



Ruhr-Universität Bochum
Fakultät für Bau- und
Umweltingenieurwissenschaften
Lehrstuhl für Hydrologie, Wasserwirtschaft
und Umwelttechnik
Prof. Dr. rer. nat. habil. A. Schumann



EXANTO

- Extreme Value Analysis Tool -

Dokumentation

Version 1.14

Autor: Dr.-Ing. Bastian Klein
Stand: 13.07.2010

EXANTO Copyright © 2010 Ruhr- University Bochum, Institute of Hydrology and
Water Management. All rights reserved.

Kontakt: Dr.-Ing. Bastian Klein bastian.klein@rub.de

Vorwort

Hochwasserwahrscheinlichkeiten werden für vielfältige wasserwirtschaftliche Bemessungsaufgaben benötigt. Verfahren der mathematischen Statistik bilden dabei die gebräuchlichste Grundlage zur Ermittlung von Wahrscheinlichkeitsaussagen. Sie haben den Vorteil einer direkten Nutzung der Datenbasis in Form gemessener Höchstabflüsse. Eine Bedingtheit der Ergebnisse dieser Analysen von den spezifischen Annahmen der genutzten Verfahren ist jedoch zu beachten. Zwar gibt es für die Anwendung der mathematischen Statistik eine Vielzahl von etablierten Vorgehensweisen, diese setzen jedoch ein grundlegendes Verständnis des Anwenders zu den mathematischen Zusammenhängen und Voraussetzungen voraus. Die hier vorgestellte Software-Lösung macht hiervon keine Ausnahme. Die Vielzahl der vorhandenen Verteilungsfunktionen, Parameterschätzverfahren und Anpassungstests erfordert sowohl ein systematisches Vorgehen als auch eine eigenständige Wertung der Ergebnisse durch den Anwender. Insofern bietet die hier bereitgestellte Software-Lösung eine wesentliche Grundlage für die Anwendung dieser Verfahren, der Anwender ist jedoch gefordert, die Ergebnisse zu bewerten und problemorientiert eine Entscheidung über die zu verwendenden Wahrscheinlichkeitsaussagen zu treffen. Um diese Herausforderung annehmen zu können, werden nachfolgend die wesentlichen statistischen Grundlagen kurz dargestellt. Diese Darstellung kann ein Lehrbuch nicht ersetzen, fokussiert sie doch auf die, hier software-technisch umgesetzten Verfahren und Methoden, aber nicht auf deren Voraussetzungen und Restriktionen.

1. Sie ist freeware, d.h. der Nutzer kann sie ohne weitere Kosten anwenden.
2. Sie ist zwar methodisch umfassend, jedoch eingeschränkt im Hinblick auf die Ergebnispräsentation. Der Anwender muss die Ergebnisse selbst auswerten (z.B. in dem er Grafiken in Excel erstellt). In Hinblick auf die erforderlichen Vorkenntnisse ist dies jedoch kein Makel, fördert die knappe Ergebnisdarstellung doch die Auseinandersetzung mit den Resultaten der Berechnungen.

Abschließend danke ich an dieser Stelle Herrn Dr. Klein, der während seiner Tätigkeit am Lehrstuhl für Hydrologie, Wasserwirtschaft und Umwelttechnik der Ruhr-Universität die Erstellung dieses Softwarepaketes realisierte. Der hierbei gezeigte Fleiß sowie die Geduld und Ausdauer bei der Durchführung zahlreicher Iterationen bis hin zum hier vorgestellten Paket sind nicht hoch genug zu schätzen. Wir wünschen dem Anwender viel Erfolg bei der Nutzung dieses Pakets und würden uns über positive, aber auch kritische Rückmeldungen freuen.

Bochum, den 23.08.2010

Prof. Dr. Andreas Schumann

Inhaltsverzeichnis

I	Lizenzvereinbarung	3
1	Einführung	3
2	EXANTO	7
2.1	Disclaimer.....	7
2.2	Copyright Notice	7
2.3	Systemvoraussetzungen:.....	7
2.4	Ausführen von EXANTO.....	7
2.5	Stichprobe.....	8
2.6	Kontrolldatei.....	9
2.7	Ergebnisdateien.....	11
3	Ermittlung des Bemessungshochwassers auf statistischer Grundlage	13
3.1	Statistische Analyse von Hochwasserabflüssen	13
3.2	Verteilungsfunktionen	14
4	HQ-Werte für kleine Wiederholungszeitspannen	16
5	Parameterschätzung	17
5.1	Herkömmliche Momentenmethode.....	17
5.2	L-Momentenmethode	20
5.3	Maximum-Likelihood-Methode.....	21
6	Anpassungstests	23
6.1	Kolmogorov-Smirnov-Test	23
6.2	Cramér–von-Mises ($n\omega^2$)-Test	24
6.3	Quantil-Korrelationstest	24
6.4	Akaiikesche Informationskriterium AIC und das Bayessche Informationskriterium BIC	25

6.5	Grafische Überprüfung.....	25
7	Konfidenzgrenzen	27
8	Verteilungsfunktionen	28
8.1	Normal-Verteilung.....	28
8.2	2-parameterige Log-Normal-Verteilung	29
8.3	3-parameterige Log-Normal-Verteilung	30
8.4	Allgemeine Extremwertverteilung (AEV).....	31
8.5	Gumbel-Verteilung (E1).....	33
8.6	Weibull-Verteilung.....	34
8.7	Log-Weibull-Verteilung	36
8.8	Frechet-Verteilung.....	37
8.9	Pearson III-Verteilung (P III)	38
8.10	Log-Pearson III-Verteilung (Log-P III).....	40
8.11	Exponentialverteilung.....	41
8.12	Gammaverteilung	42
8.13	3-parametrische Allgemeine Pareto Verteilung.....	44
8.14	2-parametrische Pareto Verteilung	45
9	Verwendete externe Javabibliotheken und deren Urheberrechte	47
9.1	A numerical library in Java for scientists and engineers (Lau, 2004)	47
9.2	FPL Statistics Group Nonlinear Optimization Java Package (Verrill, 2010).....	47
9.3	Colt Project (CERN, 2010).....	47
	FAQ-Frequently Asked Questions	49
	<u>Literaturverzeichnis</u>	51

I Nutzungsbedingungen

Mit der Überlassung der Software und der Dokumentation erkennt der Nutzer (nachfolgend als Lizenznehmer bezeichnet) folgende Lizenzvereinbarung der Ruhr-Universität Bochum, Lehrstuhl für Hydrologie, Wasserwirtschaft und Umwelttechnik, 44780 Bochum an:

Nutzungsrechte

Die Ruhr-Universität Bochum gewährt hiermit dem Lizenznehmer das nicht ausschließliche und nicht übertragbare Recht zur Nutzung der Software EXANTO und der mitgelieferten Dokumentation auf der Grundlage der nachfolgenden Bedingungen. Im Zweifelsfall kann keine Passage dieser Vereinbarung als Übertragung von Eigentumsrechten oder Urheberrechten interpretiert werden. Die Lizenz erlaubt die Verwendung der EXANTO-Software für wissenschaftliche und kommerzielle Zwecke.

Der Lizenznehmer erkennt etwaige Lizenzrechte Dritter, insbesondere die Urheberrechte, die in Kapitel 9 (verwendete externe Javabibliotheken und deren Urheberrechte) genannt sind an und verpflichtet sich zu deren Wahrung.

Die Lizenznehmer sind für die Installation und Anpassung der EXANTO-Software an ihre Computer und die Anpassung der Computerkonfiguration selbst verantwortlich. Seitens der Ruhr-Universität Bochum wird keine Garantie für die Funktionsfähigkeit der EXANTO-Software, für deren Kompatibilität mit den Computern des Lizenznehmers oder etwaige Schäden an Hard- und Software des Lizenznehmers gewährt. Die Lizenznehmer sind für die Anwendung der Software und für die Ergebnisse der Software-Anwendung eigenverantwortlich.

Die Nutzung ist wie folgt beschränkt:

Die Lizenz entsprechend dieser Vereinbarung gibt das Recht, die EXANTO-Software auf jedem Computer des Lizenznehmers zu nutzen. Eine Verwendung in einem Netzwerk, das einer dritten Partei Zugriff auf Software und/ oder Dokumentation erlaubt, wird ausgeschlossen. Der Lizenznehmer hat nicht das Recht, die EXANTO-Software einer dritten Partei zugänglich zu machen, Unterlizenzen zu vergeben, die Software und die Dokumentation zu veräußern oder entgeltlich oder unentgeltlich deren Nutzung für eine dritte Partei zu ermöglichen. Der Lizenznehmer hat nicht das Recht ohne eine vorherige schriftliche Übereinkunft mit der Ruhr-Universität Bochum Veränderungen am Quellcode vorzunehmen.

Haftungsausschluss

Die Software EXANTO wurde mit großer Sorgfalt und auf der Grundlage umfangreicher Testanwendungen entwickelt. Trotzdem kann für die Richtigkeit der Ergebnisse einer Anwendung von EXANTO keine Garantie übernommen werden. Die Ruhr-Universität Bochum ist weder verantwortlich noch haftbar für etwaige Verluste, die dem Nutzer aus der Verwen-

dung der EXANTO-Software entstehen. Die Ruhr-Universität Bochum übernimmt keine Garantie, dass die EXANTO-Software virusfrei ist. Die Ruhr-Universität Bochum haftet weder dem Lizenznehmer noch einer anderen Partei für etwaige Schäden oder Verluste, die sich aus Nutzungen der EXANTO-Software, der Dokumentation der Beispieldaten oder anderer Produkte im Zusammenhang mit EXANTO ergeben. Dieser Ausschluss betrifft materielle und immaterielle Schäden und Verluste einschließlich Datenverluste und entgangene Gewinne. Der Lizenznehmer erkennt an, dass die EXANTO-Software eine Reihe numerischer und methodischer Näherungsverfahren beinhaltet und akzeptiert die Verantwortung für die korrekte Interpretation der Berechnungsergebnisse. Der Lizenznehmer ist dafür verantwortlich, dass die Ruhr-Universität Bochum gegenüber dritten Parteien weder Verluste noch Schäden aus der Anwendung der EXANTO-Software oder der Dokumentation durch den Lizenznehmer entstehen. Ungeachtet der oben genannten Bedingungen, nach der eine Haftung der Ruhr-Universität Bochum für Verluste oder Schäden (einschließlich Personenschäden oder Todesfälle) ausgeschlossen wird, beschränkt sich eine etwaige Haftung in der Höchstsumme auf die Lizenzgebühren, die der Lizenznehmer für die Nutzung der EXANTO-Software gezahlt hat.

Vertraulichkeit

Der Lizenznehmer erkennt an, dass die EXANTO-Software und die Dokumentation strikter Vertraulichkeit unterliegen. Eine Weitergabe oder Kenntnismachung der Software oder der Dokumentation an eine dritte Partei, die gegenüber der Ruhr-Universität diese Nutzungsbedingungen nicht ausdrücklich anerkannt hat, wird ohne schriftliche Erlaubnis der Ruhr-Universität Bochum ausgeschlossen.

1 Einführung

Das Programm EXANTO - **Ext**reme **V**alue **A**nalysis **T**ool - dient der statistischen Auswertung von Stichproben (z.B. Jahreshöchstwerte des Abflusses) unter Verwendung verschiedener Verteilungsfunktionen. Folgende Verteilungsfunktionen werden zur Verfügung gestellt:

- Gumbelverteilung
- Allgemeine Extremwertverteilung GEV
- Weibullverteilung
- Log-Weibullverteilung
- Exponentialverteilung
- Pearson III-Verteilung
- Log-Pearson III-Verteilung
- Gammaverteilung
- Normalverteilung
- Log-Normal-Verteilung
- 3-Parametrische Log-Normal-Verteilung
- Frechet-Verteilung
- 2-Parametrische Generalized Pareto Verteilung
- 3-Parametrische Generalized Pareto Verteilung

Die Parameter der Verteilungsfunktionen können wahlweise mit der Momentenmethode, mit L-Momenten oder mit der Maximum-Likelihood-Methode geschätzt werden. Es besteht die Möglichkeit, verschiedene Jährlichkeiten vorzugeben, für die eine Berechnung durchgeführt werden soll. Kenngrößen zur Erfassung der Anpassungsgüte und Konfidenzintervalle werden berechnet. Diese Dokumentation beschreibt die statistischen Grundlagen der verwendeten Verteilungsfunktionen, Parameterschätzmethoden, Anpassungstests und Momentenschätzverfahren, die in EXANTO verwendet werden. Diese Dokumentation ist kein statistisches Lehrbuch, hierfür sei auf die in den Kapiteln zitierte Literatur verwiesen.

EXANTO ist Freeware!!! D. h.:

- Für die Richtigkeit der Ergebnisse wird keine Garantie gegeben, die Ergebnisse von EXANTO wurden mit großen Aufwand auf die Richtigkeit überprüft, es kann aber

Einzelfälle geben, bei denen EXANTO keine korrekten Ergebnisse liefert. Daher sind die Ergebnisse immer auf Plausibilität zu überprüfen (Q-Q-Plot, P-P-Plot, Anpassungsgüte) und nicht ungeprüft für eine Bemessung zu verwenden.

- Fehler, z.B. bei der Eingabe der Parameter in die Kontrolldatei, werden nicht abgefangen oder erklärt!!!
- Die Verteilungsfunktionen werden an alle Stichproben angepasst (z.B. eine nach oben offene Verteilungsfunktion an eine Stichprobe, die einem oberen Grenzwert zustrebt), durch das Programm erfolgt keine Prüfung der Sinnhaftigkeit der Verteilungsfunktion. Daher sind Ergebnisse und die verwendeten Verteilungsfunktionen immer auf Plausibilität zu prüfen!!!
- Es besteht kein Anspruch auf Support.

Aber!!! treten Fehler auf, wird um eine Fehlerbeschreibung per email (bastian.klein@rub.de) gebeten. Nach Möglichkeit ist die entsprechende Stichprobe, bei der der Fehler aufgetreten ist, als CSV-Datei mitzuschicken. Wir werden gern versuchen, den Fehler zu beheben bzw. Antwort auf das Problem zu geben.

2 EXANTO

2.1 Disclaimer

The author of this software EXANTO accepts no responsibility for direct, indirect, special, incidental, or consequential damages resulting from the use of this product and makes no guarantees, expressed or implied, about its quality, reliability, or any other characteristic. This software is provided "AS IS" and its author makes no guarantees that the software is without bugs, or that it will run on any particular platform.

2.2 Copyright Notice

EXANTO Copyright © 2010 Ruhr University, Institute of Hydrology and Water Management. All rights reserved. Permission is hereby granted, without written agreement and without license or royalty fees, to use and copy this software and its documentation for any purpose.

2.3 Systemvoraussetzungen:

EXANTO kann auf allen Systemen ausgeführt werden (LINUX, Windows, etc.). Für das Ausführen von EXANTO ist nur eine Java Virtual Machine erforderlich.

Java Virtual Machine ist verfügbar unter <http://www.oracle.com/technetwork/java/javase/downloads/index.html> (Java Runtime Environment (JRE) 6 herunterladen und installieren)

2.4 Ausführen von EXANTO

In der Eingabeaufforderung (Unter Windows: Start-Ausführen... cmd eingeben) zu dem Verzeichnis wechseln, in dem die Datei ExAnTo.jar liegt.

Laufwerk wechseln in Eingabeaufforderung: Laufwerksbuchstabe eingeben

z.B.

c:

Verzeichnis wechseln: cd Verzeichnisname

z.B.

```
cd c:\programme\exanto
```

Für den Start ist eine Kontrolldatei erforderlich, in der die Parameter für die Berechnung festgelegt werden.

Start von EXANTO:

```
java -jar ExAnTo.jar „Pfad Name der Kontrolldatei“
```

z.B. Liegt die Kontrolldatei in dem Verzeichnis von ExAnTo.jar

```
java -jar ExAnTo.jar control.txt
```

Liegt die Kontrolldatei in einem anderen Verzeichnis als ExAnTo.jar

```
java -jar ExAnTo.jar c:\statistics\dam\control.txt
```

Liegt in dem Verzeichnis mit ExAnTo.jar die Kontrolldatei mit dem Namen control.txt, dann kann EXANTO auch mit Doppelklick auf die Datei ExAnTo.jar gestartet werden !!! Programmausgaben auf den Bildschirm und somit Fehlermeldungen werden dann aber nicht auf dem Bildschirm ausgegeben! In diesem Fall muss natürlich die Endung .jar mit der Java Virtual Machine (Java(TM) Platform SE binary) verknüpft sein. Dies funktioniert im Allgemeinen automatisch bei der Installation der Java Runtime Environment (JRE). Ansonsten im Datei-Explorer unter Extras-Ordneroptionen-Dateitypen bei der Erweiterung .jar den Dateityp zu Java(TM) Platform SE binary ändern.

Nach dem Aufruf läuft EXANTO automatisch durch und gibt die Ergebnisse aus, es sind keine weiteren Aktionen des Anwenders erforderlich.

2.5 Stichprobe

Das Dateiformat ist .csv (kann über Excel gespeichert werden) mit Semikolon ";" als Trennzeichen (!!! im US Format ist das Komma Standardtrennzeichen für csv-Dateien). In den Spalten stehen die zu untersuchenden Stichproben mit dem Reihennamen in der ersten Zeile. Die Werte beginnen in der zweiten Zeile. In der Reihe sind keine noData-Werte zugelassen. Das Dezimaltrennzeichen ist Punkt. Beispiel für eine Stichprobendatei siehe in dem Beispielordner „Sample.csv“ (in diesem, hier zur Verfügung gestellten Beispiel werden 2 Datenreihen untersucht) in dem Verzeichnis Data.

Ausschnitt aus " Sample.csv " (Zu beachten: Der Reihename ist in erster Zeile anzugeben. Die Reihen sind durch ein Semikolon ";" zu trennen. Das zu verwendende Dezimaltrennzeichen ist ein Punkt ".") Eine Beispielsdatei:

```
Sample1;Sample2  
23.6;4.25
```

6.9;5.14
16.6;4.69
17.2;3.43
5.86;8.86

2.6 Kontrolldatei

Beispiel „control.txt“ im Anwendungsbeispiel.

Kommentarzeilen sind mit „%“ am Anfang zu kennzeichnen. Bei der Angabe der Parameter unbedingt auf Groß- und Kleinschreibung achten.

1. Zeile: Datei mit den Stichproben, absoluter oder auch relativer (bezogen auf den Speicherort der ExAnTo.jar Datei) Pfad möglich.

z.B.

#absolute Pfadangabe

D:\Ergebnisse\Extremwertstatistik\Data\sample.csv

#relative Pfadangabe, sample.csv liegt im Ordner Data

Der Ordner Data befindet sich im gleichen Verzeichnis wie ExAnTo.jar

Data\sample.csv

Die Datei sample.csv liegt im gleichen Verzeichnis wie ExAnTo.jar

sample.csv

2. Zeile: Ordner in dem die Ergebnisse gespeichert werden sollen (wenn dieser Ordner nicht vorhanden ist, wird er erstellt), absoluter oder auch relativer (bezogen auf den Speicherort der ExAnTo.jar Datei) Pfad möglich.

z.B.

#absolute Pfadangabe

D:\Ergebnisse\Extremwertstatistik\results

#relative Pfadangabe

results

3. Zeile: Schlüssel ob die Wiederkehrintervalle aus Zeile 4, für die die Quantilen berechnet werden, Wiederkehrintervallen der partiellen oder der jährlichen Reihe entsprechen (siehe Kapitel 4).

Wiederkehrintervalle in der partiellen Serie: Tp

Wiederkehrintervalle in der jährlichen Serie: Tn

z.B.

Tp

4. Zeile: Wiederkehrintervalle, für die Quantile aus den angepassten Verteilungsfunktionen berechnet werden sollen. Die verschiedenen Wiederkehrintervalle sind mit tab zu trennen.

z.B.

1.05 2 5 10 20 25 50 100

!!! Wenn die Wiederkehrintervalle als Wiederkehrintervalle der jährlichen Serie interpretiert werden, müssen die angegebenen Jährlichkeiten größer 1,0 sein, Eine Jährlichkeit von 1.0 kann nur aus partiellen Serien ermittelt werden. !!!

5. Zeile: Signifikanzniveau (Angabe als Wert $1-\alpha$) für das die Konfidenzgrenzen der Verteilungsfunktionen berechnet werden sollen.

!!! Da die Konfidenzgrenzen über Bootstrap resampling berechnet werden, kann es bei großen Stichprobengrößen und insbesondere bei der Maximum-Likelihood Methode vorkommen, dass die Berechnung sehr lange dauert. !!!!

Die Berechnung der Konfidenzintervalle kann ausgeschaltet werden, indem 0 oder 1 eingegeben wird.

z.B.

0.68

In den darauf folgenden Zeilen werden die gewünschten Verteilungsfunktionen und die Schätzmethoden angegeben:

Abkürzungen für die Verteilungsfunktionen:

- GU Gumbelverteilung
- AE Allgemeine Extremwertverteilung
- GP2 2-parametrische Generalized Pareto
- GP3 3-parametrische Generalized Pareto
- W Weibullverteilung
- LW Log-Weibullverteilung
- E Exponentialverteilung
- P3 Pearson III-Verteilung
- LP3 Log-Pearson III-Verteilung
- GA Gammaverteilung
- N Normalverteilung
- LN Log-Normal-Verteilung
- LN3 3-Parametrische Log-Normal-Verteilung
- F Frechet-Verteilung

Abkürzungen für die Parameterschätzmethode

- MM Momentenmethode
- LM L-Momenten Methode
- ML Maximum-Likelihood-Methode

z.B.

AE
MM
GU
LM
P3
ML

Beispiel Kontrolldatei:

D:\Exanto\Data\sample.csv

D:\Exanto\results

Tp

1.05 2 5 10 20 25 50 100 200 500 1000 2000 5000 10000

0.68

AE
MM
GU
LM
P3
ML

2.7 Ergebnisdateien

Es wird ein Ausgabeordner für jede Datenreihe in der Eingabedatei erstellt.

In diesem Ausgabeordner werden zwei unterschiedliche Ausgabedateien abgelegt:

- Für jede Verteilungsfunktion wird eine eigene Ausgabedatei in dem jeweiligen Ordner der Datenreihe erstellt. Name: Stichprobenname + Kürzel Verteilungsfunktion + Parameterschätzmethode, z.B. Sample1_GP3_ML.csv (Abb. 2.1)
- Für jede Datenreihe wird in den Ergebnisordner eine Zusammenfassung zum Vergleich von allen angepassten Verteilungsfunktionen erstellt (Abb. 2.2)

Die Daten werden mit Trennzeichen „;“ als .csv Dateien gespeichert, das Dezimaltrennzeichen ist „Punkt“. CSV-Dateien können direkt mit Excel geöffnet werden.

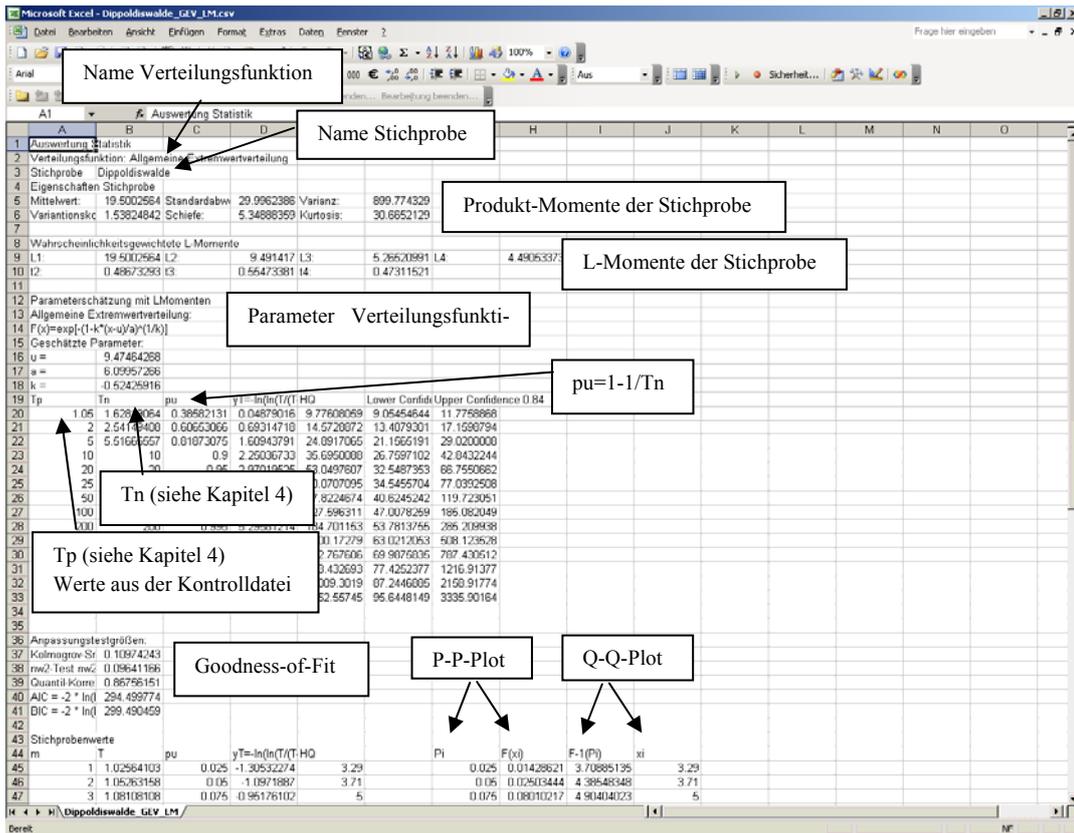


Abb. 2.1: Beispiel für eine Ergebnisdatei für eine Stichprobe und Verteilungsfunktion

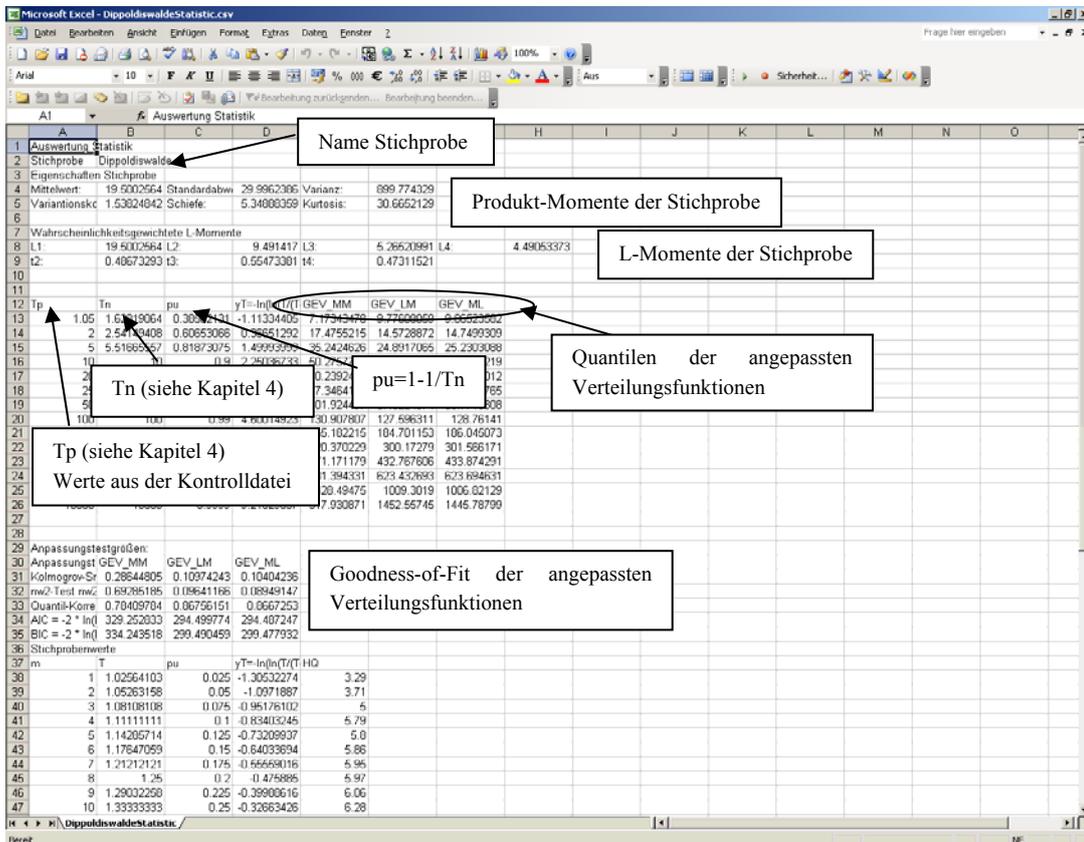


Abb. 2.2: Beispiel für eine Zusammenfassung der angepassten Verteilungsfunktion für eine Stichprobe

3 Ermittlung des Bemessungshochwassers auf statistischer Grundlage

3.1 Statistische Analyse von Hochwasserabflüssen

Durch die Anpassung einer theoretischen Verteilungsfunktion an eine gemessene Stichprobe von Jahreshöchstwerten des Abflusses können Werte mit einer kleineren Überschreitungswahrscheinlichkeit extrapoliert werden. Die statistische Analyse von Abflusswerten kann für eine jährliche oder partielle Serie erfolgen. Bei der jährlichen Serie werden aus jedem Beobachtungsjahr der höchste Scheitelabfluss und bei der partiellen Serie alle über einem bestimmten Grenzwert liegenden unabhängigen Hochwasserscheitelabflüsse verwendet. Die hier vorgestellten Grundlagen der statistischen Analyse von Hochwasserabflüssen gelten genauso für die statistische Analyse anderer hydrologischer Zufallsvariablen wie z. B. die Hochwasserfülle. Weitergehende Grundlagen zur statistischen Analyse von Zufallsvariablen können z. B. Hosking & Wallis (1997), Plate (1993), Salvadori et al. (2007) und Stedinger et al. (1993) entnommen werden.

Empfehlungen für die statistische Analyse von Hochwasserabflüssen gibt in Deutschland das DVWK-Merkblatt 251 (1999). Darin werden Empfehlungen für die Datengewinnung, Datenprüfung, die Wahl der Verteilungsfunktion und die Verifizierung der Ergebnisse gegeben. In den USA werden zum gleichen Thema Empfehlungen im Bulletin 17B (IACWD, 1982) gegeben (Kommentare und Anmerkungen zur Bulletin 17B siehe Griffis & Stedinger, 2007; Stedinger & Griffis, 2008).

Statistische Analysen sind in ihren Ergebnissen stark von den Datengrundlagen abhängig. Ihre Ergebnisse sollten auf jeden Fall über einen Vergleich mit der Hochwasserstatistik an benachbarten Pegeln mit längerer Beobachtungsdauer, einen Vergleich mit historischen Hochwasserereignissen am betrachteten Pegel oder im betrachteten Gebiet oder dem Vergleich mit Abflüssen, die über eine Niederschlag-Abflussmodellierung aus extremen Niederschlägen ermittelt worden sind, überprüft werden.

Für eine statistische Analyse werden an die Stichprobe folgende Anforderungen gestellt:

Unabhängigkeit der Stichprobenwerte: Aufeinanderfolgende Beobachtungen beeinflussen sich nicht gegenseitig, die Reihenfolge ist also gleichgültig. Bei Hochwasserscheitelwerten kann die Unabhängigkeit aufeinander folgender Stichprobenelemente mit Hilfe von Abb. 3.1 (LAWA, 1997) überprüft werden. Beide Ereignisse gelten als unab-

hängig, wenn der zeitliche Abstand zwischen ihnen größer als 7 Tage oder $A > B$ ist.

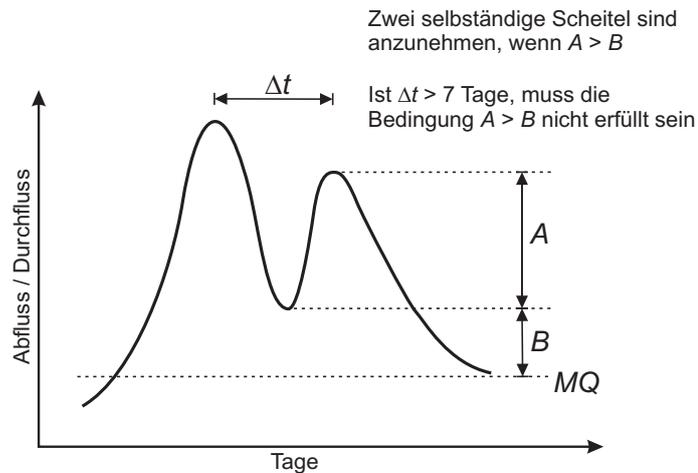


Abb. 3.1: Kriterien für statistisch unabhängige Hochwasserscheitelwerte (LAWA, 1997)

In Bacchi et al. (1992) werden zwei aufeinander folgende Gipfel dann als unabhängig betrachtet, wenn der Abstand zwischen den beiden Scheiteln mindestens das 20-fache der Scheitelanstiegszeit ist und der Abfluss zwischen den beiden Scheiteln um mehr als 80% zurückgeht.

Repräsentanz der Stichprobe: Die Stichprobe sollte die Grundgesamtheit abbilden.

Homogene Grundgesamtheit: Die Hochwasserabflüsse dürfen in der zugrunde liegenden Beobachtungszeit nicht durch Veränderungen im Einzugsgebiet, wie z. B. Urbanisierung, beeinflusst sein. Inhomogenitäten sind grundsätzlich durch sprunghafte oder kontinuierliche Veränderungen der Hochwasserabflüsse erkennbar.

Stationäre Grundgesamtheit: Forderung nach Unabhängigkeit mit der Zeit. Die statistischen Kenngrößen wie Erwartungswert und Varianz verändern sich nicht mit der Zeit, es darf auch kein Trend, wie z. B. Klimaänderung, vorliegen.

3.2 Verteilungsfunktionen

Eine Verteilungsfunktion beschreibt die Grundgesamtheit, also die Menge aller Realisationen einer stetigen Zufallsvariable X . Die Häufigkeitsverteilung einer stetigen Zufallsvariablen X kann durch die Wahrscheinlichkeitsdichtefunktion pdf „probability density function“ $f(x)$ beschrieben werden. Das Integral der Funktion von $-\infty$ bis $+\infty$ ergibt 1.

$$\int_{-\infty}^{\infty} f(x) dx = 1. \quad (2.1)$$

Die Verteilungsfunktion cdf „cumulative density function“ $F(x)$, beschreibt die Wahrscheinlichkeit, dass die Zufallsvariable X kleiner oder gleich dem Wert x ist.

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(x) dx. \quad (2.2)$$

Der Wert von $F(x)$ wird auch als Unterschreitungswahrscheinlichkeit $P_u(x)$ bezeichnet. Die Überschreitungswahrscheinlichkeit $P_{\bar{u}}(x)$ ergibt sich dann zu

$$P_{\bar{u}}(x) = 1 - P_u(x) = 1 - F(x). \quad (2.3)$$

Beschreibt die Zufallsvariable X jährliche Höchstwerte, ergibt sich das statistische Wiederkehrintervall zu

$$T = \frac{1}{P_{\bar{u}}} = \frac{1}{1 - P_u}. \quad (2.4)$$

Das Quantil der Unterschreitungswahrscheinlichkeit P_u , also der Wert x , der mit einer Wahrscheinlichkeit P_u von der Zufallsvariable X unterschritten wird, ergibt sich aus der inversen Verteilungsfunktion. Die Quantilfunktion lautet dann

$$x_{P_u} = F^{-1}(P_u). \quad (2.5)$$

Nach DVWK-Merkblatt 251 (1999) sind für die statistische Analyse von Hochwasserabflüssen die Verteilungen Allgemeine Extremwertverteilung, Extremwertverteilung Typ 1 (Gumbel-Verteilung), Pearson-Typ 3-Verteilung, Weibullverteilung, die 3-parametrische logarithmische Normalverteilung und die logarithmische Pearson-Typ 3-Verteilung geeignet. In EXANTO werden zusätzlich die Verteilungsfunktionen: Log-Weibull-Verteilung, 3- und 2-parametrische Allgemeine Pareto-Verteilung, Frechet-Verteilung, Exponential-Verteilung, Gamma-Verteilung, Normal-Verteilung, Log-Normal-Verteilung berücksichtigt.

Ein Problem der Extremwertstatistik ist die Wahl der richtigen Verteilungsfunktion, da sich die unterschiedlichen, an die Stichprobe angepassten Verteilungsfunktionen mit abnehmender Überschreitungswahrscheinlichkeit im Extrapolationsbereich immer stärker voneinander unterscheiden und für große Wiederkehrintervalle sehr unterschiedliche Scheitelabflüsse liefern können. Bei der Aufstellung von Hochwasserstatistiken für mehrere Pegel in einer Region ist es daher ratsam, eine regional gültige Verteilungsfunktion zu ermitteln, da hierbei eine größere Datenmenge für die Wahl der richtigen Verteilungsfunktion genutzt werden kann.

In Vogel & Wilson (1996) sind Studien zu extremen Hochwasserereignissen in der ganzen Welt zusammengefasst. Dabei herrscht allgemeiner Konsens, dass insbesondere bei der Wahl von regionalen Verteilungsfunktionen die Verteilung der jährlichen maximalen Abflüsse am besten durch die Allgemeine Extremwertverteilung angenähert werden kann.

4 HQ-Werte für kleine Wiederholungszeitspannen

In einer Beobachtungsreihe von n Jahren sind neben den n Jahres-HQ-Werten noch weitere HW-Ereignisse enthalten. Dadurch, dass nur Jahreshöchstabflüsse betrachtet werden, kann es vorkommen, dass in einem Jahr mit mehreren Hochwassern nur das höchste Ereignis betrachtet wird und die kleineren Ereignisse vernachlässigt werden. Die aus den Jahreshöchstabflüssen abgeleiteten Wiederkehrintervalle T_n sind daher nur für Jahreshöchstwerte gültig. In der Praxis wird jedoch das mittlere statistische Wiederkehrintervall T_p zwischen Hochwasserereignissen einer bestimmten Größe gesucht, unabhängig ob sie Jahreshöchstwerte sind oder nicht. Da diese Hochwasserereignisse häufiger zu erwarten sind, ist das mittlere statistische Wiederkehrintervall T_p immer kleiner als T_n . Nennenswerte Unterschiede sind allerdings nur bis zu einem statistischen Wiederkehrintervall T_p von ca. 5 Jahren zu erwarten.

In EXANTO werden die in der Kontrolldatei angegebenen Wiederkehrintervalle als das mittlere statistische Wiederkehrintervall T_p zwischen Hochwasserereignissen einer bestimmten Größe angesehen. Um die HQ-Werte für die angegebenen Wiederkehrintervalle aus der über die Jahreshöchstwerte angepassten Verteilungsfunktion zu berechnen, werden die Wiederkehrintervalle T_p für Jährlichkeiten ≤ 10 a in T_n über folgende Beziehung umgerechnet (DVWK, 1999):

$$T_n = \frac{\exp(1/T_p)}{\exp(1/T_p) - 1}. \quad (2.6)$$

Für $T_p = 1$ a ergibt sich somit nach Gl. (2.6) ein statistisches Wiederkehrintervall von $T_n = 1,58$ a.

Um die Jährlichkeit T_p für ein festgelegtes HQ zu bestimmen, wird die sich aus der Verteilungsfunktion ergebende Jährlichkeit T_n über folgende Beziehung umgerechnet

$$T_p = \frac{1}{\ln(T_n / (T_n - 1))}. \quad (2.7)$$

5 Parameterschätzung

Ziel ist es, die Parameter der Verteilungsfunktion aus den Stichprobenwerten zu bestimmen, und somit eine möglichst gute Anpassung der Verteilungsfunktion an die Stichprobenwerte zu erreichen. Die Parameterschätzung sollte folgende Eigenschaften besitzen:

- **Erwartungstreue:** Ein Schätzer heißt erwartungstreu für einen Parameter, wenn er im Mittel den richtigen Wert liefert.
- **Konsistenz:** Der Schätzer soll umso genauere Werte liefern, je größer die Stichprobe ist.
- **Effizienz:** Der Schätzer sollte eine kleinere Varianz haben, als jeder andere Schätzer.
- **Suffizienz:** Alle in der Stichprobe enthaltenen Informationen über den geschätzten Parameter sollten genutzt werden.
- **Robustheit:** Der Schätzer sollte unempfindlich gegenüber Ausreißern sein.

Für die Parameterschätzung stehen verschiedene Methoden zur Verfügung:

5.1 Herkömmliche Momentenmethode

Es wird ein Zusammenhang zwischen den Produktmomenten, im folgenden Momente genannt, der Verteilung und den Momenten der Stichprobe aufgestellt und daraus die Parameter der Verteilung bestimmt. Für die Ermittlung von m Parametern sind m Gleichungen erforderlich. Der Vorteil dieser Parameterschätzung ist die einfache Handhabung, der Nachteil ist die unsichere Bestimmung der Momente höherer Ordnung bei vorliegender kleiner Stichprobe.

Die Ursprungsmomente¹ μ'_r der Ordnung r einer stetigen Verteilung ergeben sich zu:

$$\mu'_r = \int_{-\infty}^{\infty} x^r f(x) dx \quad (2.8)$$

Das erste Ursprungsmoment μ'_1 entspricht dem Erwartungswert μ der Verteilung:

¹ Hinweise zur Notation: die Parameter der Grundgesamtheit werden mit griechischen und großen lateinischen Buchstaben und die Schätzwerte aus der Stichprobe mit kleinen lateinischen Buchstaben beschrieben

$$\mu = E(x) = \mu'_1. \quad (2.9)$$

Bezieht man das Moment auf den Mittelwert, ergibt sich das zentrale Moment μ_r der Ordnung r der stetigen Verteilung zu:

$$\mu_r = \int_{-\infty}^{\infty} (x - \mu)^r f(x) dx. \quad (2.10)$$

$\mu_0 = 1$, $\mu_1 = 0$, das zweite Zentralmoment entspricht der Varianz σ^2

$$\mu_2 = \sigma^2. \quad (2.11)$$

Die ersten fünf Zentralmomente können folgendermaßen aus den Ursprungsmomenten ermittelt werden:

$$\mu_1 = 0 \quad (2.12)$$

$$\mu_2 = -\mu_1'^2 + \mu_2' \quad (2.13)$$

$$\mu_3 = 2\mu_1'^3 - 3\mu_1'\mu_2' + \mu_3' \quad (2.14)$$

$$\mu_4 = -3\mu_1'^4 + 6\mu_1'^2\mu_2' - 4\mu_1'\mu_3' + \mu_4' \quad (2.15)$$

$$\mu_5 = 4\mu_1'^5 - 10\mu_1'^3\mu_2' + 10\mu_1'^2\mu_3' - 5\mu_1'\mu_4' + \mu_5'. \quad (2.16)$$

Die Wurzel der Varianz entspricht der Standardabweichung σ . Die Schiefe C_S ergibt sich über das dritte Zentralmoment zu

$$C_S = \gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{\mu_3}{\sigma^3} \quad (2.17)$$

und die Kurtosis wird über das vierte Zentralmoment ermittelt:

$$C_k = \gamma_2 = \frac{\mu_4}{\mu_2^2} - 3 = \frac{\mu_4}{\sigma^4} - 3. \quad (2.18)$$

Die Ursprungsmomente m'_r der Ordnung r aus der Stichprobe (x_1, x_2, \dots, x_n) ergeben sich zu

$$m'_r = \frac{1}{n} \sum_{i=1}^n x_i^r, \quad (2.19)$$

und die Zentralmomente m_r der Stichprobenwerte der Ordnung r zu

$$m_r = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^r. \quad (2.20)$$

Die Zentralmomente von der Stichprobe nach Gl. 2.20 sind nicht erwartungstreu (biased), die erwartungstreuen (unbiased) Zentralmomente ergeben sich zu:

$$\hat{m}_2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2.21)$$

$$\hat{m}_3 = \frac{n}{(n-1)(n-2)} \sum_{i=1}^n (x_i - \bar{x})^3 \quad (2.22)$$

$$\hat{m}_4 = \frac{n^2}{(n-1)(n-2)(n-3)} \sum_{i=1}^n (x_i - \bar{x})^4. \quad (2.23)$$

Der Mittelwert der Stichprobe ist:

$$\bar{x} = m'_1 = \frac{1}{n} \sum_{i=1}^n x_i, \quad (2.24)$$

Die Varianz der Stichprobe ist:

$$s^2 = \hat{m}_2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \quad (2.25)$$

die Standardabweichung:

$$s = \sqrt{\hat{m}_2} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}, \quad (2.26)$$

der Variationskoeffizient:

$$c_v(x) = \frac{s(x)}{\bar{x}} = \frac{\hat{m}_2^{1/2}}{m'_1}, \quad (2.27)$$

der Schiefekoeffizient:

$$c_s(x) = \frac{n}{(n-1)(n-2)} \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{s^2(x)^{3/2}} = \frac{\hat{m}_3}{\hat{m}_2^{3/2}}, \quad (2.28)$$

und die Kurtosis

$$c_k = \frac{n^2}{(n-1)(n-2)(n-3)} \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{s^2(x)^2} - 3 = \frac{\hat{m}_4}{\hat{m}_2^2} - 3. \quad (2.29)$$

Aufgrund des Exponenten r in Gl. 2.20 wird die Schätzung der Momente höherer Ordnung immer unsicherer. Ist in der Stichprobe ein Extremwert, werden die Momente höherer Ordnung, insbesondere bei kleinen Stichprobenumfängen stark beeinflusst, da die Abweichung vom Mittelwert mit dem Exponenten r in die Summenbildung eingeht.

5.2 L-Momentenmethode

L-Momente sind eine andere Möglichkeit, die statistischen Eigenschaften von Verteilungsfunktionen zu beschreiben. Die L-Momente können als Funktion der wahrscheinlichkeitsgewichteten Momente (Greenwood et al., 1979; Hosking, 1990) beschrieben werden. Die wahrscheinlichkeitsgewichteten Momente $M_{p,r,s}$ einer Zufallsvariable X mit der Verteilungsfunktion $F(x)$ sind definiert als (Greenwood et al., 1979)

$$M_{p,r,s} = E \left[X^p \{F(X)\}^r \{1-F(X)\}^s \right], \quad (2.30)$$

mit dem Sonderfall für $p=1, s=0$ für eine Verteilung mit der Quantilfunktion $x(F)$:

$$\beta_r = M_{1,r,0} = \int_0^1 x(F) F^r dF. \quad (2.31)$$

Aus einer aufsteigend geordneten Stichprobe ($x_1 \leq x_2 \leq \dots \leq x_n$) des Umfangs n ergibt sich der erwartungstreue Schätzer für b_r nach Landwehr et al. (1979) zu

$$b_r = \frac{1}{n} \sum_{i=1}^n \frac{(i-1)(i-2)\dots(i-r)}{(n-1)(n-2)\dots(n-r)} x_i; \text{ mit } r = 1, 2, \dots, (n-1) \quad (2.32)$$

Es ist schwierig, die wahrscheinlichkeitsgewichteten Momente als Lage-, Streuungs- oder Formmaß der Verteilungsfunktion zu interpretieren. Diese Information ergibt sich viel besser aus den L-Momenten. Diese sind wie folgt definiert:

$$\lambda_r = \int_0^1 x(F) P_{r-1}^*(F) dF \quad (2.33)$$

mit den Polynomen

$$P_r^*(F) = \sum_{k=0}^r p_{r,k}^* F^k \quad (2.34)$$

mit

$$p_{r,k}^* = (-1)^{r-k} \binom{r}{k} \binom{r+k}{k}. \quad (2.35)$$

Die L-Momente können als Linearkombination der wahrscheinlichkeitsgewichteten Momente beschrieben werden:

$$\lambda_{r+1} = \sum p_{r,k}^* \beta_k. \quad (2.36)$$

Die ersten vier L-Momente ergeben sich dann zu

$$\lambda_1 = \beta_0 \quad (2.37)$$

$$\lambda_2 = 2\beta_1 - \beta_0 \quad (2.38)$$

$$\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0 \quad (2.39)$$

$$\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0. \quad (2.40)$$

Das erste L-Moment λ_1 entspricht dem Mittelwert, das zweite, λ_2 , ist ein Maß für die Varianz, das dritte, λ_3 , ein Maß für die Schiefe und das vierte, λ_4 , ein Maß für die Kurtosis. Der Variationskoeffizient auf Grundlage der L-Momente ergibt sich zu

$$\tau_2 = \frac{\lambda_2}{\lambda_1}, \quad (2.41)$$

der Schiefekoeffizient zu

$$\tau_3 = \frac{\lambda_3}{\lambda_2} \quad (2.42)$$

und der Wölbungskoeffizient zu

$$\tau_4 = \frac{\lambda_4}{\lambda_2}. \quad (2.43)$$

Da bei der Ermittlung der wahrscheinlichkeitsgewichteten Momente nach Gl. 2.3.2 keine Potenzwerte auftreten, wird die Schätzung hierbei nicht so stark von Extremwerten in der Stichprobe beeinflusst wie bei der Schätzung der Produktmomente.

Weitere Definitionen zu den wahrscheinlichkeitsgewichteten Momenten und L-Momenten sind in Hosking & Wallis (1997) beschrieben. In derselben Referenz sind auch Gleichungen der L-Momente für einige Verteilungsfunktionen angegeben.

5.3 Maximum-Likelihood-Methode

Bei der Maximum-Likelihood-Methode werden die Parameter u_j einer Verteilungsfunktion so bestimmt, dass die Wahrscheinlichkeit für das Auftreten der Beobachtungswerte $x_1, x_2 \dots x_n$ maximal wird. Hierzu wird die Likelihood-Funktion

$$L(u_j) = f(x_1)f(x_2)\dots f(x_n) = \prod_{i=1}^n f(x_i) \quad (2.44)$$

bzw. die Log-Likelihood-Funktion

$$\ln L(u_j) = \ln f(x_1) + \ln f(x_2) \dots + \ln f(x_n) = \sum_{i=1}^n \ln f(x_i) \quad (2.45)$$

als Maß für die Wahrscheinlichkeit der Beobachtungswerte maximiert. Die Maximum-Likelihood-Methode ist eine robuste Schätzmethode, da sie sich nicht auf die unsichere Bestimmung höherer Momente stützt. In vielen Fällen lassen sich die Parameter mit der

Maximum-Likelihood-Methode allerdings nicht analytisch bestimmen, so dass sie mit numerischen Methoden geschätzt werden müssen.

Formeln und Näherungen für die Parameterschätzung nach der Momentenmethode, der L-Momentenmethode und der Maximum-Likelihood-Methode sind in DVWK (1999), Hosking & Wallis (1997), Rao & Hamed (2000) und Stedinger et al. (1993) gegeben. Im Anhang A sind die in dieser Arbeit verwendeten Verteilungsfunktionen und Parameterschätzungsmethoden aufgeführt.

6 Anpassungstests

Über Anpassungstests wird häufig fälschlicherweise versucht, die Verteilungsfunktion zu identifizieren, die die „beste“ Anpassung an die Stichprobe liefert. Im DVWK-Merkblatt 251 (1999) werden drei Maße empfohlen, um die Güte der Anpassung zu charakterisieren:

- Kolmogorov-Smirnov-Test
- Cramér–von-Mises ($n\omega^2$)-Test
- Quantil-Korrelationstest.

Häufig wird als Maß für die Anpassungsgüte die Kombination aus den drei oben aufgeführten Anpassungstests ($D + n\omega^2 + (1 - r_p^2)$) verwendet. In Raschke & Thürmer (2008) wird die Anwendung der im DVWK-Merkblatt 251 (1999) empfohlenen Anpassungstests und insbesondere das kombinierte Maß (Zitat: „*Dies ist natürlich zu verwerfen, da plastisch formuliert mm, cm und m addiert werden.*“) für die Anpassungsgüte kritisiert.

Stattdessen werden von Raschke & Thürmer das Akaike Informationskriterium AIC und das Bayessche Informationskriterium BIC als Kriterium für die Modellwahl empfohlen.

Generell ist die Anpassungsgüte kein trennscharfes Maß zur Unterscheidung guter Anpassungen. Es ist zu beachten, dass Anpassungstests keine „richtige“ Verteilung identifizieren können, sondern nur Verteilungen mit schlechter Anpassung ausweisen. Trotz der hier beschriebenen Anpassungstests und Maße für die Anpassungsgüte bleibt bei der Wahl der Verteilungsfunktion für den Bearbeiter noch ein großer subjektiver Spielraum.

6.1 Kolmogorov-Smirnov-Test

Bei dem KS-Test wird die betragsmäßig größte Differenz D zwischen der theoretischen und der empirischen Unterschreitungswahrscheinlichkeit an den Stellen der beobachteten Stichprobenwerte als Prüfgröße verwendet:

$$D_N = \text{Max}_{m=1 \dots N} (|P_m - F(x_m)|, |P_{m+1} - F(x_m)|) \quad (2.46)$$

mit x_m dem m -ten Element der geordneten Stichprobe, der Stichprobengröße N und der empirischen Unterschreitungswahrscheinlichkeit:

$$P_m = \frac{m-1}{N} \quad (2.47)$$

ergibt sich:

$$D_N = \text{Max}_{m=1 \dots N} \left(\left| \frac{m-1}{N} - F(x_m) \right|, \left| \frac{m}{N} - F(x_m) \right| \right). \quad (2.48)$$

6.2 Cramér–von-Mises ($n\omega^2$)-Test

Bei dem Cramér–von-Mises Test wird die Summe der quadrierten Abweichungen zwischen der theoretischen und der empirischen Unterschreitungswahrscheinlichkeit der beobachteten Abflüsse als Prüfgröße verwendet:

$$n\omega^2 = \sum_{m=1}^N [F(x_m) - P_m]^2 + \frac{1}{12N} \quad (2.49)$$

mit x_m dem m-ten Element der geordneten Stichprobe, der Stichprobengröße N und der empirischen Unterschreitungswahrscheinlichkeit:

$$P_m = \frac{m-0,5}{N} \quad (2.50)$$

ergibt sich:

$$n\omega^2 = \sum_{m=1}^N \left[F(x_m) - \frac{m-0,5}{N} \right]^2 + \frac{1}{12N}. \quad (2.51)$$

6.3 Quantil-Korrelationstest

Bei dem Quantil-Korrelationstest wird der Korrelationskoeffizient r_p zwischen den geordneten Stichprobenwerten x_m und den aus der theoretischen Quantilfunktion mit der empirischen Wahrscheinlichkeit P_m der Stichprobenwerte ergebenden Werte x_m^* als Prüfgröße verwendet:

$$r_p = \frac{\sum_{m=1}^N x_m x_m^* - N \cdot m_x \cdot m_{x^*}}{\sqrt{\left(\sum_{m=1}^N x_m^2 - N m_x^2 \right) \left(\sum_{m=1}^N x_m^{*2} - N m_{x^*}^2 \right)}} \quad (2.52)$$

mit $x_{(m)}$ dem m-ten Element der geordneten Stichprobe, der Stichprobengröße N

$$x_m^* = F^{-1}(P_m) \quad (2.53)$$

$$m_x = \frac{1}{N} \sum_{m=1}^N x_m \quad (2.54)$$

$$m_{x^*} = \frac{1}{N} \sum_{m=1}^N x_m^* . \quad (2.55)$$

Die empirischen Unterschreitungswahrscheinlichkeiten P_m ergibt sich für alle Verteilungsfunktionen zu:

$$P_m = \frac{m}{N+1} . \quad (2.56)$$

In EXANTO werden keine verteilungsabhängigen Verteilungsfunktionen verwendet (wie z.B. in DVWK, 1999). Über die Wahl von von Gl. (2.56) als verteilungsunabhängige plotting position siehe Makkonen (2008).

6.4 Akaikesche Informationskriterium AIC und das Bayessche Informationskriterium BIC

Das von Akaike (1974) entwickelte AIC verwendet wie das BIC den Wert der Log-Likelihood-Funktion $\ln L(u_j)$ nach Gl. 2.45 als Maß für die Anpassungsgüte:

$$AIC = (-2) \cdot \ln L(u_j) + 2 \cdot k , \quad (2.57)$$

mit der Anzahl der Parameter $= k$.

Das BIC geht zurück auf die Arbeiten von Schwarz (1978) und ergibt sich zu

$$BIC = (-2) \cdot \ln L(u_j) + k \cdot \ln(n) , \quad (2.58)$$

mit der Anzahl der Stichprobenwerte n .

Die beiden Kriterien führen insbesondere bei einer großen Anzahl von Beobachtungen zu unterschiedlichen Ergebnissen. Das Modell mit dem kleinsten AIC bzw. BIC ergibt danach die beste Anpassung.

Liegen Stichprobenwerte außerhalb des Definitionsbereiches von der Verteilungsfunktion, ergibt sich für diese Werte $f(x_i) = 0$ und somit die Likelihoodfunktion zu 0 bzw. die Log-Likelihoodfunktion zu $-\infty$. In EXANTO wird in diesem Fall NaN in der Ergebnisdatei ausgegeben. Definitionsgrenzen der Verteilungsfunktionen siehe Kapitel 8.

6.5 Grafische Überprüfung

Bei der visuellen Überprüfung der Anpassungsgüte der Verteilungsfunktion werden die gemessenen Werte mit der jeweiligen empirischen Wahrscheinlichkeit mit dem Verlauf der Verteilungsfunktion verglichen. Vor allem ist hierbei die Skalierung der Wahrscheinlichkeitsachse von großer Bedeutung. Daher wird in EXANTO zur Darstellung der Anpassungsgüte für die Skalierung der Wahrscheinlichkeitsachse zusätzlich die standardisierte

Gumbel-Variable

$$y_T = -\ln\{\ln[T/(T-1)]\} = -\ln[\ln(1/P_u)] \quad (2.59)$$

ausgegeben, da hiermit im Bereich großer Unterschreitungswahrscheinlichkeiten die Anpassungsgüte der Verteilung an die Stichprobe besser überprüft werden kann.

Weiter grafische Überprüfungsmöglichkeiten sind der Probability-Probability-Plot und der Quantile-Quantile-Plot. Bei dem Probability-Probability-Plot werden die theoretischen Werte der Unterschreitungswahrscheinlichkeit der geordneten Stichprobenwerte x_m über die angepasste Verteilungsfunktion $F(x_m)$ den empirischen Unterschreitungswahrscheinlichkeiten P_m der Stichprobenwerte gegenübergestellt, bei dem Quantile-Quantile-Plot werden die geordneten Stichprobenwerte x_m den aus der theoretischen Quantilfunktion mit der empirischen Wahrscheinlichkeit P_m der Stichprobenwerte ergebenden Werten x_m^* gegenübergestellt. Bei optimaler Anpassung liegen die Punkte $(F(x_m), P_m)$ bzw. (x_m, x_m^*) auf einer Ursprungsgerade mit dem Winkel 45° . Für die empirischen Unterschreitungswahrscheinlichkeiten wird für alle Verteilungen die plotting position nach Gl. (2.56) verwendet (Makkonen, 2008).

7 Konfidenzgrenzen

Die Konfidenzgrenzen in Abhängigkeit des gewählten Konfidenzkoeffizienten $(1-\alpha)$ werden für alle Verteilungsfunktionen durch Bootstrap resampling berechnet (Burn, 2003; Hall et al., 2004). Die Lage der Konfidenzgrenzen sind abhängig von der Verteilungsfunktion und der Parameterschätzmethode. Da es sich um ein resampling-Verfahren handelt, werden die Konfidenzgrenzen für die gleiche Stichprobe, Verteilungsfunktion und Parameterschätzmethode bei jeder Berechnung immer etwas unterschiedlich sein, durch eine relativ große Anzahl von bootstraps können die Unterschiede zwischen unterschiedlichen Berechnungsläufen verringert werden.

1. Aus der Stichprobe $X=[x_1, x_2, \dots, x_N]^T$ werden unter Zurücklegen der gezogenen Elemente B unabhängige Stichproben $X^*_1, X^*_2, \dots, X^*_B$ der Größe N zufällig generiert.
2. Für jede Stichprobe $X^*_1, X^*_2, \dots, X^*_B$ werden die Parameter $\hat{\theta}_i^*$ der Verteilungsfunktion mit der gewählten Parameterschätzmethode geschätzt.
3. Für jede der unter Punkt 2 angepassten Verteilungsfunktionen $\hat{F}_i(x, \hat{\theta}_i^*)$ werden die Quantile $\hat{x}_{T,i}$ für das Wiederkehrintervall T_n ermittelt.
4. Die Quantile der Wiederkehrintervalle $\hat{x}_{T,1}, \hat{x}_{T,2}, \dots, \hat{x}_{T,B}$ werden der Größe nach sortiert.
5. Die untere Konfidenzgrenze $\hat{x}_{T,(l)}$ mit $l = \alpha/2 (B + 1)$ und die obere Konfidenzgrenze $\hat{x}_{T,(u)}$ mit $u = (1 - \alpha/2)(B + 1)$ werden ermittelt.
6. Die Schritte 1 bis 5 werden für alle in der Kontrolldatei vorgegebenen Wiederkehrintervalle durchgeführt.

Die Anzahl B der Bootstraps wird in EXANTO in Abhängigkeit von der Parameterschätzmethode gewählt. Auf Grund der Rechenzeit werden bei der Maximum-Likelihood 999 Bootstraps und bei den anderen Schätzverfahren 1999 Bootstraps für die Schätzung der Konfidenzgrenzen verwendet.

8 Verteilungsfunktionen

In diesem Abschnitt werden die in EXANTO verwendeten Verteilungsfunktionen und Parameterschätzmethoden für die einzelnen Verteilungsfunktionen beschrieben. Weitere Verteilungsfunktionen sowie Gleichungen und Näherungen für die Parameterschätzung sind Rao und Hamed (2000), Hosking und Wallis (1997), Stedinger et al. (1993), Bury (1999) und DVWK (1999) zu entnehmen.

8.1 Normal-Verteilung

Parameter

m, s

$-\infty < x < \infty$

Verteilungsfunktion:

$$F(x) = \int_{-\infty}^x \frac{1}{s\sqrt{2\pi}} e^{-\frac{1}{2s^2}(x-m)^2} dx = \int_{-\infty}^{\frac{x-m}{s}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = \Phi\left(\frac{x-m}{s}\right) \quad (2.60)$$

mit $\Phi(x)$ der Wahrscheinlichkeitsdichtefunktion der Standardnormalverteilung ($m=0, s=1$)

Wahrscheinlichkeitsdichtefunktion:

$$f(x) = \frac{1}{s\sqrt{2\pi}} e^{-\frac{1}{2s^2}(x-m)^2} \quad (2.61)$$

Quantile:

$$x(F) = m + s\Phi^{-1}(F) \quad (2.62)$$

Parameterschätzung Momenten-Methode:

Die Parameter ergeben sich zu:

$$m = m'_1 \quad (2.63)$$

$$s = \sqrt{\hat{m}_2} \quad (2.64)$$

Parameterschätzung L-Momenten-Methode:

Die Parameter ergeben sich zu:

$$m = l_1 \quad (2.65)$$

$$s = \sqrt{\pi} l_2. \quad (2.66)$$

Parameterschätzung Maximum-Likelihood-Methode:

Bei der Normalverteilung ergeben die Parameterschätzmethoden nach der Momenten-Methode und der Maximum Likelihood Methode die gleichen Ergebnisse.

8.2 2-parameterige Log-Normal-Verteilung

Parameter

m, s

Verteilungsfunktion nach unten begrenzt mit dem unteren Grenzwert $0 < x < \infty$

Verteilungsfunktion:

$$F(x) = \int_{-\infty}^{\frac{\ln x - m}{s}} \frac{1}{xs\sqrt{2\pi}} e^{-\frac{1}{2s^2}(\ln x - m)^2} dx = \int_{-\infty}^{\frac{\ln x - m}{s}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = \Phi\left(\frac{\ln x - m}{s}\right) \quad (2.67)$$

mit $\Phi(x)$ der Wahrscheinlichkeitsdichtefunktion der Standardnormalverteilung ($m=0, s=1$)

Wahrscheinlichkeitsdichtefunktion:

$$f(x) = \frac{1}{xs\sqrt{2\pi}} e^{-\frac{1}{2s^2}(\ln x - m)^2} \quad (2.68)$$

Quantile:

$$x(F) = \exp\left(m + s\Phi^{-1}(F)\right) \quad (2.69)$$

Parameterschätzung Momenten-Methode:

Die Parameter ergeben sich zu:

$$s = \sqrt{\ln\left(\frac{\hat{m}_2}{m_1'^2} + 1\right)} \quad (2.70)$$

$$m = \ln m_1' - s^2/2. \quad (2.71)$$

Parameterschätzung L-Momenten-Methode:

Die Parameter werden mit den logarithmierten Werten der Stichprobe nach Gl. (2.65) und (2.66) ermittelt.

Parameterschätzung Maximum-Likelihood-Methode:

Die Parameter ergeben sich zu:

$$m = \frac{1}{N} \sum_{i=1}^N \ln x_i. \quad (2.72)$$

$$s = \sqrt{\frac{1}{N} \sum_{i=1}^N (\ln x_i - m)^2} \quad (2.73)$$

8.3 3-parameterige Log-Normal-Verteilung**Parameter**

a, m, s

Die Verteilungsfunktion ist nach unten begrenzt, der untere Grenzwert lautet $a < x < \infty$

Verteilungsfunktion:

$$F(x) = \int_{-\infty}^{\infty} \frac{1}{(x-a)s\sqrt{2\pi}} e^{-\frac{1}{2s^2}(\ln(x-a)-m)^2} dx = \int_{-\infty}^{\frac{\ln(x-a)-m}{s}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = \Phi\left(\frac{\ln(x-a)-m}{s}\right) \quad (2.74)$$

mit $\Phi(x)$ der Wahrscheinlichkeitsdichtefunktion der Standardnormalverteilung ($m=0, s=1$)

Wahrscheinlichkeitsdichtefunktion:

$$f(x) = \frac{1}{(x-a)s\sqrt{2\pi}} e^{-\frac{1}{2s^2}(\ln(x-a)-m)^2} \quad (2.75)$$

Quantile:

$$x(F) = \exp\left(m + s\Phi^{-1}(F)\right) + a \quad (2.76)$$

Parameterschätzung Momenten-Methode:

Die Parameter ergeben sich zu:

$$s = \sqrt{\ln(z_2^2 + 1)} \quad (2.77)$$

mit

$$z_2 = \frac{1 - w^{2/3}}{w^{1/3}} \quad (2.78)$$

$$w = \frac{-c_s + (c_s^2 + 4)^{0.5}}{2} \quad (2.79)$$

$$m = \ln\left(\sqrt{\hat{m}_2}/z_2\right) - 0.5 \ln(z_2^2 + 1) \quad (2.80)$$

$$a = m'_1 - \sqrt{\hat{m}_2}/z_2. \quad (2.81)$$

Parameterschätzung L-Momenten-Methode:

Die Parameter werden mit den logarithmierten Werten der Stichprobe nach Gl. (2.65) und (2.66) ermittelt.

Parameterschätzung Maximum-Likelihood-Methode:

In EXANTO werden die Parameter für die ML-Methode über ein numerisches Optimierungsverfahren (Numerisches Optimierungsverfahren Nonlinear Optimization Java Package der FPL Statistics Group (Verrill)) für die Maximierung der log-Likelihood-Funktion nach Gl. 2.45 ermittelt. Als erste Schätzung für das Optimierungsverfahren werden die Parameter über die L-Momente geschätzt. Eine anderes numerisches Lösungsverfahren für die Ermittlung der Parameter ist bei Rao und Hamed (2000) dargestellt.

8.4 Allgemeine Extremwertverteilung (AEV)

Parameter

$u, a > 0, k$

Verteilungsfunktion nach unten begrenzt für $k < 0$ mit dem unteren Grenzwert $x > u + a/k$

Verteilungsfunktion nach oben begrenzt für $k > 0$ mit dem oberen Grenzwert $x < u + a/k$

Verteilungsfunktion:

$$F(x) = \exp\left\{-\left[1 - k\left(\frac{x-u}{a}\right)\right]^{\frac{1}{k}}\right\}; \quad k \neq 0 \quad (2.82)$$

für $k = 0$ wird die AEV zu der 2-parametrischen Gumbelverteilung (siehe Kapitel 8.5):

$$F(x) = \exp\left\{-\exp\left[-\frac{1}{a}(x-u)\right]\right\}; \quad k = 0 \quad (2.83)$$

Wahrscheinlichkeitsdichtefunktion:

$$f(x) = \frac{1}{a} \left[1 - k\left(\frac{x-u}{a}\right)\right]^{-1/k-1} \exp\left\{-\left[1 - k\left(\frac{x-u}{a}\right)\right]^{\frac{1}{k}}\right\}; \quad k \neq 0 \quad (2.84)$$

für $k = 0$:

$$f(x) = \frac{1}{a} \exp \left[- \left(\frac{x-u}{a} \right) - \exp \left[- \left(\frac{x-u}{a} \right) \right] \right]; \quad k = 0 \quad (2.85)$$

Quantile:

$$x(F) = u + \frac{a}{k} \left[1 - (-\ln F)^k \right]; \quad k \neq 0 \quad (2.86)$$

für $k = 0$:

$$x(F) = u - a \ln(-\ln F); \quad k \neq 0 \quad (2.87)$$

Parameterschätzung Momenten-Methode:

Die Parameter der Verteilungsfunktion in Abhängigkeit der Momente ergeben sich zu:

$$c_s = \frac{k \left(-\Gamma(1+3k) + 3\Gamma(1+k)\Gamma(1+2k) - 2\Gamma^3(1+k) \right)}{|k| \left[\Gamma(1+2k) - \Gamma^2(1+k) \right]^{3/2}} \quad (2.88)$$

In EXANTO wird die Gl. (2.88) numerisch gelöst, um k zu ermitteln. Näherungslösungen für die Ermittlung von k sind z. B. in Rao und Hamed (2000) beschrieben.

 $k < 0$ für $C_s > 1.1396$ $k > 0$ für $0 < C_s < 1.1396$ $k > 0$ und $k < 0$ für $C_s < 0$ möglich

Für eine negative Schiefe bestehen zwei Möglichkeiten für k nach Gl. (2.88), ein Wert k größer 0 für eine nach oben begrenzte Verteilungsfunktion und ein Wert k kleiner 0 für eine nach unten begrenzte Verteilungsfunktion. Um den richtigen Wert für k zu ermitteln werden in EXANTO bei negativer Schiefe erst die Parameter nach der L-Momenten-Methode geschätzt, um zu ermitteln ob die Verteilungsfunktion nach oben oder unten begrenzt ist.

Die weiteren Parameter ergeben sich zu:

$$a = \frac{|k| \cdot \sqrt{\hat{m}_2}}{\sqrt{\Gamma(1+2k) - \Gamma^2(1+k)}} \quad (2.89)$$

$$u = m'_1 - \frac{a}{k} \left[1 - \Gamma(1+k) \right] \quad (2.90)$$

Parameterschätzung L-Momenten-Methode:

Die Parameter der Verteilungsfunktion in Abhängigkeit der L-Momente ergeben sich zu:

$$\frac{3+t_3}{2} = (1-3^{-k}) / (1-2^{-k}) \quad (2.91)$$

In EXANTO wird Gl. (2.91) numerisch gelöst um k zu ermitteln. In Hosking und Wallis

(1997) wird folgende Näherungslösung verwendet:

$$k \approx 7.8590c + 2.9554c^2 \quad (2.92)$$

mit

$$c = \frac{2}{3+t_3} - \frac{\ln 2}{\ln 3}. \quad (2.93)$$

Die anderen Parameter ergeben sich dann zu:

$$a = \frac{l_2 k}{(1-2^{-k})\Gamma(1+k)} \quad (2.94)$$

$$u = l_1 - \frac{a}{k} [1 - \Gamma(1+k)]. \quad (2.95)$$

Parameterschätzung Maximum-Likelihood-Methode:

In EXANTO werden die Parameter für die ML-Methode über ein numerisches Optimierungsverfahren (Numerisches Optimierungsverfahren Nonlinear Optimization Java Package der FPL Statistics Group (Verrill)) für die Maximierung der log-Likelihood-Funktion nach Gl. 2.45 ermittelt. Als erste Schätzung für das Optimierungsverfahren werden die Parameter über die L-Momente geschätzt. Eine andere Lösungsmethode für die Ermittlung der Parameter ist in Rao und Hamed (2000) dargestellt.

8.5 Gumbel-Verteilung (E1)

Parameter

u, a

$$-\infty < x < \infty$$

Verteilungsfunktion:

$$F(x) = \exp\left\{-\exp\left[-\frac{1}{a}(x-u)\right]\right\} \quad (2.96)$$

Wahrscheinlichkeitsdichtefunktion:

$$f(x) = \frac{1}{a} \exp\left[-\left(\frac{x-u}{a}\right) - \exp\left[-\left(\frac{x-u}{a}\right)\right]\right] \quad (2.97)$$

Quantile:

$$x(F) = u - a \ln(-\ln F) \quad (2.98)$$

Parameterschätzung Momenten-Methode:

Die Parameter ergeben sich zu:

$$a = \frac{\sqrt{6}}{\pi} \sqrt{\hat{m}_2} \quad (2.99)$$

$$u = m'_1 - \gamma a \quad (2.100)$$

mit Euler-Mascheroni-Konstante $\gamma = 0,5772156\dots$

Parameterschätzung L-Momenten-Methode:

Die Parameter ergeben sich zu:

$$a = \frac{l_2}{\ln(2)} \quad (2.101)$$

$$u = l_1 - \gamma a . \quad (2.102)$$

Parameterschätzung Maximum-Likelihood-Methode:

Die Parameter nach der Maximum Likelihood Methode (Herleitung siehe Rao & Hamed, 2000) ergeben sich zu:

$$F(a) = \sum_{i=1}^N x_i e^{-x_i/a} - \left(\frac{1}{N} \sum_{i=1}^N x_i - a \right) \sum_{i=1}^N e^{-x_i/a} = 0 \quad (2.103)$$

$$u = a \ln \left(\frac{N}{\sum_{i=1}^N x_i e^{-x_i/a}} \right) \quad (2.104)$$

Der Parameter a wird über Gl. (2.103) nach dem Newton-Verfahren iterativ ermittelt (siehe Rao & Hamed, 2000).

8.6 Weibull-Verteilung

Parameter

m, a>0, b>0

Verteilungsfunktion nach unten begrenzt mit unterem Grenzwert $x \geq m$

Verteilungsfunktion:

$$F(x) = 1 - \exp \left[- \left(\frac{x-m}{a} \right)^b \right] \quad (2.105)$$

Wahrscheinlichkeitsdichtefunktion:

$$f(x) = \frac{b}{a} \left(\frac{x-m}{a} \right)^{b-1} \exp \left[- \left(\frac{x-m}{a} \right)^b \right] \quad (2.106)$$

Quantile:

$$x(F) = m + a \left[-\ln(1-F) \right]^{1/b} \quad (2.107)$$

Parameterschätzung Momenten-Methode:

Die Parameter ergeben sich zu:

$$c_s = \frac{\Gamma(3/b+1) - 3\Gamma(1/b+1)\Gamma(2/b+1) + 2\Gamma^3(1/b+1)}{\left[\Gamma(2/b+1) - \Gamma^2(1/b+1) \right]^{3/2}} \quad (2.108)$$

Da die Weibull-Verteilung nur für $b > 0$ definiert ist, ergibt sich aus Gl. (2.108) der mögliche **Wertebereich** für die Schiefe der Verteilungsfunktion von **-1,1396** < C_s . Wenn die Stichprobe, aus der die Parameter der Verteilung geschätzt werden soll, eine andere Schiefe aufweist, kann b **nicht** über die Momentenmethode geschätzt werden. In EXANTO wird b nach Gl. (2.108) numerisch ermittelt. Aus numerischen Gründen liegt der Grenzwert in EXANTO bei **-1,138** < C_s . Weitere Lösungsverfahren über das Newton-Verfahren siehe Rao und Hamed (2000). Die weiteren Parameter ergeben sich zu:

$$a = \sqrt{\hat{m}_2} / \left[\Gamma(1+2/b) - \Gamma^2(1+1/b) \right]^{1/2} \quad (2.109)$$

$$m = m'_1 - a\Gamma(1+1/b). \quad (2.110)$$

Parameterschätzung L-Momenten-Methode:

Die Parameter ergeben sich zu:

$$\frac{3-t_3}{2} = (1-3^{-1/b}) / (1-2^{-1/b}) \quad (2.111)$$

Da die Weibull-Verteilung nur für $b > 0$ definiert ist, ergibt sich aus Gl. 2.111 der mögliche **Wertebereich** für die L-Schiefe t_3 der Verteilungsfunktion von

$$(3\ln(2) - 2 \cdot \ln(3)) / \ln(2) \approx -0.169925 < t_3.$$

In EXANTO wird der Parameter b numerisch über Gl. (2.111) ermittelt. Aus numerischen Gründen liegt der Grenzwert in EXANTO bei -0.16. Eine Näherungsverfahren ist in Rao und Hamed (2000) beschrieben.

Die anderen Parameter ergeben sich dann zu:

$$a = \frac{l_2}{(1-2^{-1/b})\Gamma(1+1/b)} \quad (2.112)$$

$$m = l_1 - a\Gamma(1+1/b). \quad (2.113)$$

Parameterschätzung Maximum-Likelihood-Methode:

In EXANTO werden die Parameter für die ML-Methode über ein numerisches Optimierungsverfahren (Numerisches Optimierungsverfahren Nonlinear Optimization Java Package der FPL Statistics Group (Verrill)) für die Maximierung der log-Likelihood-Funktion nach Gl. 2.45 ermittelt. Als erste Schätzung für das Optimierungsverfahren werden die Parameter über die L-Momente geschätzt. Ein anderes numerisches Lösungsverfahren für die Ermittlung der Parameter wird von Rao und Hamed (2000) dargestellt.

8.7 Log-Weibull-Verteilung**Parameter**

$m, a > 0, b > 0, x > 0$

Verteilungsfunktion nach unten begrenzt mit unterem Grenzwert $\ln(x) > m$

Verteilungsfunktion:

$$F(x) = 1 - \exp\left[-\left(\frac{\ln(x) - m}{a}\right)^b\right] \quad (2.114)$$

Wahrscheinlichkeitsdichtefunktion:

$$f(x) = \frac{b}{ax} \left(\frac{\ln(x) - m}{a}\right)^{b-1} \exp\left[-\left(\frac{\ln(x) - m}{a}\right)^b\right] \quad (2.115)$$

Quantile:

$$x(F) = \exp\left\{m + a\left[-\ln(1-F)\right]^{1/b}\right\} \quad (2.116)$$

Parameterschätzung Momenten-Methode:

Die Parameter werden mit den logarithmierten Werten der Stichprobe nach Gl. (2.108), (2.109) und (2.110) ermittelt.

Parameterschätzung L-Momenten-Methode:

Die Parameter werden mit den logarithmierten Werten der Stichprobe nach Gl. (2.111), (2.112) und (2.113) ermittelt.

Parameterschätzung Maximum-Likelihood-Methode:

In EXANTO werden die Parameter nach der ML-Methode über ein numerisches Optimierungsverfahren (Numerisches Optimierungsverfahren Nonlinear Optimization Java Package der FPL Statistics Group (Verrill)) für die Maximierung der log-Likelihood-Funktion nach Gl. 2.45 ermittelt. Als erste Schätzung für das Optimierungsverfahren werden die

Parameter über die L-Momente geschätzt. Eine anderes numerisches Lösungsverfahren für die Ermittlung der Parameter ist in Rao und Hamed (2000) dargestellt.

8.8 Frechet-Verteilung

Spezialfall der Allgemeinen Extremwertverteilung mit unterem Grenzwert 0

$$\rightarrow u = -\frac{a}{k}$$

Parameter

$$a > 0, k < 0$$

Verteilungsfunktion nach unten begrenzt mit unterem Grenzwert $0 < x < \infty$

Verteilungsfunktion:

$$F(x) = \exp \left\{ - \left[-\frac{kx}{a} \right]^{\frac{1}{k}} \right\} \quad (2.117)$$

Wahrscheinlichkeitsdichtefunktion:

$$f(x) = \frac{1}{a} \left(-\frac{kx}{a} \right)^{1/k-1} \exp \left[- \left(-\frac{kx}{a} \right)^{\frac{1}{k}} \right] \quad (2.118)$$

Quantile:

$$x(F) = -\frac{a}{k} (-\ln F)^k \quad (2.119)$$

Parameterschätzung Momenten-Methode:

Die Parameter der Verteilungsfunktion in Abhängigkeit der Momente ergeben sich zu:

$$c_v = \sqrt{\frac{\Gamma(1+2k)}{\Gamma^2(1+k)}} - 1 \quad (2.120)$$

In EXANTO wird die Gl. (2.120) numerisch gelöst, um k zu ermitteln.

$$a = \frac{|k| \cdot \sqrt{\hat{m}_2}}{\sqrt{\Gamma(1+2k) - \Gamma^2(1+k)}} \quad (2.121)$$

Parameterschätzung L-Momenten-Methode:

Die Parameter ergeben sich zu:

$$k = -\frac{\ln(t_2 + 1)}{\ln 2} \quad (2.122)$$

$$a = \frac{kl_2}{(1 - 2^{-k})\Gamma(1 + k)}. \quad (2.123)$$

Parameterschätzung Maximum-Likelihood-Methode:

In EXANTO werden die Parameter für die ML-Methode über ein numerisches Optimierungsverfahren (Numerisches Optimierungsverfahren Nonlinear Optimization Java Package der FPL Statistics Group (Verrill)) für die Maximierung der log-Likelihood-Funktion nach Gl. 2.45 ermittelt. Als erste Schätzung für das Optimierungsverfahren werden die Parameter über die L-Momente geschätzt. Ein anderes numerisches Lösungsverfahren für die Ermittlung der Parameter ist bei Bury (1999) dargestellt.

8.9 Pearson III-Verteilung (P III)

Parameter

$u, a \neq 0, b > 0$

Verteilungsfunktion nach unten begrenzt für $a > 0$ mit dem unteren Grenzwert $x > u$

Verteilungsfunktion nach oben begrenzt für $a < 0$ mit dem oberen Grenzwert $x < u$

Verteilungsfunktion:

$$F(x) = \frac{1}{\Gamma(b)} \int_0^{\frac{x-u}{a}} t^{b-1} e^{-t} dt = \frac{1}{\Gamma(b)} \gamma\left(b, \frac{x-u}{a}\right); \quad a > 0 \quad (2.124)$$

$$F(x) = 1 - \frac{1}{\Gamma(b)} \int_0^{\frac{x-u}{a}} t^{b-1} e^{-t} dt = 1 - \frac{1}{\Gamma(b)} \gamma\left(b, \frac{x-u}{a}\right); \quad a < 0 \quad (2.125)$$

mit der unvollständigen Gammafunktion der oberen Grenze:

$$\gamma(a, x) = \int_0^x t^{a-1} e^{-t} dt \quad (2.126)$$

Wahrscheinlichkeitsdichtefunktion:

$$f(x) = \frac{1}{a\Gamma(b)} \left(\frac{x-u}{a}\right)^{b-1} e^{-\left(\frac{x-u}{a}\right)}; \quad a > 0 \quad (2.127)$$

$$f(x) = -\frac{1}{a\Gamma(b)} \left(\frac{x-u}{a}\right)^{b-1} e^{-\left(\frac{x-u}{a}\right)}; \quad a < 0 \quad (2.128)$$

Quantile:

$$x(T) = \mu + \sigma \cdot k(T) \quad (2.129)$$

In EXANTO wird die Quantile numerisch aus Gl. (2.124) bzw. Gl. (2.125) ermittelt. Näherungslösungen für $k(T)$ sind z. B. in Rao und Hamed (2000) beschrieben.

Parameterschätzung Momenten-Methode:

Die Parameter der Verteilungsfunktion in Abhängigkeit der Momente der Stichprobe ergeben sich zu:

$$b = \frac{4}{c_s^2} \quad (2.130)$$

$$a = \frac{1}{2} \sqrt{\hat{m}_2} c_s \quad (2.131)$$

$$u = m_1' - 2 \frac{\sqrt{\hat{m}_2}}{c_s} \quad (2.132)$$

Parameterschätzung L-Momenten-Methode:

Die L-Momente der Pearson III Verteilung ergeben sich zu:

Parameter der Verteilungsfunktion in Abhängigkeit der L-Momente ergeben sich zu:

$$l_1 = u + ab \quad (2.133)$$

$$l_2 = \frac{1}{\sqrt{\pi}} a \frac{\Gamma(b+0.5)}{\Gamma(b)} \quad (2.134)$$

$$\tau_3 = 6I_{1/3}(b, 2b) - 3, \quad (2.135)$$

mit der relativen unvollständigen Betafunktion:

$$I_x(p, q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \int_0^x t^{p-1} (1-t)^{q-1} dt. \quad (2.136)$$

Der Schätzung des Parameters b wird in EXANTO aus dem L-Schiefekoeffizient t_3 der Stichprobe über Gl. (2.135) numerisch ermittelt. Eine Näherungslösung ist z.B. in Hosking und Wallis (1997) beschrieben. Die anderen Parameter ergeben sich aus den L-Momenten der Stichprobe zu:

$$a = \frac{l_2 \Gamma(b) \pi^{0.5}}{\Gamma(b+0.5)} \cdot \operatorname{sgn}|t_3| \quad (2.137)$$

$$u = l_1 - ab \quad (2.138)$$

Parameterschätzung Maximum-Likelihood-Methode:

In EXANTO werden die Parameter für die ML-Methode über ein numerisches Optimierungsverfahren (Numerisches Optimierungsverfahren Nonlinear Optimization Java Package der FPL Statistics Group (Verrill)) für die Maximierung der log-Likelihood-Funktion nach Gl. 2.45 ermittelt. Als erste Schätzung für das Optimierungsverfahren werden die Parameter über die L-Momente geschätzt. Eine andere Lösungsmethode für die Ermittlung der Parameter ist in Rao und Hamed (2000) dargestellt.

8.10 Log-Pearson III-Verteilung (Log-P III)

Parameter

$u, a \neq 0, b > 0, x > 0$

Verteilungsfunktion nach unten begrenzt für $a > 0$ mit dem unteren Grenzwert $x > u$

Verteilungsfunktion nach oben begrenzt für $a < 0$ mit dem oberen Grenzwert $x < u$

Verteilungsfunktion:

$$F(x) = \frac{1}{\Gamma(b)} \int_0^{\frac{\ln(x)-u}{a}} t^{b-1} e^{-t} dt = \frac{1}{\Gamma(b)} \gamma\left(b, \frac{\ln(x)-u}{a}\right); \quad a > 0 \quad (2.139)$$

$$F(x) = 1 - \frac{1}{\Gamma(b)} \int_0^{\frac{\ln(x)-u}{a}} t^{b-1} e^{-t} dt = 1 - \frac{1}{\Gamma(b)} \gamma\left(b, \frac{\ln(x)-u}{a}\right); \quad a < 0 \quad (2.140)$$

mit der unvollständigen Gammafunktion der oberen Grenze:

$$\gamma(a, x) = \int_0^x t^{a-1} e^{-t} dt \quad (2.141)$$

Wahrscheinlichkeitsdichtefunktion:

$$f(x) = \frac{1}{ax\Gamma(b)} \left(\frac{\ln(x)-u}{a}\right)^{b-1} e^{-\left(\frac{\ln(x)-u}{a}\right)}; \quad a > 0 \quad (2.142)$$

$$f(x) = -\frac{1}{ax\Gamma(b)} \left(\frac{\ln(x)-u}{a}\right)^{b-1} e^{-\left(\frac{\ln(x)-u}{a}\right)}; \quad a < 0 \quad (2.143)$$

Quantile:

$$x(T) = \mu + \sigma \cdot k(T) \quad (2.144)$$

In EXANTO wird die Quantile numerisch aus Gl. (2.124) bzw. Gl. (2.125) ermittelt. Näherungslösungen für $k(T)$ sind z. B. in Rao und Hamed (2000) beschrieben.

Parameterschätzung Momenten-Methode:

Die Parameter werden nicht indirekt über die logarithmierten Stichprobenwerte sondern direkt aus den Momenten der Zufallsvariable X . Der Parameter a wird in EXANTO in Abhängigkeit von den Ursprungsmomente m'_1 , m'_2 und m'_3 über folgende Gleichung numerisch geschätzt:

$$\frac{\ln m'_3 - 3 \ln m'_1}{\ln m'_2 - 2 \ln m'_1} = \frac{\ln \left[\frac{(1-a)^3}{(1-3a)} \right]}{\ln \left[\frac{(1-a)^2}{(1-2a)} \right]} \quad (2.145)$$

Die weiteren Parameter ergeben sich dann zu

$$b = \frac{\ln(1+c_v^2)}{\ln(1-a)^2 - \ln(1-2a)} \quad (2.146)$$

$$u = \ln m'_1 + b \ln(1-a) \quad (2.147)$$

Parameterschätzung L-Momenten-Methode:

Die Parameter werden mit den logarithmierten Werten der Stichprobe nach Gl. (2.135), (2.137) und (2.138) ermittelt.

Parameterschätzung Maximum-Likelihood-Methode:

In EXANTO werden die Parameter für die ML-Methode über ein numerisches Optimierungsverfahren (Numerisches Optimierungsverfahren Nonlinear Optimization Java Package der FPL Statistics Group (Verrill)) für die Maximierung der log-Likelihood-Funktion nach Gl. 2.45 ermittelt. Als erste Schätzung für das Optimierungsverfahren werden die Parameter über die L-Momente ermittelt. Eine andere Lösungsmethode für die Ermittlung der Parameter ist bei Rao und Hamed (2000) dargestellt.

8.11 Exponentialverteilung

Spezialfall der Pearson III Verteilung mit dem Parameter $b = 1$.

Parameter

u, a

Verteilungsfunktion nach unten begrenzt mit unterem Grenzwert $u < x < \infty$

Verteilungsfunktion:

$$F(x) = 1 - \exp \left[- \left(\frac{x-u}{a} \right) \right] \quad (2.148)$$

Wahrscheinlichkeitsdichtefunktion:

$$f(x) = \frac{1}{a} \exp \left[- \left(\frac{x-u}{a} \right) \right] \quad (2.149)$$

Quantile:

$$x(F) = u - a \ln(1-F) \quad (2.150)$$

Parameterschätzung Momenten-Methode:

Die Parameter ergeben sich zu:

$$a = \sqrt{\hat{m}_2} \quad (2.151)$$

$$u = m'_1 - a \quad (2.152)$$

Parameterschätzung L-Momenten-Methode:

Die Parameter ergeben sich zu:

$$a = 2l_2 \quad (2.153)$$

$$u = l_1 - 2l_2. \quad (2.154)$$

Parameterschätzung Maximum-Likelihood-Methode:

In EXANTO werden die Parameter für die ML-Methode über ein numerisches Optimierungsverfahren (Numerisches Optimierungsverfahren Nonlinear Optimization Java Package der FPL Statistics Group (Verrill)) für die Maximierung der log-Likelihood-Funktion nach Gl. 2.45 ermittelt. Als erste Schätzung für das Optimierungsverfahren werden folgende Parameter verwendet:

$$a = \frac{N(m'_1 - x_{\min})}{N-1} \quad (2.155)$$

$$u = \frac{Nx_{\min} - m'_1}{N-1}, \quad (2.156)$$

mit N der Größe der Stichprobe und dem kleinsten Wert x_{\min} der Stichprobe.

8.12 Gammaverteilung

Spezialfall der Pearson III Verteilung mit dem Parameter $u = 0$.

Parameter

$b > 0, a > 0$

Verteilungsfunktion nach unten begrenzt mit unterem Grenzwert $0 < x < \infty$

Verteilungsfunktion:

$$F(x) = \frac{1}{\Gamma(b)} \int_0^{x/a} x^{b-1} e^{-t} dt = \frac{1}{\Gamma(b)} \gamma\left(b, \frac{x}{a}\right) \quad (2.157)$$

Wahrscheinlichkeitsdichtefunktion:

$$f(x) = \frac{1}{a^b \Gamma(b)} x^{b-1} e^{-(x/a)} \quad (2.158)$$

Quantile:

$$x(T) = \mu + \sigma \cdot k(T) \quad (2.159)$$

In EXANTO wird die Quantile numerisch aus Gl. (2.157) ermittelt. Näherungslösungen für $k(T)$ sind z. B. in Rao und Hamed (2000) beschrieben.

Parameterschätzung Momenten-Methode:

Die Parameter ergeben sich zu:

$$a = \frac{\hat{m}_2}{m'_1} \quad (2.160)$$

$$b = \frac{m_1^2}{\hat{m}_2} \quad (2.161)$$

Parameterschätzung L-Momenten-Methode:

Die Parameter ergeben sich zu:

Für $0 < t_2 = l_2/l_1 < 0.5$, $z = \pi t_2^2$

$$b = (1 - 0.3080z) / (z - 0.05812z^2 + 0.01765z^3) \quad (2.162)$$

Für $0.5 \leq t_2 < 1$, $z = 1 - t_2$

$$b = (0.7213z - 0.5947z^2) / (1 - 2.1817z + 1.2113z^2) \quad (2.163)$$

$$a = \frac{l_1}{b} \quad (2.164)$$

Parameterschätzung Maximum-Likelihood-Methode:

In EXANTO werden die Parameter für die ML-Methode über ein numerisches Optimierungsverfahren (Numerisches Optimierungsverfahren Nonlinear Optimization Java Package der FPL Statistics Group (Verrill)) über die Maximierung der log-Likelihood-Funktion nach Gl. 2.45 ermittelt.

8.13 3-parametrische Allgemeine Pareto Verteilung

Parameter

u, a, k

Verteilungsfunktion nach unten begrenzt für $k \leq 0$: $u \leq x < \infty$

Verteilungsfunktion nach oben und unten begrenzt für $k > 0$: $u \leq x < u + a/k$

Verteilungsfunktion:

$$F(x) = 1 - \left[1 - \frac{k}{a}(x-u) \right]^{1/k} \quad (2.165)$$

für $k = 0$ reduziert sich die Allgemeine Pareto Gleichung zur Exponentialverteilung, für $k = 1$ reduziert sich die Allgemeine Pareto Gleichung zur Gleichverteilung $[u, u+a]$.

Wahrscheinlichkeitsdichtefunktion:

$$f(x) = \frac{1}{a} \left[1 - \frac{k}{a}(x-u) \right]^{1/k-1} \quad (2.166)$$

Quantile:

$$x(F) = u + \frac{a}{k} \left[1 - (1-F)^k \right] \quad (2.167)$$

Parameterschätzung Momenten-Methode:

Die Parameter ergeben sich zu:

$$c_s = \frac{2(1-k)(1+2k)^{0.5}}{(1+3k)} \quad (2.168)$$

In EXANTO wird k über Gl. (2.108) numerisch ermittelt.

Die weiteren Parameter ergeben sich zu:

$$a = \left[\hat{m}_2 (1+k)^2 (1+2k) \right]^{-0.5} \quad (2.169)$$

$$u = m'_1 - a/(1+k). \quad (2.170)$$

Bei gegebenen unterem Grenzwert u (z.B. x_{\min}) ergeben sich die Parameter zu:

$$k = \frac{1}{2} \left[(m'_1 - u)^2 / \hat{m}_2 - 1 \right] \quad (2.171)$$

$$a = (m'_1 - u)(1+k) \quad (2.172)$$

Parameterschätzung L-Momenten-Methode:

Die Parameter ergeben sich zu:

$$k = \frac{(1-3t_3)}{(1+t_3)} \quad (2.173)$$

$$a = l_2(1+k)(2+k) \quad (2.174)$$

$$u = l_1 - l_2(2+k). \quad (2.175)$$

Parameterschätzung Maximum-Likelihood-Methode:

In EXANTO werden die Parameter für die ML-Methode über ein numerisches Optimierungsverfahren (Numerisches Optimierungsverfahren Nonlinear Optimization Java Package der FPL Statistics Group (Verrill)) für die Maximierung der log-Likelihood-Funktion nach Gl. 2.45 ermittelt. Als erste Schätzung für das Optimierungsverfahren werden die Parameter über die Momenten-Methode mit dem kleinsten Stichprobenwert x_{\min} als untere Grenze u nach Gl. (2.171) und (2.172) ermittelt.

8.14 2-parametrigere Pareto Verteilung

Sonderfall der Allgemeinen Paretoverteilung mit unterem Grenzwert $u=0$

Parameter

a, k

Verteilungsfunktion nach unten begrenzt für $k \leq 0$: $0 \leq x < \infty$

Verteilungsfunktion nach oben und unten begrenzt für $k > 0$: $0 \leq x < a/k$

Verteilungsfunktion:

$$F(x) = 1 - \left[1 - \frac{k}{a}x \right]^{1/k} \quad (2.176)$$

für $k = 0$ reduziert sich die Allgemeine Paretoequation zur Exponentialverteilung, für $k = 1$ reduziert sich die Allgemeine Paretoequation zur Gleichverteilung $[u, u+a]$.

Wahrscheinlichkeitsdichtefunktion:

$$f(x) = \frac{1}{a} \left[1 - \frac{k}{a}x \right]^{1/k-1} \quad (2.177)$$

Quantile:

$$x(F) = \frac{a}{k} \left[1 - (1-F)^k \right] \quad (2.178)$$

Parameterschätzung Momenten-Methode:

Die Parameter ergeben sich zu:

$$k = \frac{1}{2} [m_1^2 / \hat{m}_2 - 1] \quad (2.179)$$

$$a = m_1' (1 + k) \quad (2.180)$$

Parameterschätzung L-Momenten-Methode:

Die Parameter ergeben sich zu:

$$k = l_1 / l_2 - 2 \quad (2.181)$$

$$a = (1 + k) l_1. \quad (2.182)$$

Parameterschätzung Maximum-Likelihood-Methode:

In EXANTO werden die Parameter für die ML-Methode über ein numerisches Optimierungsverfahren (Numerisches Optimierungsverfahren Nonlinear Optimization Java Package der FPL Statistics Group (Verrill)) für die Maximierung der log-Likelihood-Funktion nach Gl. 2.45 ermittelt. Als erste Schätzung für das Optimierungsverfahren werden die Parameter über die Momenten-Methode nach Gl. (2.171) und (2.172) ermittelt.

9 Verwendete externe Javabibliotheken und deren Urheberrechte

EXANTO ist vollständig in Java™ programmiert. Folgende externe Bibliotheken werden in EXANTO verwendet. Den Entwicklern dieser Bibliotheken wird an dieser Stelle gedankt. Dem US-amerikanischen Urheberrecht entsprechend werden z.T. umfangreiche Anmerkungen zu Haftungsausschlüssen und Urheberrechten gemacht, die aus rechtlichen Gründen nachfolgend aufgeführt sind. Der Nutzer von EXANTO erkennt die Nutzungsrechte und Garantiausschlüsse an.

9.1 A numerical library in Java for scientists and engineers (Lau, 2004)

Quelle: Lau, H. T. (2004): A numerical library in Java for scientists and engineers., Chapman & Hall.

Autoren: Hang T. Lau

Copyright Notice: © Copyrighth CHAPMAN & HALL/CRC PRESS 2004. All rights preserved.

9.2 FPL Statistics Group Nonlinear Optimization Java Package (Verrill, 2010)

Quelle: <http://www1.fpl.fs.fed.us/optimization.html>

DISCLAIMER OF WARRANTIES: THIS SOFTWARE IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND. THE AUTHOR/TRANSLATOR DOES NOT WARRANT, GUARANTEE OR MAKE ANY REPRESENTATIONS REGARDING THE SOFTWARE OR DOCUMENTATION IN TERMS OF THEIR CORRECTNESS, RELIABILITY, CURRENTNESS, OR OTHERWISE. THE ENTIRE RISK AS TO THE RESULTS AND PERFORMANCE OF THE SOFTWARE IS ASSUMED BY YOU. IN NO CASE WILL ANY PARTY INVOLVED WITH THE CREATION OR DISTRIBUTION OF THE SOFTWARE BE LIABLE FOR ANY DAMAGE THAT MAY RESULT FROM THE USE OF THIS SOFTWARE.

9.3 Colt Project (CERN, 2010)

Quelle: <http://acs.lbl.gov/~hoschek/colt/>

Copyright Notice: Packages cern.colt* , cern.jet*, cern.clhep

Copyright © 1999 CERN - European Organization for Nuclear Research.

Permission to use, copy, modify, distribute and sell this software and its documentation for any purpose is hereby granted without fee, provided that the above copyright notice appear in all copies and that both that copyright notice and this permission notice appear in supporting documentation. CERN makes no representations about the suitability of this software for any purpose. It is provided "as is" without expressed or implied warranty.

FAQ-Frequently Asked Questions

1. In den Ergebnisdateien wird für AIC und BIC „NaN“ ausgegeben!

Liegen Stichprobenwerte außerhalb des Definitionsbereiches der Verteilungsfunktion, ergebenergibt sich für diese Werte $f(x_i) = 0$ und somit die Likelihoodfunktion zu 0 bzw. die Log-Likelihoodfunktion zu $-\infty$. In EXANTO wird in diesem Fall „NaN“ in der Ergebnisdatei ausgegeben. Die Definitionsgrenzen der Verteilungsfunktionen sind in Kapitel 8 benannt..

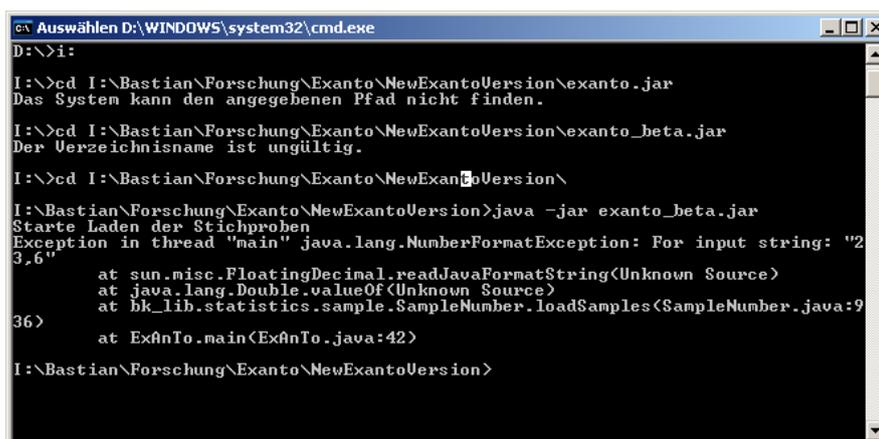
2. In den Ergebnisdateien wird für die Anpassungstests NaN ausgegeben!

Bei bestimmten Parameterwerten (z.B. k-Wert bei der AEV-Verteilung sehr groß) kann es vorkommen, dass die $F(x)$ und $f(x)$ aufgrund des großen Exponenten nicht mehr berechnet werden können, in diesem Fall wird NaN ausgegeben.

3. Programm endet mit der Ausgabe: "Falsche Bezeichnung für die Parameterschätzmethode in der Kontrolldatei!!" oder "Falsche Bezeichnung für die Verteilungsfunktionen in der Kontrolldatei!!"

Kontrolldatei überprüfen ob die Abkürzungen für die Verteilungsfunktionen richtig eingetragen sind (!!! Unbedingt auf Groß- und Kleinschreibung achten !!!)

4. Folgender Fehler:



```
ex Auswählen D:\WINDOWS\system32\cmd.exe
D:\>i:
I:\>cd I:\Bastian\Forschung\Exanto\NewExantoVersion\exanto.jar
Das System kann den angegebenen Pfad nicht finden.
I:\>cd I:\Bastian\Forschung\Exanto\NewExantoVersion\exanto_beta.jar
Der Verzeichnisname ist ungültig.
I:\>cd I:\Bastian\Forschung\Exanto\NewExantoVersion\
I:\Bastian\Forschung\Exanto\NewExantoVersion>java -jar exanto_beta.jar
Starte Laden der Stichproben
Exception in thread "main" java.lang.NumberFormatException: For input string: "23,6"
    at sun.misc.FloatingDecimal.readJavaFormatString(Unknown Source)
    at java.lang.Double.valueOf(Unknown Source)
    at bk_lib.statistics.sample.SampleNumber.loadSamples(SampleNumber.java:9)
    at ExAnTo.main(ExAnTo.java:42)
I:\Bastian\Forschung\Exanto\NewExantoVersion>
```

Das Dezimaltrennzeichen in der Stichprobendatei ist Komma und nicht Punkt → in den Systemeinstellungen anpassen (unter Windows XP: Start→Einstellungen→Systemsteuerung→Regions-und Sprachoptionen→Regionale Einstellungen→anpassen...→Dezimaltrennzeichen anpassen) Punkt als Dezimaltrennzeichen einstellen.

Literaturverzeichnis

- Akaike, H. (1974): A New Look at Statistical Model Identification. *Ieee Transactions on Automatic Control* AC19(6), 716-723
- Bacchi, B., A. Brath & N. T. Kottegoda (1992): Analysis of the Relationships Between Flood Peaks and Flood Volumes Based on Crossing Properties of River Flo Processes. *Water Resources Research* 28(10), 2773-2782
- Bertie, A. (2010). "JSC Java Statistical Classes." from <http://www.jsc.nildram.co.uk/about.html>.
- Burn, D. H. (2003): The use of resampling for estimating confidence intervals for single site and pooled frequency analysis. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques* 48(1), 25-38
- Bury, K. (1999): *Statistical Distributions in Engineering*, Cambridge University Press, Cambridge.
- CERN. (2010). "Colt Project." from <http://acs.lbl.gov/~hoschek/colt/>.
- DVWK (1999): *Statistische Analyse von Hochwasserabflüssen*. DVWK-Merkblatt Heft 251, Kommissionsvertrieb Paul Parey, Hamburg und Berlin.
- Greenwood, J. A., J. M. Landwehr, N. C. Matalas & J. R. Wallis (1979): Probability Weighted Moments: Definition and Relation to Parameters of Several Distributions Expressible in Inverse Form. *Water Resources Research* 15(5), 1049-1054
- Griffis, V. W. & J. R. Stedinger (2007): Evolution of flood frequency analysis with Bulletin 17. *Journal of Hydrologic Engineering* 12(3), 283-297
- Hall, M. J., H. F. P. van den Boogaard, R. C. Fernando & A. E. Mynett (2004): The construction of confidence intervals for frequency analysis using resampling techniques. *Hydrology and Earth System Sciences* 8(2), 235-246
- Hicklin, J., C. Moler, P. Webb, R. F. Boisvert, B. Miller, R. Pozo & K. Remington. (2010). "JAMA : A Java Matrix Package." from <http://math.nist.gov/javanumerics/jama/>.
- Hosking, J. R. M. (1990): L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society Series B-Methodological* 52(1), 105-124
- Hosking, J. R. M. & J. R. Wallis (1997): *Regional frequency analysis: an approach based on L-moments*, Cambridge University Press, Cambridge.
- IACWD (1982): *Guidelines for determining flood flow frequency*. Bulletin No. 17B (revised and corrected), Interagency Committee on Water Data, Washington, D.C.
- Landwehr, J. M., N. C. Matalas & J. R. Wallis (1979): Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. *Water Resources Research* 15(5), 1055-1064
- Lau, H. T. (2004): *A numerical library in Java for scientists and engineers.*, Chapman & Hall.
- LAWA (1997): *Pegelvorschrift-Stamtext*, Kulturbuchverlag Berlin GmbH, Berlin.
- Lee, E. A. & C. Brooks (2010): *Ptplot 5.7 - Java Plotter*.
- Makkonen, L. (2008): Bringing closure to the plotting position controversy. *Communications in Statistics-Theory and Methods* 37(3), 460-467

- Plate, E. J. (1993): Statistik und angewandte Wahrscheinlichkeitslehre für Bauingenieure, Verlag Ernst & Sohn, Berlin.
- Rao, A. R. & K. H. Hamed (2000): Flood Frequency Analysis, CRC Press LLC.
- Raschke, M. & K. Thürmer (2008): Defizite der Modellselektion in der Hochwasserstatistik. Wasser und Abfall 12, 43-48
- Salvadori, G., C. De Michele, N. T. Kottegoda & R. Rosso (2007): Extremes in Nature: An Approach Using Copulas. Water Science and Technology Library, 56, Springer, Dordrecht.
- Schwarz, G. (1978): Estimating the dimension of a model. The Annals of Statistics 6(2), 461-464
- Stedinger, J. R. & V. W. Griffis (2008): Flood frequency analysis in the United States: Time to update. Journal of Hydrologic Engineering 13(4), 199-204
- Stedinger, J. R., R. M. Vogel & E. Foufoula-Georgiou (1993): Frequency Analysis of Extreme Events. In: D. R. Maidment (Ed.): Handbook of Hydrology, McGraw-Hill Inc., New York.
- Verrill, S. (2010). "Nonlinear Optimization Java Package." from <http://www1.fpl.fs.fed.us/optimization.html>.
- Vogel, R. M. & I. Wilson (1996): Probability Distribution of Annual Maximum, Mean, and Minimum Streamflows in the United States. Journal of Hydrologic Engineering 1(2), 69-76