.

Kapitel 11

Vergleich von L-Schätzern

In diesem Kapitel werden die fQ-ABLU-Schätzer im LS-Modell für die Verteilungen NV,CV und EV in Simulationsstudien mit anderen Schätzverfahren für diese Verteilungen verglichen. Anhand dieser drei Fallbeispiele soll die Frage geklärt werden, ob die fQ-ABLU-Schätzung, die aufgrund des allgemeinen Konstruktionsprinzips für beliebige stetige unimodale Verteilungen definiert ist, eine sinnvolle Alternative zu bekannten Schätzverfahren für diese spezielle Verteilungen darstellt.

Nach der zusammenfassenden Darstellung der Konstruktion der fQ-ABLU-Schätzer im nächsten Abschnitt werden anschließend die drei obigen Verteilungen gesondert untersucht. In jedem dieser Kapitel werden dabei zunächst allgemeine Aussagen über die Schätzung von μ und σ in dieser Verteilungsklasse gemacht und die speziellen Schätzer vorgestellt, die in den Vergleichsstudien berücksichtigt werden.

In allen Simulationen werden für jede Verteilung und jeden Stichprobenumfang 10000 Datensätze erzeugt und die Schätzer jeweils auf den selben Datensatz angewandt. Alle Berechnungen wurden mit S-PLUS 4.0 durchgeführt und es wurden auch die dort implementierten Zufallszahlengeneratoren benutzt. In den Tabellen werden für jeden Schätzer $\hat{\mu}$ bzw. $\hat{\sigma}$ jeweils der Mittelwert der Schätzwerte (*Est*), die Verzerrung (*Bias*), die Varianz (*Var*) und der mittlere quadratische Fehler (*MSE*) aufgeführt. Soweit nicht anders angegeben, wird in der Spalte mit der Bezeichnung *Eff* die geschätzte Effizienz des betreffenden Schätzers bezüglich der Rao-Cramér-Schranke in der jeweiligen Verteilungsklasse aufgelistet.

11.1 Konstruktion der fQ-ABLU-Schätzer

Gegeben sei eine stetige unimodale Verteilung $F \in \mathcal{F}_U$ mit fQ-Darstellung $\hat{F} = \mathcal{F}(\mu, \sigma, \alpha, \beta)$ und eine Stichprobe $X_i, i = 1, \ldots, n$ aus dem LS-Modell mit Grundverteilung F, d.h.

$$X_i \sim F(\mu + \sigma x), \sigma \in \mathbb{R}^+, \mu \in \mathbb{R}.$$

Der geordnete Datenvektor $\mathbf{X} = (X_{(1)}, \ldots, X_{(n)})$ werde in der Form

$$\tilde{\mathbf{X}} = (X_{(n_1+1)}, \ldots, X_{(n-n_2)}) := (\tilde{X}_1, \ldots, \tilde{X}_k),$$

eingeschränkt (zensiert), wobei n_1 und n_2 durch die Beziehungen

$$\begin{aligned} n_1 + 1 &= \lceil n \cdot \alpha_{\hat{F}}(s) \rceil, \\ n - n_2 &= \lfloor n \cdot \omega_{\hat{F}}(s) \rfloor, \end{aligned}$$

mit einem $s > 0$ bestimmt werden, wobei in den folgenden Beispielen der Wert $s = 0.01$ gewählt wurde.
Seien dann

$$\begin{aligned} \hat{a}_i &:= \hat{Q}(\pi_i) \qquad \text{und} \\ \hat{b}_{i,j} &:= \frac{\pi_i(1-\pi_j)}{n - \frac{\alpha+\beta}{2} + 2} \left(\widehat{fQ}(\pi_i)\widehat{fQ}(\pi_j)\right)^{-1}, \end{aligned}$$

mit

$$\pi_i := \frac{i - \alpha/2}{n - (\alpha+\beta)/2 + 1}, \quad n_1 + 1 \le i \le n - n_2$$

die Schätzungen für den Vektor der Erwartungswerte und für die Kovarianzmatrix. Dann sind die (eingeschränkten) fQ-ABLU-Schätzer nach Satz 7.7 durch die Formeln

$$\hat{\mu}_{fQ} = \sum_{i=1}^{k} c_i^* \cdot \tilde{X}_{(i)}, \quad \hat{\sigma}_{fQ} = \sum_{i=1}^{k} d_i^* \cdot \tilde{X}_{(i)}$$

gegeben, wobei für die exakte Darstellung der Koeffizienten c_i^* und d_i^* auf den Abschnitt 7.3 verwiesen sei.

11.2 Normalverteilung

Da die Theorie der Lokations- und Skalenschätzer im Normalverteilungsmodell sehr umfangreich ist (siehe Johnson, Kotz und Balakrishnan [1994, S. 123-151]), werden hier nur Schätzer für die spezielle Situation der Zensierung betrachtet. Wie Johnson, Kotz und Balakrishnan bemerken [S. 132], wäre im Falle der Normalverteilung für unzensierte Daten im übrigen der Einsatz von (A)BLU-Schätzer sinnlos, da die Effizienzen dieser Schätzer nicht besser sind als die der einfacher zu berechnenden ML-Schätzer.

Für den unzensierten Datensatz ist die Rao-Cramér-Schranke bei einer unverzerrten Schätzung von μ gegeben durch

$$Var(\hat{\mu}) = \frac{\sigma^2}{n}.$$

Der unverzerrte ML-Schätzer

$$\hat{\sigma}_{ML} = a_n \cdot S := \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \sqrt{\frac{n}{2}} \cdot S \tag{11.1}$$

für σ hat die Varianz

$$Var(\hat{\sigma}_{ML}) = \left(a_n^2(1 - \frac{1}{n}) - 1\right)\sigma^2.$$

Diese Werte werden später bei den Effizienzvergleichen herangezogen.

Um die verwendeten Schätzer direkt vergleichen zu können, werden die Indexgrenzen des zensierten Datenvektors $\tilde{\mathbf{X}} = (X_{(n_1+1)}, \ldots, X_{(n-n_2)})$ für alle betrachteten Schätzer durch die Beziehung

$$n_1 = n_2 = \lceil n \cdot \alpha_{\hat{F}_{NV}}(0.01) \rceil - 1 = \lceil n \cdot 0.033 \rceil - 1$$

mit der Approximation $\hat{F}_{NV}$ gemäß (9.2) bestimmt.

Neben dem eingeschränkten fQ-ABLU-Schätzer und $\hat{\sigma}_{ML}$ (11.1) werden folgende Schätzer betrachtet:

- Die ABLU-Schätzer von Gupta $\hat{\mu}_G$ und $\hat{\sigma}_G$ nach Satz 7.6. Dabei ist zu beachten, daß aufgrund der Konstruktion $\hat{\mu}_G = \bar{X} = \hat{\mu}_{ML}$ gilt.
- Die ABLU-Schätzer von Blom $\hat{\mu}_B$ und $\hat{\sigma}_B$ in der Gestalt (7.33) und (7.34), wobei für die Korrekturterme in der Formel für die π_i (siehe 7.35) gerade die Empfehlungen aus Kapitel 10 benutzt werden, d.h. $\hat{c}_{i,j} = \hat{d}_{i,j} = 3/8$.

- Das winsorisierte Mittel

$$\hat{\mu}_W := \frac{1}{n}\left(n_1 X_{(n_1)} + \sum_{i=n_1+1}^{n-n_1} X_{(i)} + n_1 X_{(n-n_1)}\right). \tag{11.2}$$

Ergebnisse der Simulationen

In den Simulationen werden normalverteilte Zufallszahlen mit $\mu = 0$ und $\sigma = 1$ benutzt. In den Tabellen werden die Werte für die Schätzungen und den Bias für $\hat{\mu}$ mit einem Faktor 10^3 multipliziert. Bei den beiden ersten Beispielen $(n = 10, n = 30)$ wurde der komplette Datensatz verwendet

	Est	Bias	Var	MSE	Eff
$\hat{\mu}_{fQ}$	-2.586	0.0067	0.1023	0.1024	97.703
$\hat{\mu}_B$	-3.141	0.0099	0.0983	0.0983	101.696
$\hat{\mu}_G$	-3.199	0.0102	0.0983	0.0983	101.694
$\hat{\mu}_W$	-0.201	0.0000	0.1396	0.1396	71.651
$\hat{\sigma}_{fQ}$	0.995	0.0000	0.0803	0.0804	70.939
$\hat{\sigma}_B$	1.000	0.0000	0.0569	0.0569	100.249
$\hat{\sigma}_G$	1.000	0.000	0.0579	0.0579	98.525
$\hat{\sigma}_{ML}$	1.003	0.0000	0.0565	0.0565	100.949

Tabelle 11.1: Vergleich der Schätzer für die Normalverteilung $(n = 10)$

	Est	Bias	Var	MSE	Eff
$\hat{\mu}_{fQ}$	2.408	0.0058	0.0335	0.0335	99.479
$\hat{\mu}_B$	2.789	0.0078	0.0332	0.0333	100.235
$\hat{\mu}_G$	2.816	0.0079	0.0333	0.0333	100.225
$\hat{\mu}_W$	1.152	0.0013	0.0380	0.0380	87.708
$\hat{\sigma}_{fQ}$	0.997	0.0000	0.0195	0.0195	89.154
$\hat{\sigma}_B$	1.001	0.0000	0.0177	0.0177	98.127
$\hat{\sigma}_G$	1.002	0.0000	0.0177	0.0177	98.245
$\hat{\sigma}_{ML}$	0.999	0.0000	0.0175	0.0175	99.166

Tabelle 11.2: Vergleich der Schätzer für die Normalverteilung $(n = 30)$

	Est	Bias	Var	MSE	Eff
$\hat{\mu}_{fQ}$	-0.022	0.0000	0.0102	0.0102	97.934
$\hat{\mu}_B$	-0.036	0.0000	0.0102	0.0102	97.932
$\hat{\mu}_G$	-0.105	0.0000	0.0124	0.0124	80.495
$\hat{\mu}_W$	-0.032	0.0000	0.0102	0.0102	97.961
$\hat{\sigma}_{fQ}$	1.000	0.0000	0.0061	0.0061	82.693*
$\hat{\sigma}_B$	1.001	0.0000	0.0061	0.0061	82.435*
$\hat{\sigma}_G$	1.000	0.0000	0.0065	0.0065	77.711*
$\hat{\sigma}_{ML}$	1.000	0.0000	0.0051	0.0051	98.599

Tabelle 11.3: Vergleich der Schätzer für die Normalverteilung ($n = 100, n_1 = 5$)

Beim dritten Beispiel ($n = 100$) wurde eine Zensierung ($n_1 = 5$) durchgeführt. Die Effizienzangaben für $\hat{\sigma}_{fQ}$, $\hat{\sigma}_B$ und $\hat{\sigma}_G$ beziehen sich auf den Vergleich mit $Var(\hat{\sigma}_{ML})$ für den *unzensierten* Fall und sind nur für einen relativen Vergleich dieser Schätzer untereinander geeignet.

Aus diesen Tabellen lassen sich folgende Resultate ableiten:

- Der Schätzer $\hat{\mu}_B$ ist bei der unzensierten Schätzung praktisch äquivalent zu $\bar{X} = \hat{\mu}_G$.

- Die speziell für den Fall der Normalverteilung optimierten Schätzer $\hat{\sigma}_B$ und $\hat{\sigma}_G$ haben in der Tat eine hohe Effizienz im Vergleich zu $\hat{\sigma}_{ML}$.

- Bei der Schätzung von μ im unzensierten Fall ist der fQ-Schätzer dem winsorisierten Mittel bei kleinen Stichproben überlegen.

- Die Effizienz von $\hat{\sigma}_{fQ}$ ist erst für $n = 30$ in einem akzeptablen Bereich.

- Die Stärke des fQ-Schätzers liegt im Bereich der zensierten Schätzung, da der MSE in diesem Fall nicht schlechter ist als der von $\hat{\sigma}_B$ und $\hat{\sigma}_G$, der fQ-Schätzer aber wesentlich einfacher zu berechnen ist.

11.3 Cauchyverteilung

Im Gegensatz zur Normalverteilung ist die Schätzung von μ und σ im Falle der Cauchyverteilung durch zwei wesentliche Punkte eingeschränkt:

- Aufgrund der fehlenden Momente der Cauchyverteilung entfällt die Methodik der Momentenschätzung grundsätzlich.
- Die ML-Schätzung ist mit der technischen Schwierigkeit verbunden, daß die Likelihoodfunktion multimodal sein kann (Johnson, Kotz und Balakrishnan [1994, S. 311]).

Zwangsläufig bieten sich somit die Verfahren der L-Schätzung an, die in diesem Fall die einzige, praktisch realisierbare Alternative darstellen.
Folgende Schätzer werden betrachtet:

- Der eingeschränkte fQ-ABLU-Schätzer mit Indexgrenzen $n_1 := \lceil n \cdot 0.129 \rceil$ und $n_2 := \lfloor n \cdot 0.871 \rfloor$ gemäß der Approximation 9.8.
- Die eingeschränkten Schätzer von Gupta $\hat{\mu}_G$ und $\hat{\sigma}_G$.
- Die Schätzer von Blom $\hat{\mu}_B$ und $\hat{\sigma}_B$, die mit den Schätzungen

$$\hat{a}_i := \tan\left(\pi\left(\frac{i-1.23}{n-1.46} - \frac{1}{2}\right)\right),$$

$$\hat{b}_{i,j} := \frac{1}{n}\frac{p_i(1-p_j)}{f_i f_j},$$

  für $p_i = i/(n+1), i = 3, \ldots, n-2$ nach Satz 7.7 konstruiert werden.

  Die Werte f_i werden dabei mit der exakten fQ-Funktion der Cauchyverteilung

$$f_i := fQ_{CV}(p_i) = \pi^{-1}(\sin \pi p_i)^2$$

  bestimmt.

- Ein Quickschätzer der Form

$$\hat{\mu}_Q := \frac{Q_n(p^*) + Q_n(1-p^*)}{2},$$

wobei Q_n die empirische Quantilfunktion bezeichnet. Speziell wird hier der Wert $p^* = 0.4435$ gewählt, der nach Johnson, Kotz und Balakrishnan [1994, S. 307] unter allen Schätzern dieser Form eine minimale Varianz hat.

- Der Schätzer von Weiss und Howlader (1988)

$$\hat{\sigma}_W := -\frac{1}{r}\sum_{i \in I_n} \tan(\pi \tilde{p}_i) X_{(i)},$$

mit

$$I_n := \begin{cases} \left\{3, 4, \ldots, \frac{n-1}{2}, \frac{n+3}{2}, \ldots, n-2\right\} & , \text{ n ungerade} \\ \{3, 4, \ldots, n-2\} & , \text{ n gerade} \end{cases},$$

$$r := \# I_n \quad \text{und} \quad \tilde{p}_i := \begin{cases} \frac{i-2+1/2}{n-1} & , \ i \leq \frac{n}{2} \\ 1 - \tilde{p}_{n-i+1} & , \ sonst \end{cases}.$$

Für Effizienzaussagen werden die Rao-Cramér-Schranken für die unverzerrten Schätzer herangezogen, die nach Chan (1970) durch

$$Var(\hat{\mu}) = \frac{2\sigma^2}{n}, \quad Var(\hat{\sigma}) = \frac{2\sigma^2}{n} \tag{11.3}$$

gegeben sind.

Ergebnisse der Simulationen

In den Beispielen werden jeweils cauchyverteilte Zufallszahlen mit $\mu = 0$ und $\sigma = 10$ verwendet, da die Problematik bei der Schätzung von σ bei größeren Werten vermutlich stärker hervortritt.
Für Datensätze mit Stichprobenumfang von $n = 20$ und $n = 50$ ergeben sich folgenden Ergebnisse:

	Est	Bias	Var	MSE	Eff
$\hat{\mu}_{fQ}$	0.001	0.00000	12.7644	12.7645	78.343
$\hat{\mu}_B$	0.011	0.00001	15.4763	15.4764	64.615
$\hat{\mu}_G$	-0.028	0.00008	41.8107	41.8115	23.917
$\hat{\mu}_Q$	-0.012	0.00001	13.5034	13.5035	74.055
$\hat{\sigma}_{fQ}$	10.086	0.0074	14.1757	14.1831	70.507
$\hat{\sigma}_B$	9.644	0.1268	21.9662	22.0930	45.263
$\hat{\sigma}_G$	10.041	0.0017	28.2876	28.2893	35.350
$\hat{\sigma}_W$	7.860	4.5811	9.0898	13.6710	73.148

Tabelle 11.4: Vergleich der Schätzer für die Cauchyverteilung $(n = 20)$

	Est	Bias	Var	MSE	Eff
$\hat{\mu}_{fQ}$	-0.006	0.0000	4.4650	4.4651	89.584
$\hat{\mu}_B$	-0.005	0.0000	4.8654	4.8654	82.213
$\hat{\mu}_G$	-0.032	0.0011	37.9880	37.9891	10.529
$\hat{\mu}_Q$	-0.016	0.00003	5.0317	5.0320	79.491
$\hat{\sigma}_{fQ}$	10.043	0.0019	4.6375	4.6394	86.219
$\hat{\sigma}_B$	9.732	0.0717	7.3329	7.4046	54.020
$\hat{\sigma}_G$	10.048	0.0023	21.7487	21.7510	18.390
$\hat{\sigma}_W$	8.842	1.3416	3.8438	5.1854	77.140

Tabelle 11.5: Vergleich der Schätzer für die Cauchyverteilung $(n = 50)$

Folgende Aussagen lassen sich hier festhalten:

- Der Schätzer von Gupta ist ungeeignet, da die Annahme $Cov(\mathbf{Y}) = I$ bei der Konstruktion zu einem prinzipiellen Fehler führt.
- Der Schätzer von Blom $\hat{\mu}_B$ ist erst bei einem größem Stichprobenumfang dem einfacher anzuwendenden Quickschätzer überlegen.
- Der Schätzer von Weiss und Howlader führt zwar zu einem größeren Bias, hat demgegenüber aber die geringste Varianz.
- Die fQ-Schätzer liefern bei der Schätzung von μ und σ die besten Resultate.

11.4 Extremwertverteilung

Ähnlich wie bei der Cauchyverteilung ergeben sich hier prinzipielle Probleme:

- Die Schätzung des Skalenparameters σ führt häufig zu unbefriedigenden Resultaten.
- Die ML-Schätzer können nur durch aufwendige iterative Verfahren bestimmt werden.

Folgende Schätzer werden betrachtet:

- Der fQ-ABLU-Schätzer mit den Indexgrenzen $n_1 := \lceil n \cdot 0.03 \rceil$ und $n_2 := \lfloor n \cdot 0.894 \rfloor$ gemäß der Approximation (9.9).
- Die Schätzer von Gupta $\hat{\mu}_G$ nd $\hat{\sigma}_G$.
- Die ABLU-Schätzer von Blom $\hat{\mu}_B$ und $\hat{\sigma}_B$ mit

  $$\hat{a}_i := Q_{EV}\left(\frac{i-1/2}{n}\right), \quad \hat{b}_{i,j} := \frac{1}{n}\frac{p_i(1-p_j)}{f_i f_j},$$

  für $p_i = i/(n+1), i = 1, \ldots, n$ und

  $$f_i := fQ_{EV}(p_i) = (1-p_i)\log\left(1/(1-p_i)\right).$$

- Die Momentenschätzer

  $$\hat{\sigma}_M := \frac{\sqrt{6}}{\pi}S, \qquad (11.4)$$
  $$\hat{\mu}_M := \bar{X} - \gamma\hat{\sigma}_M, \qquad (11.5)$$

  wobei $\gamma \approx 0.57722$ die Eulersche Konstante bezeichnet.

- Die approximierten ML-Schätzer von Kimball (1956)

  $$\hat{\sigma}_K := \frac{\bar{X} + \frac{1}{n}\sum_{i=1}^n X_{(i)}\left(\sum_{j=i}^n \frac{1}{j}\right)}{1+2.3n^{-1}}, \qquad (11.6)$$
  $$\hat{\mu}_K := \bar{X} - \gamma\hat{\sigma}_K \qquad (11.7)$$

Nach Downton (1966) sind die Rao-Cramér-Schranken für die unverzerrten Parameterschätzungen durch die Grenzen

$$Var(\hat{\mu}) = \left(1 + 6(1-\gamma)^2\pi^{-2}\right)\frac{\sigma^2}{n} \approx 1.10867 \cdot \frac{\sigma^2}{n}, \tag{11.8}$$

$$Var(\hat{\sigma}) = \frac{6}{\pi^2}\frac{\sigma^2}{n} \approx 0.60793 \cdot \frac{\sigma^2}{n}. \tag{11.9}$$

Damit ergeben sich mit den Varianzformeln nach Tiago de Olivera (1963)

$$Var(\hat{\mu}_M) \approx 1.1678\frac{\sigma^2}{n}, \quad Var(\hat{\sigma}_M) \approx 1.1\frac{\sigma^2}{n} \tag{11.10}$$

für die Momentenschätzer eine Effizienz von ca. 95% bzw. 55%.

Ergebnisse der Simulationen

In den Simulationen werden Stichproben mit $\mu = 0$ und $\sigma = 1$ der Längen 20 und 50 erzeugt. In den Tabellen werden die Werte für die Schätzungen und den Bias für $\hat{\mu}$ mit einem Faktor 10^5 multipliziert.

	Est	Bias	Var	MSE	Eff
$\hat{\mu}_{fQ}$	0.000	0.0000	0.0566	0.0566	97.949
$\hat{\mu}_B$	0.026	6.6683	0.0572	0.0579	95.814
$\hat{\mu}_G$	−0.012	1.3999	0.0505	0.0507	109.364
$\hat{\mu}_M$	0.014	1.8404	0.0576	0.0577	96.010
$\hat{\mu}_K$	0.112	126.5433	0.0593	0.0719	77.090
$\hat{\sigma}_{fQ}$	0.999	0.0054	0.0381	0.0382	79.676
$\hat{\sigma}_B$	0.976	5.9170	0.0365	0.0371	81.896
$\hat{\sigma}_G$	1.000	0.0017	0.0412	0.0412	73.801
$\hat{\sigma}_M$	0.977	5.2977	0.0493	0.0498	61.035
$\hat{\sigma}_K$	0.806	377.9087	0.0228	0.0606	50.164

Tabelle 11.6: Vergleich der Schätzer für die Extremwertverteilung ($n = 20$)

	Est	Bias	Var	MSE	Eff
$\hat{\mu}_{fQ}$	0.000	0.0007	0.0223	0.0223	99.307
$\hat{\mu}_B$	0.010	9.8215	0.0224	0.0225	98.616
$\hat{\mu}_G$	−0.006	4.1317	0.0215	0.0216	102.787
$\hat{\mu}_M$	0.006	3.7304	0.0228	0.0229	96.918
$\hat{\mu}_K$	0.050	247.1191	0.0227	0.0252	87.929
$\hat{\sigma}_{fQ}$	1.001	0.0571	0.0145	0.0145	83.670
$\hat{\sigma}_B$	0.993	5.3315	0.0143	0.0143	84.735
$\hat{\sigma}_G$	0.999	0.1184	0.0167	0.0167	72.769
$\hat{\sigma}_M$	0.987	15.7306	0.0208	0.0210	57.934
$\hat{\sigma}_K$	0.912	775.8533	0.0111	0.0189	64.429

Tabelle 11.7: Vergleich der Schätzer für die Extremwertverteilung ($n = 50$)

Die Tabellen lassen diese Folgerungen zu:

- Mit Ausnahme des Schätzers von Kimball liefern alle Schätzer gute Ergebnisse für μ.
- Die Schätzer von Kimball weisen in beiden Fällen eine große Verzerrung auf und sind aus diesem Grunde als ungeeignet anzusehen.
- Die theoretische Effizienzaussage (11.10) für die $\hat{\mu}_M$ wurde bei kleinem Stichprobenumfang deutlich übertroffen.
- Der beste Schätzer für μ ist $\hat{\mu}_G$; dies ist überraschend, da der Schätzer von Gupta keinerlei Informationen über die Kovarianzen der Erwartungswerte dieser asymmetrischen Verteilung nutzt.
- Die Effizienzen für die Schätzer von Blom und die fQ-Schätzer sind vergleichbar und dominieren im Falle der Schätzung von σ die anderen Verfahren.
- Alle Verfahren, die eine bessere Effizienz als der fQ-Schätzer haben, weisen jeweils einen höhreren Bias auf.

Zusammenfassung

Mit Ausnahme der Schätzung bei der Normalverteilung mit geringem Stichprobenumfang und dem überraschenden Resultat des Schätzers von Gupta für die Extremwertverteilung liefert die fQ-Methode stets Ergebnisse, die sich mit den jeweils besten Schätzern messen können. Im Falle der Cauchyverteilung scheint es für den Problemfall der Schätzung von σ bei großem Stichprobenumfang den bekannten Verfahren überlegen zu sein.

Der Vorteil der fQ-Schätzung liegt darin, daß zur Konstruktion der Schätzer lediglich die Kenntnis der fQ-Parameter der betreffenden Verteilung erforderlich ist, im Gegensatz zu den umfangreichen Tabellenwerken für die bekannten BLU- und ABLU-Verfahren.

Kapitel 12

Dichteschätzung unimodaler Verteilungen

Die nichtparametrische Schätzung von Dichtefunktionen stetiger Verteilungen nimmt in der statistischen Literatur einen großen Raum ein, wobei der Methode der Kernschätzung die Schlüsselrolle zukommt. Die Grundlagen dieser Technik gehen auf Rosenblatt (1956) und Parzen (1962) zurück, die den Kernschätzer für einen Datensatz $\mathbf{x} = (x_1, \ldots, x_n)$ in der Form

$$\hat{f}_b(x) := \frac{1}{b} \sum_{i=1}^{n} K\left(\frac{x - x_i}{b}\right) \tag{12.1}$$

definieren, wobei die Funktion K (die *Kernfunktion* oder der *Kern*) eine stetige, beschränkte und symmetrische Funktion mit

$$\int K(u)\, du = 1$$

ist. Der Skalenparameter b wird die *Bandbreite* genannt und bestimmt den Grad der Glättung (siehe auch die folgenden Beispiele).

Grundsätzlich stellt die Kernschätzung ein *lokales* Verfahren dar, da der Ausdruck $\hat{f}_b(x)$ einer gewichteten Summe der Abstände $x - x_j$ entspricht, wobei das Maß für die Lokalität gerade durch den Parameter b bestimmt wird. Das Ergebnis einer Kernschätzung ist somit eine *Punktschätzung* für den Wert der zugrundeliegenden Dichte an einer vorgegebenen Stelle x. Demgegenüber ist das Resultat einer fQ-Dichteschätzung die Approximation der Dichtefunktion in einer parametrischen Darstellung, die in weiteren Analysen in geschlossener Form verwendet werden kann.

Bevor anhand konkreter Beispiele diese beiden Schätzverfahren miteinander verglichen werden, seien in einer kurzen Übersicht einige Aussagen über die Kernschätzung zusammengefaßt:

- Für die Kernfunktion K gibt es verschiedene Vorschläge (siehe z.B. Silverman [1986] oder Härdle [1991]). Die bekanntesten sind dabei der Rechteck-Kern, der Dreiecks-Kern, der Gauß-Kern, der Parzen-Kern und der Epanechnikow-Kern, der gewissen Optimalitätskriterien genügt (Härdle [1991, S. 138]).

- Die Schätzung ist extrem abhängig von der Wahl der Bandbreite b: bei einem kleinen Wert ist die Schätzung multimodal (im Grenzfall $b \to 0$ werden die Originaldaten reproduziert), für große Werte von b wird die geschätzte Dichte sehr glatt.

  Mathematisch bedeutet dies, daß für $b \to \infty$ der Bias der Schätzung steigt während gleichzeitig die Varianz sinkt (für $b \to 0$ kehren sich diese Beziehungen um).

- Für die Wahl von b gibt es verschiedenste empirische Aussagen, siehe zum Beispiel Silverman (1986) oder Scott (1992)

$$\hat{b}_1 = 1.06 \cdot \min\left(\hat{\sigma}, \frac{Q_n(0.75) - Q_n(2.25)}{1.34}\right) \cdot n^{-1/5} \tag{12.2}$$

$$\hat{b}_2 = 1.144 \cdot \hat{\sigma} \cdot n^{-1/5}. \tag{12.3}$$

- Für eine automatische Wahl von b werden einerseits Optimalitätskriterien wie der *Mean Integrated Square Error* (MISE) in der Form (Scott [1992])

$$MISE = \int E\left((\hat{f}_b(x) - f(x))^2\right) dx$$

  oder rechnerische Methoden, wie die *Cross validation* (Stone [1974]) vorgeschlagen.

Letztendlich kann die Optimalität einer Kernschätzung jedoch nur anhand einer visuellen Prüfung der Anpassungsqualität beurteilt werden.

Beispiele

Anhand einiger Beispiele soll ein optischer Eindruck von der Qualität der fQ-Dichteschätzung im Vergleich zur Kernschätzung vermittelt werden. In den folgenden Diagrammen wird soweit nicht anders angegeben die exakte Dichtefunktion (durchgezogene Linie), die Kernschätzung (gestrichelte Linie) sowie die fQ-Dichteschätzung (Strich-Punkt-Linie) abgebildet.
Im ersten Beispiel werden zunächst für eine normalverteilte Stichprobe mit $n = 100$ einige der oben beschriebenen Varianten bei der Kernschätzung angewendet.

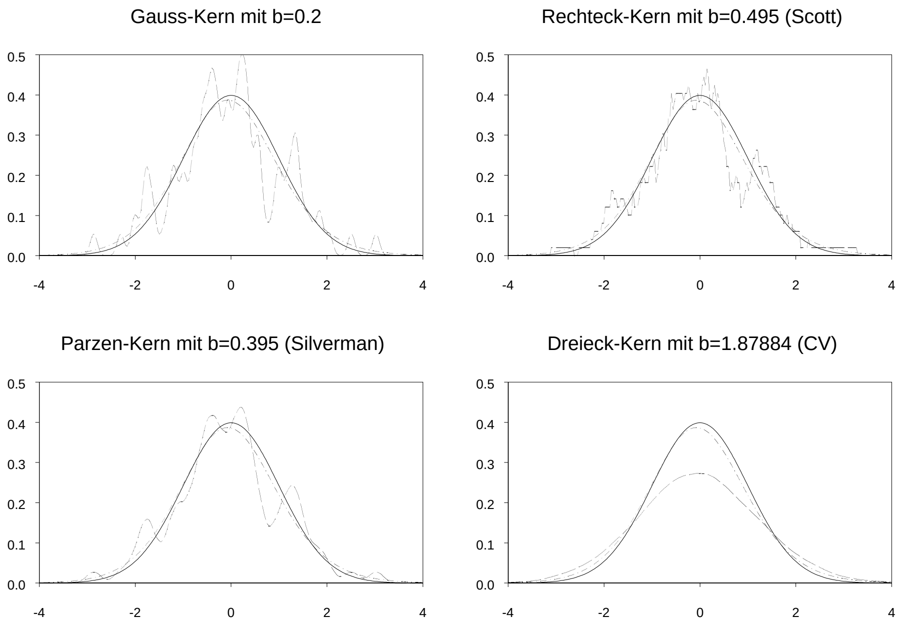

Abbildung 12.1: Verschiedene Kernschätzungen der NV ($n = 100$)

Weder mit der frei gewählten Bandbreite von $b = 0.2$ noch mit den nach den Regeln von Scott und Silverman oder durch Cross validation automatisch bestimmten Werten wird die tatsächliche Dichte angemessen beschreiben. Beim Vergleich der unterschiedlichen Kerne fällt die deutlich gröbere Struktur der resultierenden Kurve für den Rechteck-Kern auf.

Um einen standardisierten Vergleich durchführen zu können, wird bei allen weiteren Kernschätzungen der Gauß-Kern verwendet und die Bandbreite durch einen optischen Vergleich so bestimmt, daß sie die tatsächliche Dichte möglichst gut approximiert. Die Kernschätzungen in den folgenden Abbildungen sind also als das Resultat von manuellen Anpassungsprozesssen zu verstehen.
Die Einteilung der ausgewählten Beispiele orientiert sich an der Symmetrie und an dem Tailverhalten der Dichten und beginnt mit dem symmetrischen Fall mit einem Stichprobenumfang von $n = 300$ und den Verteilungen

- Normalverteilung $NV(0,1)$
- Logistische Verteilung $LV(0,1)$
- Trapez-Verteilung $Tr(0,3,1/3,2/3)$
- Rechteck-Verteilung $U(0,2)$.

In allen Diagrammen wird zusätzlich die gewählte Bandbreite bei der jeweiligen Kernschätzung angegeben.

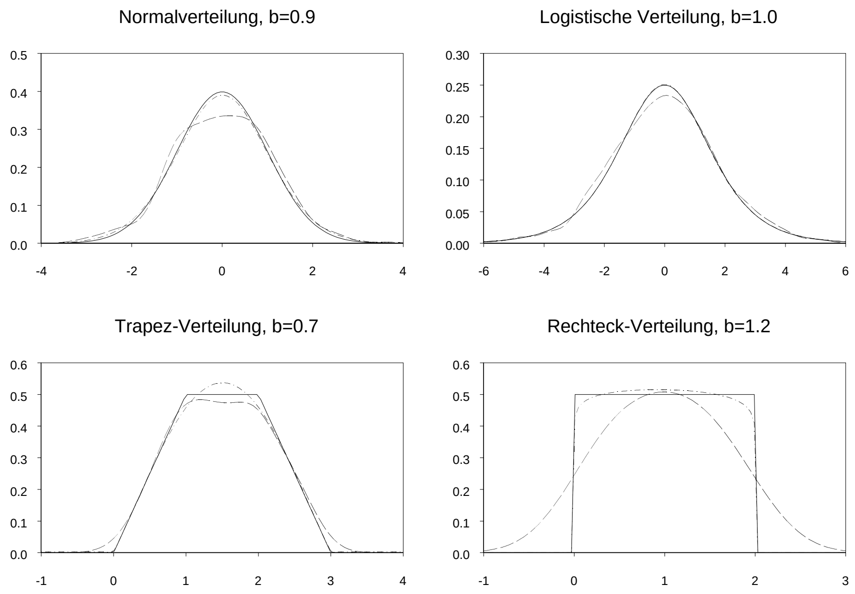

Abbildung 12.2: Dichteschätzung für symmetrische Verteilungen ($n = 300$)

Auffallend ist hier die gute Annäherung der fQ-Dichteschätzung an die Gestalt der exakten Dichtefunktion, wobei selbst die schwach unimodale Rechteck-Verteilung gut approximiert wird.
Für die Untersuchung von symmetrischen Verteilungen mit großen Tails werden die Cauchyverteilung und die t-Verteilung mit 3 Freiheitsgraden herangezogen. Für eine Stichprobe vom Umfang $n = 300$ ergibt sich folgendes Bild, wobei zur Verdeutlichung zusätzlich zu den geschätzten Dichten in einem zweiten Diagramm die Abstände der Schätzungen zu der exakten Dichte abgebildet sind.

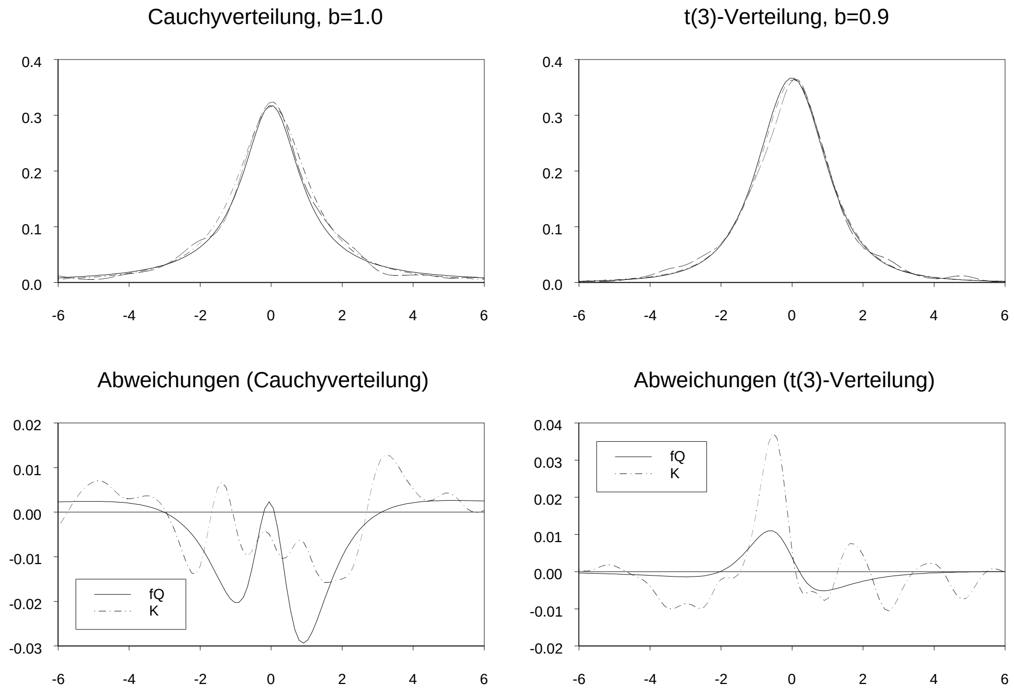

Abbildung 12.3: Dichteschätzung für Verteilungen mit großen Tails

In diesem Beispiel zeigt es sich, daß beide Schätzungen gut geeignet sind, die korrekte Dichte anzunähern, da die Größenordnungen der Differenzen in beiden Fällen gering sind.

In der nächsten Abbildung werden die Schätzungen für schiefe Verteilungen gezeigt, wobei bei einem Stichprobenumfang von $n = 300$ folgende Verteilungen betrachtet werden:

- Beta-Verteilung $B(4, 1.5)$
- F-Verteilung $F(4, 10)$
- Log-Normalverteilung $LNV(1, 1/2)$
- Exponentialverteilung $E(1)$.

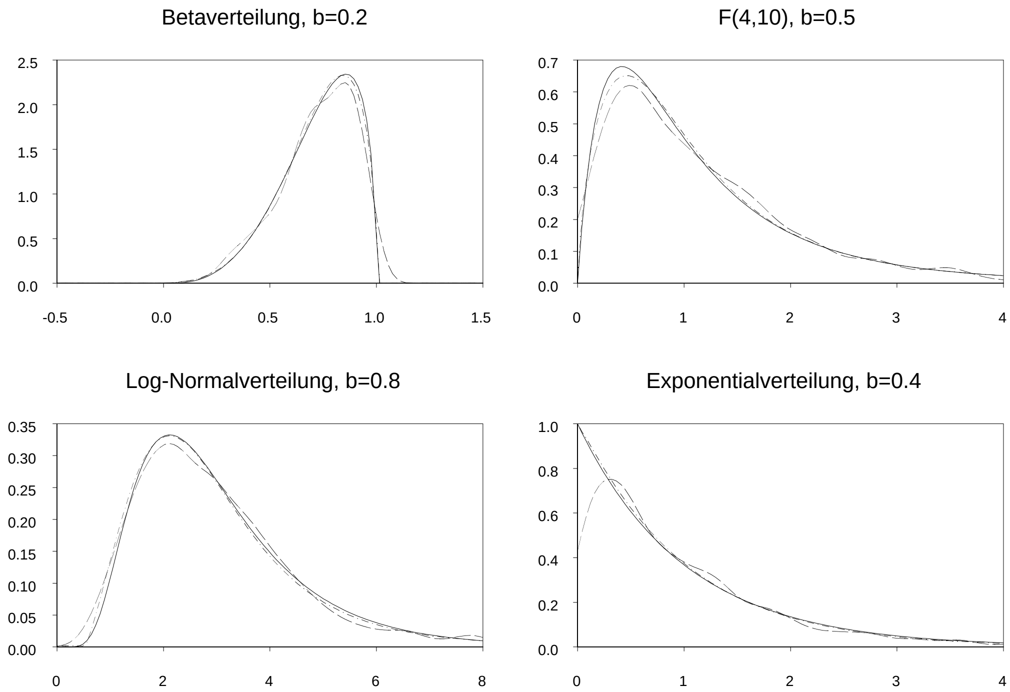

Abbildung 12.4: Dichteschätzung für schiefe Verteilungen

Erneut ist zu beobachten, daß sowohl hinsichtlich der lokalen als auch der globalen Anpassungsgüte die fQ-Schätzung der Kernschätzung überlegen ist. Insbesondere zeigt das Beispiel der Exponentialverteilung eine prinzipielle Schwäche der Kernschätzung bei Verteilungen mit endlichem Träger.

Im abschließenden Beispiel werden die Stichproben von geringem Umfang ($n = 30, 50$) betrachtet und dabei speziell die Standardnormalverteilung und die F-Verteilung mit Parametern $(4, 10)$ untersucht:

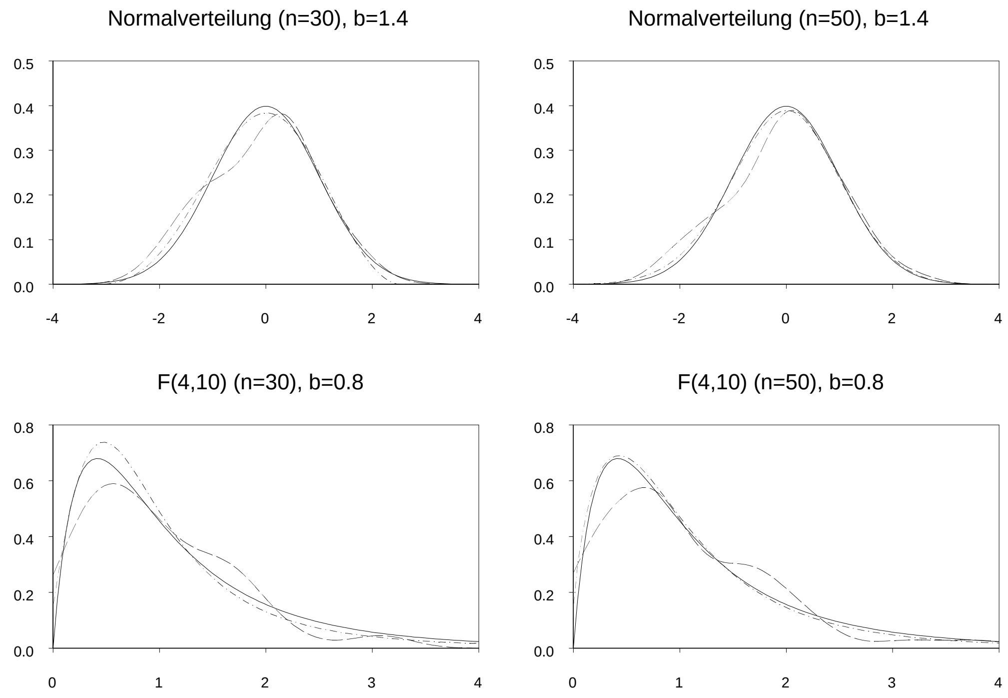

Abbildung 12.5: Dichteschätzung bei kleinem Stichprobenumfang ($n = 30, 50$)

In keinem Fall gelang es, durch eine Variation der Bandbreite die lokalen Schwingungen der Kernschätzung angemessen zu glätten. Dies liegt daran, daß gerade bei geringem Stichprobenumfang der Einfluß einer einzelnen Beobachtung bei der Berechnung der Kernschätzung an einem Punkt x groß ist, wodurch es in Bereichen auf der x-Achse, in denen ungewöhnlich viele Beobachtungen auftreten, zwangsläufig zu Ausschlägen der Dichteschätzung kommt.

Zusammenfassung

Die betrachteten Beispiele lassen den Schluß zu, daß in der Situation einer vermuteten zugrundeliegenden stetigen und unimodalen Verteilung die fQ-Methode ein sinnvolles Verfahren zur Dichteschätzung darstellt. Gegenüber der Kernschätzung weist sie folgende Vorteile auf:

- Es muß keine Kernfunktion bestimmt werden und es ist nicht erforderlich, ein empirisches oder analytisches Verfahren zur Bestimmung der Bandbreite durchzuführen. Die fQ-Dichteschätzung ist in diesem Sinne also als ein nichtparametrisches Glättungsverfahren anzusehen.

- Die angepaßte Kurve schätzt den *globalen* Verlauf der Dichte im Gegensatz zur prinzipiellen Lokalität der Kernschätzung und liefert demzufolge optisch bessere Anpassungen.

- Die Schätzung führt auch bei schwach unimodalen Verteilungen und bei Verteilungen mit endlichem Träger zu akzeptablen Resultaten.

- Auch bei geringem Stichprobenumfang kann die tatsächliche Dichte gut angenähert werden.

Kapitel 13

Zusammenfassung und Ausblick

Die Grundidee des hier eingeführten fQ-System besteht darin, unimodale stetige Verteilungen anhand optischer Merkmale in analytischer Form zu beschreiben. Die Untersuchungen im theoretischen Teil bestätigen, daß die gewählte funktionale Form in der Tat geeignet ist, mit Hilfe der charakterisierenden Parameter unmittelbare Aussagen über globale Merkmale der Verteilungen wie zum Beispiel die Existenz von Momenten oder die Endlichkeit des Trägers treffen zu können. Es zeigt sich weiter, daß die vorgestellten Ordnungsbegriffe und Kennmaße in konsistenter Weise miteinander verbunden und zudem durch einfache parametrische Bedingungen überprüfbar sind. Mit Hilfe des neuen Begriffs der fQ-Metrik läßt sich die Definition der fQ-Darstellung bekannter unimodaler Verteilungen leicht formulieren.

Als Ergebnis der Approximation einiger ausgewählter Verteilungen läßt sich festhalten, daß die Qualität der Anpassung innerhalb definierter Grenzen zu nutzbaren Ergebnissen führt. Dies kann zunächst direkt verwendet werden, um mit Hilfe der fQ-Anpassung gegebener Verteilungen Schätzungen für die Momente der Ordnungsstatistiken zu erlangen.

Im einem weiteren Schritt können diese Schätzungen genutzt werden, um mit ihnen die bekannte Problematik bei der Konstruktion der besten linearen Schätzer zu umgehen, die in der aufwendigen Bestimmung oder unzureichenden Schätzung der Momente der Ordnungsstatistiken der bekannten Grundverteilug besteht. Die Simulationsstudien im praktischen Teil zeigen, daß der Fehler, der durch die Verwendung der fQ-Approximation anstelle der exakten fQ-Funktion begangen wird,

die Effizienz der resultierenden Schätzer nicht beeinträchtigt. Der praktische Vorteil der fQ-Methode liegt in der Tatsache, daß zur Konstruktion der Schätzer lediglich die fQ-Parameter der Grundverteilung sowie ein Verfahren zur Bestimmung der fQ-Quantilfunktion bekannt sein müssen.
Wie die praktischen Versuche im Umgang mit dem fQ-System zeigen, können dabei die numerischen Probleme bei der Berechnung der Dichte-, Quantil- und Verteilungsfunktion dank der vorgestellten numerischen Verfahren als gelöst betrachtet werden.
Für die Praxis der Dichteschätzung unimodaler Verteilungen kann die Methode der fQ-Dichteschätzung eine sinnvolle Alternative zur klassischen Kernschätzung sein, da die Resultate der Vergleichsstudie darauf hindeuten, daß einige der bekannten Probleme der Kernschätzung bei der fQ-Schätzung vermieden werden. Insbesondere ist hier die Tatsache von Bedeutung, daß zur Durchführung der fQ-Schätzung keinerlei zusätzliche Vorüberlegungen wie die Bestimmung einer optimalen Bandbreite notwendig sind.

Ausblick

Aufgrund der Bedeutung der fQ-Funktion und der Scorefunktion vor allem im Bereich der nichtparametrischen Verfahren sind dort weitere praktische Anwendungen denkbar. Da durch die Anpassung einer fQ-Funktion an empirische Daten zugleich eine Schätzung der Scorefunktion gegeben ist, ergibt sich damit zum Beispiel die Möglichkeit, mit Hilfe dieser geschätzten Scorefunktion $\hat{J}$ unter Verwendung der Ergebnisse von Hájek und Šidák (1967) approximative, lokal optimale lineare Rangstatistiken zu konstruieren. Anhand umfangreicher Simulationen wäre zu untersuchen, ob mit dieser *feinadaptiven* Methode die gängigen *grobadaptiven* Testverfahren verbessert werden können.

Symbolverzeichnis

$\alpha(f)$	Untere Grenze des Trägers einer Verteilung F mit Dichte f
$\alpha_{\hat{F}}^{(s)}$	Untere Grenze des s-Approximationsintervalls
$B(a,b)$	Vollständige Betafunktion
$\mathbf{C}_{L_2}([\,0,1])$	Raum der quadratintegrierbaren Funktionen auf $[\,0,1]$
D_f	Träger einer Verteilung einer Verteilung F mit Dichte f
$D_{\hat{F}}^{(s)}$	Approximationsmenge (bzw. -intervall)
d_J	Metrik von Hájek und Šidák
d_{fQ}	fQ-Metrik
$f(x_{mod})$	Modalwert einer Dichte
fQ, fQ_F	Dichte-Quantil-Funktion (einer Verteilung F)
$\mathcal{F}(\mu,\sigma,\alpha,\beta)$	fQ-Verteilung
$\mathcal{FQ}$	fQ-System
$\mathcal{FQ}_I$	fQ-Verteilungen mit endlicher Fisher-Information
$\mathcal{FQ}_U$	Unimodale fQ-Verteilungen
$\mathcal{FQ}_{WU}$	Schwach unimodale fQ-Verteilungen
$\mathcal{F}_S$	Menge der stetigen Verteilungen
$\mathcal{F}_U$	Menge der stetigen unimodalen Verteilungen
$\mathcal{F}_I$	Menge der stetigen Verteilungen mit endlicher Fisher-Information
$\mathcal{F}_{WU}$	Menge der stetigen, schwach unimodalen Verteilungen
γ_F	Exponentfunktion einer Verteilung
H_2	Schätzer für die Peakedness von Hogg
$I(F)$	Fisher-Information
J, J_F	Score-Funktion (einer Verteilung F)
J^*, J_F^*	Normierte Score-Funktion (einer Verteilung F)

$\mu(a,b)$	Verallgemeinerter Mittelwert
μ_D	Dispersionsmaß
μ_D^F	Familiendispersion
μ_D^S	Skalendispersion
μ_S	Schiefemaß
μ_{appr}	ABLU-Lokationsschätzer
μ_B	ABLU-Lokationsschätzer nach Blom
μ_G	ABLU-Lokationsschätzer nach Gupta
μ_{opt}	BLU-Lokationsschätzer
$\omega(f)$	Obere Grenze des Trägers einer Verteilung
$\omega_{\hat{F}}^{(s)}$	Obere Grenze des s-Approximationsintervalls
Q, Q_F	Quantil-Funktion (einer Verteilung F)
$\Psi_{F,G}$	Transferfunktion
ρ_J	Ähnlichkeitsmaß von Hájek und Šidák
σ_{appr}	ABLU-Skalenschätzer
σ_B	ABLU-Skalenschätzer nach Blom
σ_G	ABLU-Skalenschätzer nach Gupta
σ_{opt}	BLU-Skalenschätzer
τ	Funktion zur Umparametrisierung einer fQ-Verteilung
$T(a,b)$	Schätzer für die Tailstärke
x_{mod}	Modus einer Dichte
$\overset{d}{<}, \overset{d}{\sim}, \overset{d}{>}$	Dispersionsordnung
$\overset{k}{<}, \overset{k}{\sim}$	Konvexordnung
$\overset{p}{<}, \overset{p}{\sim}$	Tailordnung
$<f,g>$	Skalarprodukt auf $\mathbf{C}_{L_2}([0,1])$
$\|f\|$	Norm auf $\mathbf{C}_{L_2}([0,1])$

Literatur

Alzaid, A. A. and M. Al-Osh (1989). Ordering probability distributions by tail behavior. *Stat. and Prob. Let. 8*(2), 185–188.

Arnold, B. C., N. Balakrishnan, and H. Nagarja (1992). *A first course in order statistics.* New York: Wiley.

Balakrishnan, N. and A. C. Cohen (1991). *Order Statistics and inference.* London: Academic press.

Benett, C. A. (1952). *Asymptotic properties of ideal linear estimators.* Ph. D. thesis, University of Michigan.

Björck, A. and G. Dahlquist (1972). *Numerische Methoden.* München, Wien: Oldenbourg.

Blom, G. (1958). *Statistical Estimates and Transformed Beta-Variables.* New York: Wiley.

Bronstein, I. N. and K. A. Semendjajew (1984a). *Taschenbuch der Mathematik.* Thun und Frankfurt/Main: Verlag Harri Deutsch.

Bronstein, I. N. and K. A. Semendjajew (1984b). *Taschenbuch der Mathematik, Ergänzungen.* Thun und Frankfurt/Main: Verlag Harri Deutsch.

Brown, G. and J. W. Tukey (1948). *Annals of Mathematical Statistics 19*, 76–81.

Büning, H. (1985). *Adaptive verteilungsfreie Tests - Nichtparametrische Maße zur Klassifizierung von Verteilungen.* Neuere Verfahren der nichtparametrischen Statistik. Berlin: Springer.

Büning, H. (1991). *Robuste und adaptive Tests.* Berlin, New York: Walter de Gruyter.

Burden, R. L. and J. D. Faires (1989). *Numerical Analysis.* Boston: PWS-Kent.

Chan, L. K. (1970). Linear estimation of the location and scale parameters of the cauchy distribution based on sample quantiles. *JASA 65*, 851–859.

David, F. N. and N. Johnson (1954). Statistical treatment of censored data. i. fundamental formulae. *Biometrika 12*, 408–412.

David, H. A. (1980). *Order Statistics* (2nd ed.). New York: Wiley.

de Oliveira, J. T. (1963). Decision results for the parameter of the extreme value (gumbel) distribution ased on the mean and the standard deviation. *Trabajos de Estadistica 14*, 61–81.

Dongarra, J. J. and E. Grosse (1987). Distribution of mathematical software via electronic mail. *Communications of the ACM 30*, 403–407.

Downton, F. (1966). Linear estimates of parameters in the extreme value distribution. *Technometrics 8*, 3–17.

Groeneveld, R. and G. Meeden (1984). Measuring skewness and kurtosis. *The Statistician 33*, 391–399.

Gupta, A. K. (1952). Estimation of the mean and standard deviation of a normal population from a censored sample. *Biometrica 39.*

Hajek, J. and Z. Sidak (1967). *Theory of rank tests.* New York: Academic Press.

Handl, A. (1985). *Maßzahlen zur Klassifizierung von Verteilungen bei der Konstruktion adaptiver verteilungsfreier Test.* Ph. D. thesis, Freie Universität Berlin, Berlin.

Härdle, W. (1991). *Smoothing techniques with Implementation in S.* New York: Springer.

Hogg, R. V. (1974). Adaptive robust procedures. a partial review and some suggestions for future applications and theory. *JASA 69*, 909–927.

Hogg, R. V. (1982). On adapative statistical inference. *Commun. Statist., Ser. A, Theor. and Meth. 11*, 2531–2542.

Ibragimov, J. A. (1956). On the composition of unimodal distributions. *Teoriya verayatnostey I*, 283–288.

Johnson, N. L. (1949). Systems of frequency curves generated by methods of translation. *Biometrika 36*, 149–176.

Johnson, N. L., S. Kotz, and N. Balakrishnan (1994). *Continuous univariate distributions* (second ed.), Volume 1. New York: Wiley.

Johnson, N. L., S. Kotz, and N. Balakrishnan (1995). *Continuous univariate distributions* (second ed.), Volume 2. New York: Wiley.

Joiner, B. L. and D. L. Hall (1983). The ubiquitous role of f'/f in efficient estimation of location. *The American Statistician 37*(2), 128–133.

Karian, Z. A., E. J. Dudewicz, and P. McDonald (1996). The extended generalized lambda distribution. *Commun. Statist. Simul. 25*, 611–642.

Kimball, B. F. (1956). The bias in certain estimates of the extreme-value distribution. *Annals of Mathematical Statistics 27*, 758–767.

Lloyd, E. H. (1952). Least-squares estimation of location and scale parameters using order statistics. *Biometrica 39*, 88–95.

Parzen, E. (1962). On estimation of a probability density and mode. *Annals of Mathematical Statistics 35*, 1065–1076.

Parzen, E. (1979a). Density quantile estimation approach to statistical data modelling. In *Smoothing Techniques for Curve Estimation*, pp. 155–180.

Parzen, E. (1979b). Nonparametric statistical data modeling. *JASA 74*, 105–120.

Pearson, K. (1895). Contributions to the mathematical theory of evolution. *Philosophical Transactions of the Royal Society of London, Series A 186*, 343–414.

Ramberg, J. S. and B. Schmeiser (1972). An approximative method for generating symmetric random variables. *Comm. ACM 15*, 987–990.

Ramberg, J. S. and B. Schmeiser (1974). An approximative method for generating asymmetric random variables. *Comm. ACM 17*, 78–82.

Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics 27*, 642–669.

Ryan, T. P. (1997). *Modern Regression Methods.* New York. Wiley.

Sarhan, A. E. and B. Greenberg (1962). *Contributions to Order Statistics.* New York: Wiley.

Scott, D. W. (1992). *Multivariate Density-Estimation, Theory, Practice and Visualation.* New York: Wiley.

Sedgewick, R. (1992). *Algorithmen in C.* Bonn, München: Addison Wesley.

Shaked, M. and J. G. Shanthikumar (1994). *Stochastic Orders and their applications.* London: Academic Press.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis.* New York: Chapman and Hall.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions (with discussion). *Journal of the Royal Statistical Society B 36*, 111–147.

Tukey, J. W. (1962). The future of data analysis. *Annals of Math. Statistics 33*, 1–67.

van Zwet, W. R. (1964). Convex transformations of random variables.

Voit, E. O. (1992). The s-distribution: A tool for approximation and classification of univariate, unimodal probability distributions. *Biometrical Journal 34*, 855–878.

Weiss, G. and H. Howlader (1988). Linear scale estimation. *Journal of statistical Computation and Simulation 29*, 117–126.