

Systematische Untersuchung von Spektrums-Fits in der Gammaastronomie

Masterarbeit aus der Physik

vorgelegt von

Andreas Horlbeck

Tag der Abgabe: 28.10.2014

Erlangen Centre for Astroparticle Physics
Physikalisches Institut
Friedrich-Alexander-Universität
Erlangen-Nürnberg



1. Gutachter: Prof. Dr. Christopher van Eldik
2. Gutachter: Prof. Dr. Gisela Anton

Kurzzusammenfassung

Ziel dieser Arbeit ist die Untersuchung und Verbesserung von statistischen Schätzmethoden angewandt auf die Verteilung, die man mit einem Potenzgesetz mit exponentiellem Cutoff beschreibt. Der Verlauf der Verteilung lässt sich dabei mit zwei Parameter beschreiben, dem spektralen Index und dem Cutoff-Parameter. Energiespektren von Gammastrahlungsquellen, die bei hohen Energien ein vom Potenzgesetz abweichendes Verhalten zeigen, also steiler abfallen, lassen sich beispielsweise mit solch einer Verteilung beschreiben. Im ersten Teil der Arbeit wurden die Schätzfunktionen für die beiden Parameter, die sich aus der erweiterten Maximum-Likelihood-Methode ergeben, auf ihre Schätzqualität untersucht. Diese Methode wurde auf ungebinned und gebinnte Daten angewandt. Hier zeigte sich nur für kleine Stichprobengrößen, dass die gebinnte Version bessere Ergebnisse erzielt. Für große Stichproben zeigen sich nur vernachlässigbare Unterschiede. Für die ungebinnete Version wurde eine alternative Methode entwickelt, die auf der Ersetzungsmethode beruht. Hier zeigt sich für die Schätzfunktion für den exponentiellen Cutoff für kleine Stichprobengrößen ein besseres Verhalten, im Sinne von kleineren Werten des mittleren quadratischen Fehlers. Alle vorgestellten Methoden liefern konsistente Schätzfunktionen. Im zweiten Teil der Arbeit wurden weitere Untersuchungen anhand der erweiterten Maximum-Likelihood-Methode, angewandt auf gebinnte Daten, durchgeführt. Zunächst wurde eine Korrektur durchgeführt, sodass der systematische Bias für die Schätzfunktion für den exponentiellen Cutoff kleiner ausfällt. Des Weiteren wurden effektive Fläche und Untergrundsignale in der Analyse berücksichtigt, hier zeigen sich wiederum konsistente Schätzfunktionen. Abschließend wurden Maximum-Verhältnis-Tests vorgestellt, mit denen man zum Beispiel statistisch entscheiden kann, ob sich eine Datenreihe hinreichend durch ein Potenzgesetz beschreiben lässt oder die Erweiterung durch einen exponentiellen Cutoff notwendig ist.

Abstract

The goal of this thesis is the examination and improvement of statistical estimation methods used on distributions, that are described by a power law with an exponential cut-off. The shape of this distribution depends on two parameters, the spectral index and the cut-off parameter. An example for that kind of distribution is the energy spectrum of some gamma ray sources, that show a deviation from a power law at high energies, where it declines steeper. In the first part of the study, the extended maximum-likelihood method is tested for its estimation quality for the mentioned parameters. To do this, it was used on binned and unbinned datasets. For small samples, the binned version leads to better results, while the differences are negligible for bigger samples. An alternate method, based on the method of moments, was developed for the unbinned data. The estimator for the cut-off parameter shows a better behaviour for smaller samples, resulting in a smaller mean error. All presented methods lead to consistent estimators. The second part of the thesis consists of further studies, using the extended maximum-likelihood method on binned data. First, a correction to decrease the systematic bias of the cut-off estimator was performed. On top of that, effective areas and background signals were taken into account, which resulted in consistent estimators again. At last maximum-ratio-tests were introduced, that, for example, allows for testing statistically, if a dataset can be described by a power law properly or if the extension with an exponential cut-off is necessary.

Inhaltsverzeichnis

1 Einleitung und Motivation	6
2 Wichtige Begriffe aus der Statistik	9
2.1 Statistisches Modell	9
2.2 Parametrische Schätzfunktionen und Schätzprinzipien	9
2.3 Kriterien geeigneter Schätzfunktionen	12
2.4 Hypothesentests und Likelihood-Verhältnis-Methode	13
3 Erweiterte Maximum-Likelihood-Methode: Ungebinnte Version	15
3.1 Prinzip der erweiterten Maximum-Likelihood-Methode bei ungebinnten Daten	15
3.2 Untersuchung der Schätzfunktionen	17
4 Erweiterte Maximum-Likelihood-Methode: Gebinnte Version	25
4.1 Prinzip der erweiterten Maximum-Likelihood-Methode bei gebinnten Daten	25
4.2 Untersuchung der Schätzfunktionen	27
4.3 Vergleich der erweiterten Maximum-Likelihood-Methode hinsichtlich ungebinnter und gebinnter Daten	28
5 Ersetzungsmethode	29
6 Vergleich der erweiterten Likelihood-Methode und der Ersetzungsmethode	50
7 Verbesserung der Schätzfunktion $\hat{\lambda}$	54
8 Berücksichtigung der effektiven Fläche beim Schätzen der Parameter	61
9 Erweiterte Maximum-Likelihood-Methode unter Berücksichtigung von Untergrundergebnissen	66
10 Hypothesentests	77
10.1 Test auf reines Potenzgesetz gegen Potenzgesetz mit exponentiellem Cutoff	77
10.2 Test auf Potenzgesetz gegen beliebige Verteilung	82
11 Zusammenfassung und Ausblick	85

1 Einleitung und Motivation

In vielen Bereichen der Wissenschaft zeigt sich ein Interesse am Verständnis von Wahrscheinlichkeitsverteilungen, die sich durch ein *Potenzgesetz* beschreiben lassen. Dazu gehören unter anderem die Physik, die Geowissenschaften, die Sozialwissenschaften und das Versicherungswesen. So lassen sich beispielsweise durch Potenzgesetze die Intensitäten von Erdbeben und die Bevölkerungsgrößen von Städten beschreiben [Clauset et al., 2007].

Für die Astrophysik haben Potenzgesetzverteilungen einen besonderen Stellenwert. Sie dienen zum Beispiel der Beschreibung der Energieverteilung der kosmischen Strahlung, die hauptsächlich aus Protonen und nur zu einem geringen Anteil aus Elektronen und schwereren Kernen besteht [Longair, 2011]. Als mögliche Quellen für kosmische Strahlung und damit auch Gammastrahlung kommen Supernovea und ihre Überreste, schnell rotierende Objekte wie aktive galaktische Kerne und Schwarze Löcher in Frage. Der Gammastrahlung ordnet man Energien größer als 100 keV zu. Nach oben ist die Gammastrahlung in ihrer Energie nicht begrenzt. Man hat schon kosmische Gammaquanten mit Energien von 1 PeV gemessen [Grupen, 2000]. Die Gammastrahlung kann dabei mittels mehrerer möglicher Szenarien entstehen. Eines dieser Szenarien ist der inverse Compton-Effekt, bei dem Elektronen an Photonen streuen und so Energie übertragen. Ein anderes mögliches Szenario ist der Zerfall von neutralen Pionen in Gammaquanten, die zuvor aus der Wechselwirkung von Protonen der kosmischen Strahlung mit Kernen des umgebenden interstellaren Mediums, z.B. aus Molekülwolken, entstehen. So ergeben sich abhängig vom Entstehungsmechanismus unterschiedliche zugehörige Energiespektren.

Diese Energiespektren, die man von Gammastrahlungsquellen misst, werden gewöhnlich durch ein Potenzgesetz beschrieben, die dann abhängig vom Entstehungsmechanismus durch andere zugehörige Parameterwerte des Potenzgesetzes beschrieben werden:

$$\frac{dN}{dE} = \Phi_0 \cdot \left(\frac{E}{1TeV} \right)^{-\alpha}$$

Der Parameter Φ_0 ist die sogenannte Flussnormierung und α ist der spektrale Index des Potenzgesetzes. Abhängig vom Energiebereich und der Art der Gammastrahlungserzeugung treten Abweichungen vom reinen Potenzgesetz auf, die unter anderem Rückschlüsse auf den Produktionsprozess und die zugrundeliegende Teilchenpopulation zulassen. Im Bereich sehr hochenergetischer Gammastrahlung, also im Bereich von 100 GeV bis zu 100 TeV, wurden in den letzten Jahren durch Instrumente wie H.E.S.S über 100 neue Gammastrahlungsquellen entdeckt. Für einige dieser Quellen reicht die Photonenstatistik aus, um deutliche Abweichungen vom reinen Potenzgesetz, insbesondere bei den höchsten Energien zu sehen [Aharonian et al., 2006]. Das H.E.S.S.-Experiment besteht aus einem in Namibia stationierten Teleksop-Array, das mit bodengebundener Gammaastronomie kosmische Gammastrahlung im Energiebereich von einigen 10 GeV bis zu einigen 10 TeV misst [HES,]. Als Beispiel hat das H.E.S.S.-Experiment bei der Untersuchung der Quelle HESS-J1745-290, siehe Abbildung 1, ein Energiespektrum gemessen, das eine Abweichung vom Potenzgesetzverhalten bei hohen Energien zeigt. Es gibt hier im Vergleich zum Potenzgesetz einen schnelleren Abfall des Spektrums bei hohen Energien. Eine sinnvolle Erweiterung, die dieses Spektrum mathematisch modelliert, ist ein „exponentieller Cutoff“. Das Potenzgesetz lässt sich dabei folgendermaßen durch eine Exponentialfunktion und einen Cutoff-Parameter E_c erweitern [Aharonian et al., 2009]:

$$\frac{dN}{dE} = \Phi_0 \cdot \left(\frac{E}{1TeV} \right)^{-\alpha} \cdot e^{-\frac{E}{E_c}}$$

E_c ist der sogenannte Cutoff-Parameter bzw. die Cutoff-Energie. Die Exponentialfunktion hat bei niedrigen Energien kaum Einfluss, bei hohen Energien fällt das Spektrum aber immer steiler

ab. Der Cutoff-Parameter regelt, ab welchen Energiewerten dieses Verhalten einsetzt. Der Wert ist so definiert, dass beim Wert der Cutoff-Energie die Exponentialfunktion den Wert $1/e$ annimmt. Die Abbildung 1 zeigt dabei das gemessene Spektrum der Quelle HESS-J1745-290 und die Anpassung des Potenzgesetztes mit exponentiellem Cutoff an den gemessenen Fluss.

Für das Zustandekommen eines Cutoffs gibt es im Allgemeinen mehrere mögliche physikalische Erklärungen. Eine der Erklärungen ist, dass es sich um eine intrinsische Eigenschaft der Gammastrahlungsquelle handelt. Die Energien der Protonen bzw. die Energien, auf die die Protonen beschleunigt werden, können durch die Quelle selbst begrenzt sein. Ein Beispiel hierfür ist wiederum die galaktische Quelle HESS-J1745-290. Hier vermutet man einen Cutoff im primären Protonenspektrum.

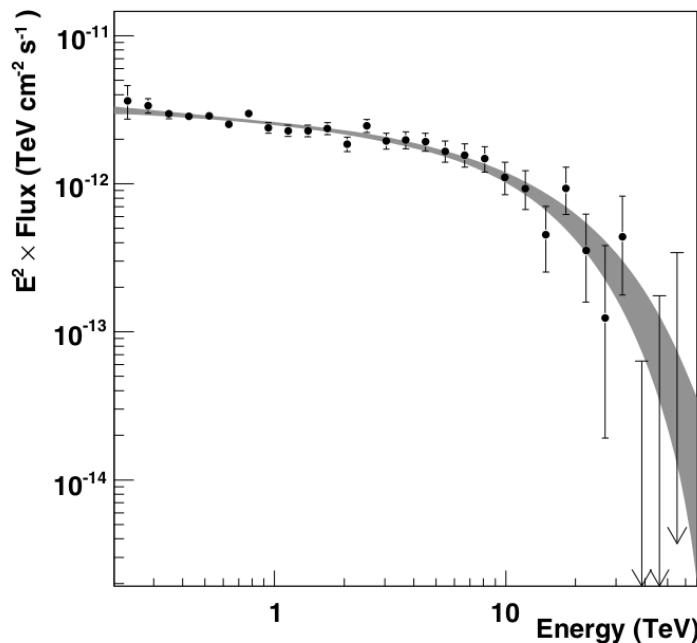


Abbildung 1: Gemessenes Spektrum der HESS-Quelle J1745-290 in doppelt-logarithmischer Auftragung [Aharonian et al., 2009]. Der Fluss wurde zur besseren Darstellung mit E^2 multipliziert. Man sieht, dass die Datenpunkte für kleine Energien auf einer Geraden liegen - in dieser Auftragung einem Potenzgesetz entsprechend - für hohe Energien dann allerdings von dem Verlauf der Geraden abweichen und stark abfallen. Dies lässt sich mit dem eingezeichneten Fit, siehe graues Band, beschreiben.

Eine weitere mögliche Erklärung wäre, dass die kosmische Strahlung weitaus energiereichere Gammastrahlung produziert, aber durch das extragalaktische Hintergrundlicht energieabhängig absorbiert wird und so das zugehörige Spektrum einen Cutoff erfährt. Beispiel hierfür bieten extragalaktische Quellen wie Blazare, siehe [H.E.S.S. Collaboration et al. \(2013\)](#).

Aus den verschiedenen Erklärungsmöglichkeiten für das Zustandekommen eines Cutoffs, gibt dieser auch Aufschluss auf das primäre Protonenspektrum bzw. auf die Absorption durch das extragalaktische Hintergrundlicht.

Quellen, deren Gammastrahlung durch die von ihnen beschleunigten kosmischen Strahlung verursacht wird, können so über das Messen der Gammastrahlung auch Auskunft über die kosmische Strahlung geben. Im Vergleich zur kosmischen Strahlung im Energiebereich um einige TeV, die aufgrund ihrer Ladung auf den Weg zur Erde signifikant von galaktischen Magnetfeldern abgelenkt wird, hat Gammastrahlung den Vorteil, dass die Bewegung ihrer ungeladenen Gammaquanten in guter Näherung auf den Ort ihrer Entstehung zurück zeigt [[Grupen, 2000](#)]. So lässt

sich letztlich mit dem Wissen um die Parameter, die das Energiespektrum der Gammastrahlung beschreiben, auch auf die zugrundliegenden physikalischen Mechanismen ihrer Quellen schließen. Die Schwierigkeit bei der Bestimmung des Cutoff-Parameters ist allerdings, dass Ereignisse mit hohen Energiewerten im Bereich, in dem das Spektrum vom Potenzgesetz abweicht, nur in geringer Zahl zu erwarten sind. Aufgrund der sich damit ergebenden geringen Statistik bedarf es guter Schätzmethoden, um diesen Cutoff-Parameter dennoch ohne systematische Verzerrung und mit möglichst kleinem statistischem Fehler bestimmen zu können.

Grundlegendes Ziel dieser Arbeit ist es, Methoden der Gammaastronomie zur Bestimmung der Parameter des Energiespektrums α und E_c , vor allem des Cutoff-Parameters E_c , zu untersuchen und zu beurteilen. Der nötige mathematische Formalismus wird dazu in Kapitel 2 vorbereitend zusammengefasst. Dazu betrachtet man zunächst ein Signal, das lediglich aus Gammaphotonen besteht. Hier werden zunächst die Schätzfunktionen für die Parameter, die sich aus der erweiterten Maximum-Likelihood-Methode ergeben, hinsichtlich ihres Schätzverhaltens untersucht. Dies wird zunächst für gebinnte Daten und anschließend auch für ungebinnte Daten durchgeführt. Des Weiteren wird ein alternatives neues Schätzprinzip vorgestellt, das auf der Ersetzungsmethode basiert und für ungebinnte Daten konzipiert ist. Nach der Untersuchung dieser Methode hinsichtlich ihres Schätzverhaltens für die Parameter werden die beiden ungebinnten Methoden miteinander verglichen.

Die nachfolgenden Untersuchungen werden mittels der gebinnten Version der erweiterten Maximum-Likelihood-Methode durchgeführt. So folgen weitere Untersuchungen zur Verbesserung der Schätzmethoden und der Untersuchung der Schätzmethoden bei Berücksichtigung der effektiven Fläche und Untergundsignalen. Abschließend wird ein statistischer Test zur Untersuchung vorgestellt, ob sich das Energiespektrum eines wiederum reinen Gammastrahlungssignals durch ein Potenzgesetz hinreichend beschrieben lässt.

2 Wichtige Begriffe aus der Statistik

Ziel der vorliegenden Arbeit ist vor allem die Untersuchung, Bewertung und den Vergleich der verwendeten Schätzmethoden. Die allgemeine Aufgabe dieser Schätzmethoden ist es, anhand von empirischen Daten, hier Messwerte der Energien der kosmischen Gammastrahlung, die zufallsgesteuert sind, also sich mit Zufallsvariablen beschreiben lassen, Aussagen über die zugrundeliegenden Parameter des Energiespektrums treffen zu können. Innerhalb dieser Arbeit wird auch die Fragestellung beantwortet, wie man statistisch entscheiden kann, ob sich die Verteilung der Messwerte hinreichend mit einem Potenzgesetz beschreiben lässt oder man eine andere Verteilung benutzt sollte, wie insbesondere die Potenzgesetzverteilung mit exponentiellem Cutoff. Dazu wird ein entsprechender Hypothesentest konstruiert, der in diesem Kapitel ebenso allgemein erklärt wird.

Dazu werden in diesem Kapitel zunächst einige wichtige allgemeine Begriffe und allgemeine Methoden aus der Statistik definiert und erklärt werden, die für die statistischen Aufgaben in den folgenden Kapiteln benötigt werden. Das folgende Kapitel beruht auf dem Buch [Georgii \(2009\)](#), insbesondere Kapitel 7 und 10.

2.1 Statistisches Modell

Bevor man Schätzmethoden entwickelt, gilt es zunächst ein statistisches Modell \mathcal{M} aufzustellen. Ein statistisches Modell ist allgemein ein Tripel $\mathcal{M} = (\mathcal{X}, \mathcal{F}, \mathcal{P}_\theta : \theta \in \Theta)$, bestehend aus einem Stichprobenraum \mathcal{X} , der alle möglichen Beobachtungsergebnisse x in einer Menge zusammenfasst, einer Menge \mathcal{F} , dem Ereignisraum und einer mindestens zweielementigen Klasse von Wahrscheinlichkeitsmaßen $\{\mathcal{P}_\theta : \theta \in \Theta\}$, definiert auf $(\mathcal{X}, \mathcal{F})$, die mit einer Indexmenge Θ indiziert ist. Ein einzelnes Wahrscheinlichkeitsmaß ist dabei als Abbildung $\mathcal{P} : \mathcal{F} \rightarrow [0, 1]$ definiert, die jedem Ereignis, sprich Element von \mathcal{F} , eine Wahrscheinlichkeit, also einen Zahlenwert zwischen 0 und 1 zuordnet und $\mathcal{P}_\theta(\mathcal{X}) = 1, \forall \theta \in \Theta$ gilt.

Allerdings steht dieses statistische Modell für ein Einzelexperiment. Im Folgenden möchten wir aber die n-fache unabhängige Wiederholung dieser identischen Einzelexperimente beschreiben, und so ergibt sich schließlich ein n-Tupel (X_1, \dots, X_n) von Zufallsvariablen, die nun jeweils mit einer Zahl zwischen 1 und n beschrieben werden, um so die Teilexperimente durchzunummerieren und sie so unterscheiden zu können. n bezeichne in dieser Arbeit eine beliebige natürliche Zahl mit $n \geq 1$ die immer als Parameter für eine Stichprobengröße zu verstehen ist.

Das zugehörige statistische Modell \mathcal{M}^* , das die n-fache unabhängige Wiederholung des Einzel-experiments beschreibt, ist dann als das Produktmodell des vorherigen statistischen Modells definiert: $\mathcal{M}^* = (\mathcal{X}^n, \mathcal{F}^{\otimes n}, \mathcal{P}_\theta^{\otimes n} : \theta \in \Theta)$. \mathcal{X}^n ist dabei das n-fache kartesische Produkt von \mathcal{X} . Die Zufallsvariable X_i bezeichne dann die Projektion $X_i : \mathcal{X}^n \rightarrow \mathcal{X}$, also den Ausgang des i-ten Teilexperiments. $\mathcal{F}^{\otimes n}$ bezeichne wiederum den zugehörigen Ereignisraum. Wegen der geforderten Unabhängigkeit der X_1, \dots, X_n bezüglich jedes \mathcal{P}_θ faktorisiert das Produktwahrscheinlichkeitsmaß: $\mathcal{P}_\theta^{\otimes n}(X_1, \dots, X_n) = \mathcal{P}_\theta(X_1) \cdot \dots \cdot \mathcal{P}_\theta(X_n)$.

Ferner sei hier noch erwähnt, dass eine zufällige Beobachtung bzw. Stichprobe beschrieben durch Zufallsvariablen X_i mit einem Großbuchstaben bezeichnet wird, eine tatsächliche Realisation bzw. konkreter Messwert x_i der Zufallsvariablen werde mit einem Kleinbuchstaben bezeichnet.

2.2 Parametrische Schätzfunktionen und Schätzprinzipien

Als eines der Vorgehen einer statistischen Analyse schätzt man die Parameter der unbekannten Verteilung \mathcal{P}_θ . Das Ergebnis dieser Schätzung ist dann jeweils ein einzelner Wert für jeden dieser Parameter. Mit dem Begriff „Schätzen“ ist gemeint, dass man anhand der Beobachtungen einen

Wert für die Parameter angibt, von denen man ausgeht, dass sie den wahren, typischerweise unbekannten Wert des Parameters möglichst genau treffen.

Dazu konstruiert man eine sogenannte Schätzfunktion T , im Folgenden Schätzer genannt. Sei ferner τ eine bijektive Abbildung $\tau : \Theta \rightarrow \Sigma$, die jedem $\theta \in \Theta$ eine Kenngröße $\tau(\theta) \in \Sigma$ zuordnet. Der Schätzer T ist dann eine Abbildung vom Stichprobenraum \mathcal{X} in den Raum Σ , $T : \mathcal{X} \rightarrow \Sigma$ und damit ein Schätzer für $\tau(\theta)$. Wichtig sei hier noch zu erwähnen, dass es sich bei dem Schätzer $T_n(X_1, \dots, X_n)$ selbst um eine Zufallsvariable handelt, da sein Wert wiederum von Zufallsvariablen abhängt. Typischerweise werden Schätzfunktionen auch mit einem „ $\hat{\cdot}$ “ -Symbol auf dem Parameter, den sie schätzen, gekennzeichnet, beispielsweise $\hat{\tau}(\theta)$.

Gegebenenfalls lässt sich $\hat{\tau}(\theta)$ nach θ analytisch auflösen oder der Wert daraus numerisch bestimmen, um so auch einen Schätzwert für θ angeben zu können.

Es werden im Folgenden einige wichtige Schätzfunktionen genannt, die sich alle von zwei wichtigen Prinzipien ableiten lassen, mit dessen Hilfe man allgemein Schätzfunktionen herleiten kann.

Ersetzungsmethode

Jede Zufallsvariable X , allgemein definiert auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathcal{P}_\theta)$, $\theta \in \Theta$, also für ein einzelnes vorgegebenes Wahrscheinlichkeitsmaß \mathcal{P}_θ , lässt sich bezüglich ihrer Verteilung \mathcal{P}_X^θ und ihrer zugehörigen Verteilungsfunktion F_X^θ mit l , $l \in \mathbb{N}$, verschiedenen theoretischen Kenngrößen $K^j(\theta)$ beschreiben. j bezeichne dabei die j -te von l Kenngrößen.

Der Ansatz der Ersetzungsmethode ist es, dass man diese Kenngrößen der Verteilung mit den entsprechenden empirischen Kenngrößen der Stichprobe ersetzt. Eine empirische Kenngröße $K_{emp}^j(X_1, \dots, X_n)$ ist dabei als eine Zufallsvariable auf einer Stichprobe (X_1, \dots, X_n) der Zufallsvariable X definiert.

Allgemein würde das wiederum so aussehen: Man ersetzt $(, \rightarrow)$ die empirische Größe mit der entsprechenden theoretischen Kenngröße und löst nach dem Wert für $\theta \in \Theta$ auf, die ebendiesem Wert entspricht.

$$\begin{aligned} K_{emp}^j(x_1, \dots, x_n) &\rightarrow K^j(\gamma) \\ K_{emp}^j(x_1, \dots, x_n) = \hat{\tau}(\theta) &\Rightarrow \hat{\theta} = \tau^{-1}(K_{emp}^j(x_1, \dots, x_n)) \end{aligned}$$

Die wichtigsten theoretischen Kenngrößen einer Zufallsvariable X und dazugehörigen empirischen Kenngrößen einer Stichprobe (X_1, \dots, X_n) in Abhängigkeit von \mathcal{P}_θ werden im Folgenden aufgelistet.

Theoretische Kenngrößen:

- **Erwartungswert** $\mathbb{E}_\theta(X)$:

allgemein: $\mathbb{E}_\theta(X) = \int_\Omega X(\omega) d\mathcal{P}_\theta$

für X stetig: $\mathbb{E}_\theta(X) = \int_{\mathbb{R}} x f_\theta(x) dx$ mit Dichte f_θ

für X diskret: $\mathbb{E}_\theta(X) = \sum_{i \in I} \xi_i \cdot \mathcal{P}_\theta(\{\xi_i\})$ mit Merkmalswerten $\xi_i, i \in I$

- **Varianz** $\mathbb{V}_\theta(X)$:

$$\mathbb{V}_\theta(X) = \mathbb{E}_\theta((X - \mathbb{E}_\theta(X))^2) = \mathbb{E}_\theta(X^2) - \mathbb{E}_\theta^2(X)$$

- **Median** t_m :

für X stetig: $t_m : F_X^\theta(t_m) = 0.5$

für X diskret: $\lim_{t \rightarrow t_m} F_X^\theta(t) \leq 0.5 \leq F_X^\theta(t_m)$

- **Minimum** m :

$$m = \min\{x | x \in X(\Omega)\}$$

Es werde vorausgesetzt, dass Erwartungswert und Varianz existieren.

Diese vier Kenngrößen geben dabei den „mittleren Wert“ der Zufallsvariablen, die mittlere quadratische Streuung der Zufallsvariablen um diesen mittleren Wert, den Wert der Zufallsvariablen, für den es gleichwahrscheinlich ist, einen größeren bzw. kleineren Wert der Zufallsvariable zu erhalten und den minimalen Wert an, den die Zufallsvariable annehmen kann.

Empirische Kenngrößen:

- **arithmetisches Mittel** \bar{X}_n :

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

- **empirische Varianz** $S_{(n)}^2$:

$$S_{(n)}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_{(n)})^2$$

- **Median** \tilde{t}_m :

$$\tilde{t}_m = \begin{cases} X_{\frac{n+1}{2}} & \text{für } n \text{ ungerade} \\ \frac{1}{2}(X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}) & \text{für } n \text{ gerade} \end{cases}$$

- **Minimum** \tilde{m} :

$$\tilde{m} = \min(X_1, \dots, X_n) = X_{(1)}$$

$(X_{(1)}, \dots, X_{(n)})$ bedeutet dabei, dass die Stichprobe nach ihren jeweiligen Werten geordnet ist. Es gilt also: $X_{(1)} \leq \dots \leq X_{(n)}$.

Die Idee der Ersetzungsmethode lässt sich mithilfe von Konvergenzsätzen der Wahrscheinlichkeitstheorie rechtfertigen. Nach dem *starken Gesetz der großen Zahl* gilt:

$$P\left(\lim_{n \rightarrow \infty} \bar{X}_{(n)} = \mathbb{E}(X)\right) = 1 \Leftrightarrow \bar{X}_{(n)} \rightarrow \mathbb{E}(X)$$

$$S_{(n)}^2 \rightarrow \mathbb{V}(X)$$

$$F_n(t; X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(-\infty, t]}(X_i) \rightarrow F_X(t) \Rightarrow \tilde{t}_m \rightarrow t_m, \tilde{m} \rightarrow m$$

Das starke Gesetz der großen Zahlen sagt damit aus, dass die Wahrscheinlichkeit, dass die empirischen Kenngrößen, $\bar{X}_{(n)}$ und $S_{(n)}^2$, gleich der entsprechenden theoretischen Kenngrößen sind, bei einer Stichprobe von $n \rightarrow \infty$ gleich 1 ist. Die Konvergenz von \tilde{t}_m und \tilde{m} gegen ihren entsprechenden theoretischen Größen, lässt sich aus der Konvergenz der empirischen Verteilungsfunktion gegen die Verteilungsfunktion ableiten.

Mit anderen Worten ist es so, dass für hohe Stichprobenzahlen n die wahren Werte der Kenngröße „gut“ approximiert werden. Aus der Approximation der empirischen Kenngrößen an die theoretischen Kenngrößen folgt auch, dass die Schätzwerte $\hat{\theta}$ den wahren Wert des Parameters θ approximieren.

Maximum-Likelihood-Methode

Ein anderes Vorgehen um Schätzfunktionen herzuleiten ist die Maximum-Likelihood-Methode. Grundlegende Annahme ist hier, dass eine Beobachtung „typisch“ für das zu bestimmende wahre

\mathcal{P}_θ ist, und diese deswegen mit hoher Wahrscheinlichkeit eintritt. Der Schätzwert für $\theta \in \Theta$ ist dann der Wert, der genau die zugehörige Wahrscheinlichkeit maximiert.

Formal geht man von einer konkreten Stichprobe (x_1, \dots, x_n) einer Zufallsvariable X mit wahrer Verteilung \mathcal{P}_θ und zugehöriger gegebener Wahrscheinlichkeitsdichte $f_\theta(x)$ aus. Die Wahrscheinlichkeit, die man dem Eintreten der Stichprobe (x_1, \dots, x_n) zuordnet, ergibt sich aus der dem Wahrscheinlichkeitsmaß des Produktmodells zugehörigen Dichtefunktion $f_\theta^{\otimes n}(x_1, \dots, x_n) = \prod_{i=1}^n f_\theta(x_i)$. Fasst man nun $f_\theta(x) = f(x, \theta)$ als Funktion von θ auf und maximiert die zugehörige Produktdichte mittels $\theta \in \Theta$, so ist $T(x_1, \dots, x_n)$ der sogenannte Maximum-Likelihood-Schätzer für θ , wenn er Maximalstelle der Dichte ist.

$$f^{\otimes n}(x_1, \dots, x_n, T(x)) = \max_{\theta \in \Theta} f^{\otimes n}(x_1, \dots, x_n, \theta), \quad \forall x \in \mathcal{X}$$

Die Funktion $f^{\otimes n}(x_1, \dots, x_n, \theta)$ wird Likelihood-Funktion, kurz L , genannt.

2.3 Kriterien geeigneter Schätzfunktionen

Es gibt drei wichtige Kriterien, um die Qualität von Schätzfunktionen zu bewerten und sie miteinander vergleichen zu können :

Konsistenz

Eine Schätzfolge $(T_n)_{n \geq 1}$ für $\tau(\theta)$ heißt konsistent, wenn

$$\mathcal{P}_\theta(|T_n - \tau(\theta)| \leq \epsilon) \xrightarrow[n \rightarrow \infty]{} 1, \quad \forall \epsilon > 0, \theta \in \Theta$$

Das bedeutet in Worten, dass mit größer werdener Stichprobenzahl n die Wahrscheinlichkeit für eine bestimmte Abweichung des Schätzers vom wahren Wert immer kleiner wird und der Schätzer somit den wahren Wert besser trifft.

Erwartungstreue

Ein Schätzer T_n gilt als erwartungstreu, wenn gilt:

$$\mathbb{E}_\theta(T_n) = \tau(\theta), \quad \forall n \in \mathbb{N}, \forall \theta \in \Theta$$

Unabhängig von der Stichprobengröße n trifft der Schätzer T_n im Mittel den wahren Wert $\tau(\theta)$. Anschaulich bedeutet das, simuliert man den Wert des Schätzers $T_n(x_1, \dots, x_n)$ unendlich oft, so ergibt sich als Mittelwert hier wieder der wahre Wert $\tau(\theta)$ unabhängig von der Wahl von n .

Gilt diese Erwartungstreue allerdings nur, wenn $n \rightarrow \infty$ strebt, also für eine hohe Stichprobenanzahl, spricht man von asymptotischer Erwartungstreue.

Zeigt eine Schätzfunktion T_n asymptotische Erwartungstreue oder ist T_n auch für hohe Stichprobenanzahlen nicht erwartungstreu, so ist

$$\mathbb{B}_{T_n}(\theta) = \mathbb{E}_\theta(T_n) - \tau(\theta).$$

der Bias bzw. systematische Fehler von T_n .

Ist eine Schätzfunktion mindestens asymptotisch erwartungstreu und konvergiert die Varianz der Schätzfunktion gegen 0, dann folgt daraus die Konsistenz der Schätzfunktion [Schmidt, 2004].

Effizienz

Um die Effizienz eines Schätzers T_n zu beurteilen, untersucht man dessen Varianz. Vergleicht man mehrere Schätzer miteinander, hier beispielsweise die erwartungstreuen Schätzer T_n und S_n , so ist der beste Schätzer unter ihnen varianzminimierend im Sinne von:

$$\mathbb{V}_\theta(T_n) \leq \mathbb{V}_\theta(S_n), \quad \forall \theta \in \Theta, n \in \mathbb{N}$$

Arbeitet man allerdings mit Schätzern, die nicht erwartungstreu sind, ist ein brauchbares Maß um zwei Schätzer miteinander zu vergleichen der mittlere quadratische Fehler $\mathbb{F}_{T_n}(\theta)$, der die Streuung um den wahren Wert beschreibt:

$$\mathbb{F}_{T_n}(\theta) := \mathbb{E}_\theta((T_n - \tau(\theta))^2) = \mathbb{B}_{T_n}(\theta)^2 + \mathbb{V}_\theta(T_n)$$

Für die Schätzfunktionen der Ersetzungsmethode, dem arithmetischen Mittel und der Stichprobenvarianz handelt es sich jeweils um erwartungstreue, konsistente, varianzminimierende Schätzer für den Erwartungswert und die Varianz [med, 2012].

Bei den Schätzern für den Median und Minimum handelt es sich um konsistente und varianzminimierende Schätzfunktionen. Die Schätzfunktionen, die sich dann für Parameter θ ergeben, gilt es dann im Detail auf diese Kriterien zu untersuchen.

Die Schätzfunktionen, die sich aus dem Maximum-Likelihood-Prinzip ergeben, sind im Allgemeinen nur konsistent und asymptotisch normalverteilt [Hujer, 2005].

2.4 Hypothesentests und Likelihood-Verhältnis-Methode

Bei den sogenannten Hypothesentests geht es im Gegensatz zur parametrischen Schätztheorie nicht darum, den zugrundeliegenden Zufallsmechanismus \mathcal{P}_θ möglichst treffend zu beschreiben, sondern ausgehend von einer Stichprobe und geeigneten statistischen Modell zu entscheiden, ob man eine vorher getroffene Annahme über \mathcal{P}_θ für zutreffend hält oder nicht. Man testet hier auf die sogenannte Alternativhypothese H_1 , wenn man tatsächlich ein statistisch signifikantes Ergebnis für H_1 erhält, darf man H_1 annehmen. Solange sich die beobachtete Stichprobe allerdings hinreichend mit der Nullhypothese H_0 beschreiben lässt, muss man diese beibehalten. H_0 und H_1 beziehen sich damit auf θ aus der Parametermenge Θ , die Hypothesenpaare zerlegen Θ damit in zwei komplementäre Bereiche:

$$H_0 : \theta \in \Theta_0; \quad H_1 : \theta \in \Theta_1, \quad \text{mit} \quad \Theta = \Theta_0 \cup \Theta_1$$

Nachdem die H_0 und H_1 gewählt wurden, gibt man sich ein sogenanntes Irrtumsniveau $\alpha \in (0, 1)$ vor, typischerweise ist $\alpha = 0.05$, womit man dann seine Entscheidungsregel mit einer geeigneten Statistik T_n konstruieren kann. Das Irrtumsniveau ist die Wahrscheinlichkeit für einen Fehler 1. Art. Diesen Fehler begeht man, wenn eine Beobachtung für H_1 spricht und man diese annimmt, obwohl in Wahrheit H_0 vorliegt. Mathematisch ist die Wahrscheinlichkeit für den Fehler 1. Art mit $\mathcal{P}_\theta(T_n \notin I)$ gegeben. Man fordert nun $\sup_{\theta \in \Theta_0} \mathcal{P}_\theta(T_n \notin I) \leq \alpha$, entsprechend dem, dass die Wahrscheinlichkeit für einen Fehler 1. Art kleiner gleich dem Wert des Irrtumsniveaus liegen soll. I ist dabei der Akzeptanzbereich für H_0 . Dieser lässt sich mit der Forderung $\sup_{\theta \in \Theta_0} \mathcal{P}_\theta(T_n \notin I) = \alpha$ aus dieser Gleichung bestimmen.

Ein Beispiel für einen bestimmten Hypothesentest ist der sogenannte *Likelihood-Ratio-Test* mit dem Hypothesenpaar:

$$H_0 : \theta = \theta_0; \quad H_1 : \theta \in \Theta_1 = \Theta \setminus \theta_0$$

Das gegebene Wahrscheinlichkeitsmaß aus dem zugehörigen statistischen Modell für das Einzel-experiment lautet:

$(\mathcal{X}, \mathcal{F}, \mathcal{P}_\gamma : \gamma = (\gamma_1, \dots, \gamma_m, \theta) \in \Gamma_1 \times \dots \times \Gamma_m \times \Theta)$, wobei die $\Gamma_i, i = 1, \dots, m, m \in \mathbb{N}$, die jeweilige Parametermenge für den Parameter γ_i ist. Diese Parameter und deren Parametermengen gelten gleichermaßen für beide Hypothesen.

Hier vergleicht man anschaulich die Wahrscheinlichkeit, die Stichprobe (x_1, \dots, x_n) unter H_0 zu beobachten, mit der entsprechenden Wahrscheinlichkeit unter H_1 mittels Quotientenbildung. Dazu setzt man voraus, es existiere eine Wahrscheinlichkeitsdichte $f(x, \theta)$. Im Produktmodell gilt: $f^{\otimes n}(x_1, \dots, x_n, \gamma_1, \dots, \gamma_m, \theta) = \prod_{i=1}^n f(x_i, \gamma_1, \dots, \gamma_m, \theta)$.

Man wählt folgende Teststatistik λ :

$$\lambda(x_1, \dots, x_n) = \frac{\sup_{\gamma_1 \in \Gamma_1, \dots, \gamma_m \in \Gamma_m} f(x_1, \dots, x_n, \gamma_1, \dots, \gamma_m, \theta_0)}{\sup_{\gamma_1 \in \Gamma_1, \dots, \gamma_m \in \Gamma_m, \theta \in \Theta_1} f(x_1, \dots, x_n, \gamma_1, \dots, \gamma_m, \theta)}$$

Da man im Nenner einen Parameter mehr hat, der sich frei variieren und die Wahrscheinlichkeit maximieren lässt, ist die Wahrscheinlichkeit im Nenner immer mindestens so groß wie im Zähler. Es gilt also für den Quotienten: $\lambda \leq 1, \forall (x_1, \dots, x_n) \in \mathcal{X}$. Unter der Nullhypothese erwartet man also Werte nahe dem Wert 1, so dass sich die Wahrscheinlichkeiten wenig unterscheiden.

Nach dem *Theorem von Wilk* gilt Folgendes [Cowan, 2013]: Für große n lässt sich die Verteilung von $\Lambda(X_1, \dots, X_n) = -2 \cdot \log(\lambda(X_1, \dots, X_n))$ mit der χ^2 -Verteilung unter der Nullhypothese approximieren:

$$P_{\theta_0}(\Lambda \leq t) \approx F_{\chi^2}(t), \quad \text{für } n \rightarrow \infty$$

Durch die Modifikation mit $-2 \cdot \log(\lambda)$ sind nun typische Werte von Λ unter H_0 eher kleine Werte, so dass sich wiederum mit der Vorgabe für den Fehler erster Art folgende Entscheidungsregel ergibt:

$$P_{\chi^2}(\Lambda(X_1, \dots, X_n) \notin I) = \alpha$$

Für I ergibt sich damit das Intervall $I = (0, c_{1-\alpha})$, wobei $c_{1-\alpha}$ das $(1 - \alpha)$ -Quantil ist. Im Falle von $\Lambda(X_1, \dots, X_n) \in (0, c_{1-\alpha})$ muss man die Nullhypothese beibehalten. Ein α -Quantil c_α einer beliebigen kontinuierlichen Verteilung mit Verteilungsfunktion F ist dabei wie folgt definiert: $F(c_\alpha) = \alpha$, mit $\alpha \in (0, 1)$.

3 Erweiterte Maximum-Likelihood-Methode: Ungebinnte Version

Wie in der Einleitung Kapitel 1 beschrieben, geht es in dieser Arbeit um die Untersuchung von Schätzmethoden, mit denen die Parameter α und E_c anhand einer Messreihe von Energien von Gammaphotonen bestimmt werden. Diese Parameter beschreiben dabei den Verlauf des Energiespektrums, $E^{-\alpha} \exp(-\frac{E}{E_c})$, der kosmischen Gammastrahlung. In diesem Kapitel wird nun untersucht, wie sich die Schätzfunktionen, die sich aus dem Prinzip der erweiterten Maximum-Likelihood-Methode ergeben, für diese beiden Parameter im Sinne der in Kapitel 2 vorgestellten Kriterien verhalten. Das allgemeine Prinzip der erweiterten Maximum-Likelihood Methode und der Unterschied zur Maximum-Likelihood-Methode aus Kapitel 2 werden im ersten Unterkapitel vorgestellt. Daran schließt sich die Bewertung dieser Methode, angewandt auf die kosmische Gammastrahlung und dessen Energiespektrum an. Dazu werden mit verschiedenen Simulationen, zu vorgegebenen Paramterwerten von α und E_c , die Verteilungen der zugehörigen Schätzfunktionen simuliert. Innerhalb dieses Kapitels werden die Messwerte der Energien als ungebinnt betrachtet.

3.1 Prinzip der erweiterten Maximum-Likelihood-Methode bei ungebinnten Daten

Dieses Unterkapitel orientiert sich an [Kolanoski \(2008\)](#) und [Hansmann-Menzemer \(2007\)](#). Das Maximum-Likelihood-Prinzip wurde bereits im Kapitel 2.2 erläutert. Die allgemeine und grundlegende Idee der Methode sei hier noch einmal erwähnt.

Man maximiert anhand der Parameter der Wahrscheinlichkeitsverteilung die Wahrscheinlichkeit dafür, dass eine bestimmte Realisation einer Stichprobe (x_1, \dots, x_n) , die der Wahrscheinlichkeitsverteilung entnommen ist, eintritt. Die Stichprobe wird daher als typische Realisation gemäß der wahren Verteilung angesehen. Sei nun beispielhaft eine beliebige reelle Wahrscheinlichkeitsdichte $f(x)$ gegeben, die von einem Satz von m unbekannten kontinuierlichen Parametern $\theta = (\theta_1, \dots, \theta_m)$, $m \in \mathbb{N}$, abhängt. Θ_i , mit $\theta_i \in \Theta_i$, $\forall i \in \{1, \dots, m\}$, sei jeweils ein geeigneter Definitionsbereich. So bedeutet das Maximum-Likelihood-Prinzip, dass die Likelihood-Funktion L , also die zum Produktmodell zugehörige Dichte in Abhängigkeit aller θ_i 's, unter Kenntnis der Stichprobe gleichzeitig maximiert wird:

$$L(\theta_1, \dots, \theta_m) := \prod_{i=1}^n f(\theta_1, \dots, \theta_m; x_i)$$

Die Maximum-Likelihood-Schätzer $(\hat{\theta}_1, \dots, \hat{\theta}_m)$ erfüllen dann:

$$\sum_{i=1}^n \left. \frac{\partial \ln L}{\partial \theta_j} \right|_{\hat{\theta}_j} = 0 \quad \wedge \quad \left. \frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j} \right|_{\hat{\theta}_i, \hat{\theta}_j} = U_{ij} \quad \text{negativ definit} \quad , \forall i, j \in \{1, \dots, m\}$$

$$\text{bzw. } L(\hat{\theta}_1, \dots, \hat{\theta}_m) \geq L(\theta_1, \dots, \theta_m) \quad \forall \theta_i \in \Theta_i$$

Häufig benutzt man den natürlichen Logarithmus „ \ln “ der Likelihood-Funktion, da sich dieser bei numerischen Maximierungen deutlich stabiler verhält und auch bei analytischen Betrachtungen das Differenzieren vereinfacht, jedoch den Wert der Extremstelle nicht ändert, da der Logarithmus eine streng monoton steigende Funktion ist.

Die sich ergebende Matrix U ist negativ definit, wenn alle Eigenwerte kleiner 0 sind. Dies ist bei nur einem Parameter gleichbedeutend mit der Bedingung, dass die 2. Ableitung kleiner 0 sein muss, damit ein echtes Maximum vorliegt.

Die zur Dichte zugehörige Normierungskonstante garantiert dabei unabhängig von den Schätzwerten der Parameter, dass die Dichte f und damit auch die Likelihood-Funktion stets normiert sind.

$$\int f(\theta_1, \dots, \theta_m; x) dx = 1 \quad , \forall \theta_i \in \Theta_i, i \in \{1, \dots, m\}$$

Um nun die Idee der *erweiterten Maximum-Likelihood-Prinzip* herzuleiten, wird diese Zwangsbedingung der Normierung aufgehoben. Abhängig vom Design des Experiments ist es vor allem in der Teilchenphysik bzw. Astrophysik so, dass die Stichprobenzahl n , also die Gesamtanzahl gemessener Ereignisse, kein beliebig wählbarer Parameter ist, sondern bei einer festen Messzeit selbst dem Zufall unterliegt. Die Stichprobenzahl n lässt sich also als Zufallsvariable beschreiben. Mit einer vorgegebenen Messzeit t , die sich als wählbarer Parameter des Experiments festlegen lässt, ist die Gesamtzahl der registrierten Ereignisse im gesamten detektierten Energiebereich poisson-verteilt. Die gemessene Zahl der Ereignisse, also n , ist somit der Schätzwert für den zugehörigen Poisson-Parameter \mathcal{N} , der die unbekannte zu erwartende Gesamtzahl der Ereignisse beschreibt. Ziel ist es nun, die Anzahl der tatsächlich beobachteten Ereignisse n mit den theoretisch zu erwarteten Ereignissen \mathcal{N} in Übereinstimmung zu bringen. \mathcal{N} hängt dabei von den Parametern $(\theta_1, \dots, \theta_m)$ ab.

Ausgehend von der gewöhnlichen Likelihood-Funktion von oben, und der Berücksichtigung dieser Poisson-Wahrscheinlichkeit, n Ereignisse bei \mathcal{N} erwarteten Ereignissen zu beobachten, ist die modifizierte Likelihood-Funktion, also die Likelihood-Funktion der *erweiterten Maximum-Likelihood-Methode* gegeben durch:

$$L(\theta_1, \dots, \theta_m) = \frac{\mathcal{N}^n e^{-\mathcal{N}}}{n!} \prod_{i=1}^n f(\theta_1, \dots, \theta_m; x_i)$$

Die zugehörigen Maximum-Likelihood-Schätzer $(\hat{\theta}_1, \dots, \hat{\theta}_m)$ sind die Maximalstelle der Produktdichte und der Poisson-Wahrscheinlichkeit.

Wichtig ist es zu erwähnen, dass die Stichprobengröße n innerhalb dieser Arbeit als Parameter zur Simulation vorgegeben wird. Dies ermöglicht eine bessere Vergleichbarkeit der Ergebnisse der Schätzfunktionen, da man sie so anhand der gleichen Stichprobengröße vergleichen kann. Dies gilt aber nur für die Zwecke der Analysen in dieser Arbeit. In der tatsächlichen Anwendung im Experiment ist dies nicht möglich. Die eigentliche Zufälligkeit der Stichprobengröße wird allerdings gerade durch die Idee der erweiterten Maximum-Likelihood-Methode, also die Anpassung an die gemessenen n Ereignisse mittels der Poisson-Wahrscheinlichkeit, berücksichtigt. Nun lässt sich mithilfe der Umformung auf eine Funktion $g(x)$ zeigen, dass die erweiterte Maximum-Likelihood-Methode, die Parameter $(\theta_1, \dots, \theta_m)$ zusätzlich so anpasst, dass $\mathcal{N} = n$ ist. Die zugehörige Log-Likelihood-Funktion ist:

$$\ln L(\theta_1, \dots, \theta_m) = n \ln \mathcal{N} - \mathcal{N} + \sum_{i=1}^n \ln f(\theta_1, \dots, \theta_m; x_i) + const$$

wobei der Term $n!$, der nicht von den θ_i abhängt mit $const$ abgekürzt wird. Diesen kann man wegen des noch zu vollziehenden Maximierungsprozesses vernachlässigen.

Mit der Umrechnung

$$n \ln \mathcal{N} + \sum_{i=1}^n \ln f(\theta_1, \dots, \theta_m; x_i) = \sum_{i=1}^n \ln (\mathcal{N} \cdot f(\theta_1, \dots, \theta_m; x_i))$$

lässt sich die Funktion $g(x) := \mathcal{N} \cdot f(x)$ definieren, die aufgrund der Normierung der Dichte f auf den Wert \mathcal{N} normiert ist: $\int g(x) dx = \mathcal{N}$.

Damit erhält die Likelihood-Funktion folgende Form:

$$\ln L(\theta_1, \dots, \theta_m) = \sum_{i=1}^n \ln g(\theta_1, \dots, \theta_m; x_i) - \int g(\theta_1, \dots, \theta_m; x) dx$$

Nimmt $\ln L$ im ersten Term durch einen größeren Wert der Normierung höhere Werte an, so nimmt der Wert durch den zweiten Term hingegen ab. Es gilt hier also so eine angemessene Balance zu finden, dass sich insgesamt ein Maximum von $\ln L$ einstellt.

Die Schätzwerte für die Parameter $(\hat{\theta}_1, \dots, \hat{\theta}_m)$ erfüllen die Bedingung:

$$\sum_{i=1}^n \left(\frac{\partial \ln g(\hat{\theta}_1, \dots, \hat{\theta}_m; x_i)}{\partial \theta_j} \right) - \frac{\partial \int g(\hat{\theta}_1, \dots, \hat{\theta}_m; x) dx}{\partial \theta_j} = 0 \quad , \forall j \in \{1, \dots, m\}$$

Dabei gilt es zu verifizieren, ob tatsächlich ein Maximum vorliegt. Skaliert man die Funktion g mit einer reellen Konstante $\beta > 0$, ist die Likelihood-Funktion dann maximal, wenn folgender Zusammenhang gilt:

$$\begin{aligned} \ln L(\theta_1, \dots, \theta_m) &= \sum_{i=1}^n \ln \beta g(\theta_1, \dots, \theta_m; x_i) - \int \beta g(\theta_1, \dots, \theta_m; x) dx \\ \frac{\partial \ln L}{\partial \beta} &= \frac{n}{\beta} - \int g(x) dx \stackrel{!}{=} 0 \quad \Rightarrow \beta_{max} = \frac{n}{\int g(x) dx} \end{aligned}$$

Da das tatsächliche β in der Likelihood-Funktion gleich 1 ist, folgt daraus, dass sie dann maximal wird, wenn die integrierte Funktion g in Abhängigkeit ihrer Parameter den Wert n annimmt. Kurzum besteht die erweiterte Maximum-Likelihood-Methode darin, die Parameter so zu variieren, dass die Funktion g mit der Nebenbedingung, dass das Integral über die Funktion gleich n ist, maximiert wird. Damit beschreibt die Dichte $g(x)$ nicht nur die Form der Verteilung, sondern auch ihre Größe bzw. ihren Inhalt [Barlow, 1990]. Die Information der gemessenen Anzahl von Ereignissen verbessert damit die Schätzung der Parameter, da diese daran angepasst werden. Hier wird demnach die Funktion $g(x)$ innerhalb der erweiterten Maximum-Likelihood-Methode angepasst. Die Parameter stellen sich innerhalb des Fits so ein, dass wiederum das Integral über diese Funktion gleich der Zahl der beobachteten Ereignisse ist:

$$\int g(\hat{\theta}_1, \dots, \hat{\theta}_m; x) dx = n$$

3.2 Untersuchung der Schätzfunktionen

In dem Kontext der Messung der Energien der kosmischen Gammastrahlung mit einem zugehörigem Energiespektrum, das sich mit einem Potenzgesetz mit exponentiellem Cutoff beschreiben lässt, gilt hier für die Funktionen f und g :

Im Zuge der mathematischen Betrachtungen verzichten wir hier auf die Beschreibung der Energie mit der Variablen E und nehmen stattdessen X als Zufallsvariable bzw. x als konkreter Messwert der Energie. Des Weiteren wird der Ausdruck E_c mit $\frac{1}{\lambda}$ ersetzt.

$$f(x) = C(\alpha, \lambda) x^{-\alpha} e^{-\lambda \cdot x} \quad \text{mit} \quad \int f(x) dx = 1 \quad , x \in [m, m'], \alpha \in (0, \infty), \lambda \in (0, \infty)$$

$$g(x) = \Phi_0 x^{-\alpha} e^{-\lambda \cdot x} \quad \text{mit} \quad \int g(x) dx = \mathcal{N}, \quad x \in [m, m'], \alpha \in (0, \infty), \lambda \in (0, \infty)$$

$[m, m']$ ist der Energiebereich in dem man die Gammastrahlung beobachtet. Da sich diese Arbeit am H.E.S.S.-Experiment orientiert, wird hier ein Bereich gewählt, in dem dieses Experiment auch misst [HES,]. Die untere Grenze m wird auf 100 GeV gesetzt, die obere auf 10 TeV.

Um nun im Folgenden zu untersuchen wie sich die Schätzfunktionen für die Parameter α , λ und Φ_0 verhalten, die sich aus der erweiterten Maximum-Likelihood-Methode ergeben, werden für verschiedene Stichprobengrößen n die Verteilungen der Schätzfunktionen $\hat{\Phi}_0$, $\hat{\alpha}$ und $\hat{\lambda}$ simuliert. Unter Simulation ist zu verstehen, dass für eine gewisse Zahl, im Folgenden mit N bezeichnet, N verschiedene Stichproben (x_1, \dots, x_n) für eine bestimmte Stichprobengröße n gezogen werden. Grund für die fest vorgegebene Stichprobengröße ist wiederum so die Schätzfunktionen besser vergleichen zu können, siehe vorheriges Unterkapitel. Stichprobengrößen, die man mit dem H.E.S.S.-Experiment misst, liegen nach Aharonian et al. (2009), im Bereich von 500 bis 5000 gemessenen Ereignissen. Die Untersuchung der Schätzfunktionen in diesem Bereich ist also von besonderer Bedeutung.

Die wahren Parameter, die zur Simulation benutzt werden, liegen dabei im Bereich gemessener Werte von Gammastrahlungsquellen, die einen Cutoff zeigen. Sie haben den Wert $\alpha_0 = 2.0$ und $\lambda_0 = 0.1$ [Aharonian et al., 2009]. Die Null im Index steht hier und im Folgenden für die verwendeten Parameter der jeweiligen Simulationen.

n	$\bar{\hat{\Phi}}_0$	$\hat{\sigma}_{\hat{\Phi}_0}$	$\bar{\hat{\alpha}}$	$\hat{\sigma}_{\hat{\alpha}}$	$\bar{\hat{\lambda}}$	$\hat{\sigma}_{\hat{\lambda}}$
300	44.8	239.6	1.98439	0.09503	0.140725	0.219383
1 000	107.0	10.0	1.99461	0.05365	0.111627	0.052620
3 000	318.6	17.1	1.99627	0.03127	0.105384	0.030106
10 000	1055.3	31.6	1.99915	0.01739	0.101442	0.016565
30 000	3161.9	52.7	1.99956	0.00984	0.100449	0.008970
100 000	10534.8	95.91	1.99975	0.00540	0.100172	0.004782

Tabelle 1: Arithmetische Mittelwerte und Standardabweichungen der simulierten Verteilungen der drei Schätzfunktionen $\hat{\Phi}_0$, $\hat{\alpha}$ und $\hat{\lambda}$ für verschiedene Stichprobengrößen n und einer Simulationsgröße $N=1000$ bei den vorgegebenen Parametern α_0 und $\lambda_0 = 0.1$.

Die Werte für Φ_0 ergeben sich dann aus den zwei Parametern und der Stichprobengröße n . Denn mit der Normierungsbedingung $\int \Phi_0 x^{-\alpha} e^{-\lambda x} dx = n$ ist der wahre Wert von Φ_0 eindeutig bestimmt. Der Index Null hat hier nicht die obige Bedeutung des Parameterwerts für die Simulation, der Index gehört zur typischen Bezeichnung der Flussnormierung. Mit jeder dieser Stichproben lässt sich dann jeweils ein Schätzwert für die drei Parameter angeben.

Die Tabelle 1 zeigt für $N=1000$ die zugehörigen arithmetischen Mittelwerte und Standardabweichungen der simulierten Verteilungen der Schätzer.

In Abbildung 1 sind für Stichprobengrößen von $n = 300$ und $n = 10000$ die simulierten Verteilungen dargestellt. Die Schätzfunktion $\hat{\Phi}_0$ wird allerdings nicht auf ihre Qualität als Schätzer beurteilt. Grund dafür ist, dass die Parameter α und λ , die die Form der Wahrscheinlichkeitsdichte beschreiben, Informationen über zugrundeliegende physikalische Prozesse liefern, Φ_0 hingegen skaliert die Funktion g nur.

Aus Tabelle 1 und den simulierten Verteilungen der drei Schätzfunktionen $\hat{\Phi}_0$, $\hat{\alpha}$ und $\hat{\lambda}$ aus Abbildung 3 lässt sich erkennen, dass die Schätzwerte für die beiden Parameter α und λ mit

wachsender Stichprobengröße jeweils gegen den wahren Wert α_0 und λ_0 streben und dabei weniger Streuung um ihren Mittelwert aufweisen. Abbildung 4 zeigt innerhalb der Auftragung der Standardabweichungen gegenüber der Stichprobengröße außerdem, dass sich dieses Verhalten in der doppelt-logarithmischen Auftragung linear anpassen lässt. Es liegt somit eine Abhängigkeit von $\hat{\sigma}_{\hat{\alpha}}$ bzw. $\hat{\sigma}_{\hat{\lambda}}$ von n vor, die hier einer Gerade mit Steigung -0.5 entspricht, siehe Abbildung 4, somit gilt für beide Schätzfunktionen: $\hat{\sigma} \propto \frac{1}{\sqrt{n}}$.

Tabelle 1 und Abbildung 4 zeigen auch, dass sich $\hat{\lambda}$ von oben an λ_0 annähert, wohingegen sich die Werte von $\hat{\alpha}$ von unten an α_0 annähern. Mit anderen Worten lässt sich feststellen: $\hat{\lambda}_n > \hat{\lambda}_{n'} > \lambda_0$ und $\hat{\alpha}_n < \hat{\alpha}_{n'} < \alpha_0$, für verschiedene Stichprobengrößen $n' > n$.

Dieses Ergebnis spricht für eine Antikorrelation der beiden Schätzfunktionen. Mit Antikorrelation bezeichnet man die Tatsache, dass ein statistisch linearer Zusammenhang zwischen $\hat{\lambda}$ und $\hat{\alpha}$ vorliegt, bei dem kleinere (größere) Werte hinsichtlich des Mittelwerts der einen Schätzfunktion, kleinere (größere) Werte hinsichtlich des Mittelwerts der anderen Schätzfunktion mit sich ziehen. Dies entspricht wiederum der Idee der erweiterten Maximum-Likelihood-Methode. Damit das Spektrum $\Phi_0 x^\alpha \exp(-\lambda x)$ aufintegriert n ergibt, muss bei einem tendenziell kleineren (größeren) Wert des einen Parameters, der jeweilige andere Wert tendenziell größer (kleiner) sein. Dies bestätigt auch der Korrelations-Plot, siehe Abbildung 2. Ein Paar $(\hat{\alpha}, \hat{\lambda})$ ergibt sich jeweils aus einer Stichprobe (x_1, \dots, x_n) . Der Korrelationskoeffizient c , der den linearen statistischen Zusammenhang misst, liegt hier bei $c_{\hat{\alpha}, \hat{\lambda}} = -0.76$. Dies ist eine Bestätigung der Antikorrelation der Schätzfunktionen.

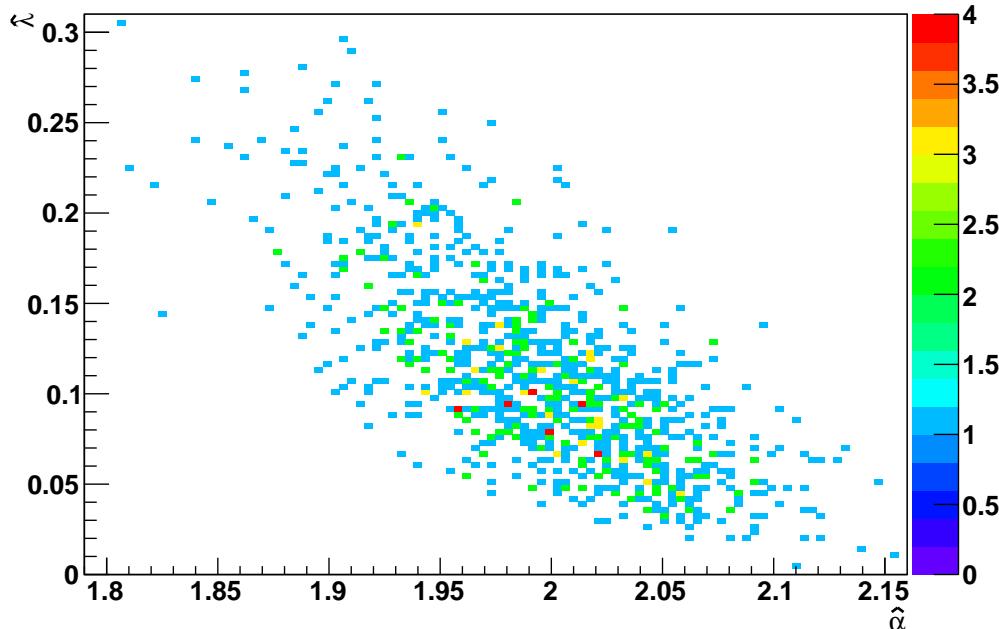


Abbildung 2: Auftragung der jeweiligen $(\hat{\alpha}, \hat{\lambda})$ aus den 1000 Simulationen. Jedes Paar ist dabei Schätzwert für die beiden Parameter für eine Stichprobe (x_1, \dots, x_n) am Beispiel von $n=1000$. Ausgehend vom Punkt der wahren Parameter $(\alpha_0, \lambda_0) = (2.0, 0.1)$, streuen die Punkte der geschätzten Parameter um eine gedachte Gerade mit Steigung von circa -1. $\hat{\alpha}$ und $\hat{\lambda}$. Sie sind also anti-korrelierte Größen. Die verwendete Farbskala beschreibt die Häufigkeiten der Wertepaare.

Ferner bedeuten obige Ergebnisse, dass das Schätzverfahren keine erwartungstreuen Schätzfunktionen für die Parameter α und λ liefert und ein stichprobenabhängiger Bias besteht. Dies wird durch die Plots aus Abbildung 4 veranschaulicht. Da die Schätzfunktionen für große Stichpro-

benzahlen n gegen den wahren Wert α_0 bzw. λ_0 streben, sie also asymptotisch erwartungstreu sind und ihre Standardabweichung abnimmt, folgt mit Kapitel 2.3, dass es sich um konsistente Schätzfunktionen handelt.

Ferner erkennt man aus den simulierten Verteilungen von $\hat{\lambda}$ und $\hat{\alpha}$ in Abbildung 3 anschaulich, dass die Standardabweichung bei der größeren Stichprobe von $n = 10000$ deutlich geringer ist verglichen mit denen bei $n = 300$. Die Verteilung von $\hat{\lambda}$ nimmt beim Vergleich der beiden Stichprobengrößen eine Form mit höherer Symmetrie an. Hier zeigt sich die Eigenschaft von Maximum-Likelihood-Schätzern asymptotisch normalverteilt zu sein. Bei $n = 300$ liegt hier noch eine sehr asymmetrische, rechtsschiefe Verteilung vor. Vorteil einer symmetrischen Verteilung ist einerseits die Möglichkeit der Approximation mit einer Normalverteilung, andererseits ergibt sich hier auch eine einfache Möglichkeit Konfidenzintervalle für die Parameter anzugeben. Diese ergeben sich aus den geschätzten Standardabweichungen. So lässt sich mit diesen das 68%-Intervall beispielsweise für $\hat{\alpha}$ mit $[\hat{\alpha} + \hat{\sigma}_{\hat{\alpha}}, \hat{\alpha} - \hat{\sigma}_{\hat{\alpha}}]$ angeben.

Abschließend soll noch der große Wert von $\hat{\sigma}_{\hat{\Phi}_0}$ für $n=300$ kommentiert werden. Beurteilt man diese Streuung des Parameters unter Berücksichtigung des Wertes $\hat{\sigma}_{\hat{\lambda}_0}$, so zeigt sich dass $\hat{\sigma}_{\hat{\lambda}_0}$ hier im Vergleich mit den Werten bei größeren Stichproben n verhältnismäßig auch stärker streut. Bedenkt man, dass λ Parameter einer Exponentialfunktion ist, trägt hier die Streuung von $\hat{\lambda}$ zu einer starken Streuung von „ $\exp(-\lambda x)$ “ bei. Zusammen mit dem eher unauffälligen Schwanken von α , lässt sich so der große Wert von $\hat{\sigma}_{\hat{\Phi}_0}$ erklären.

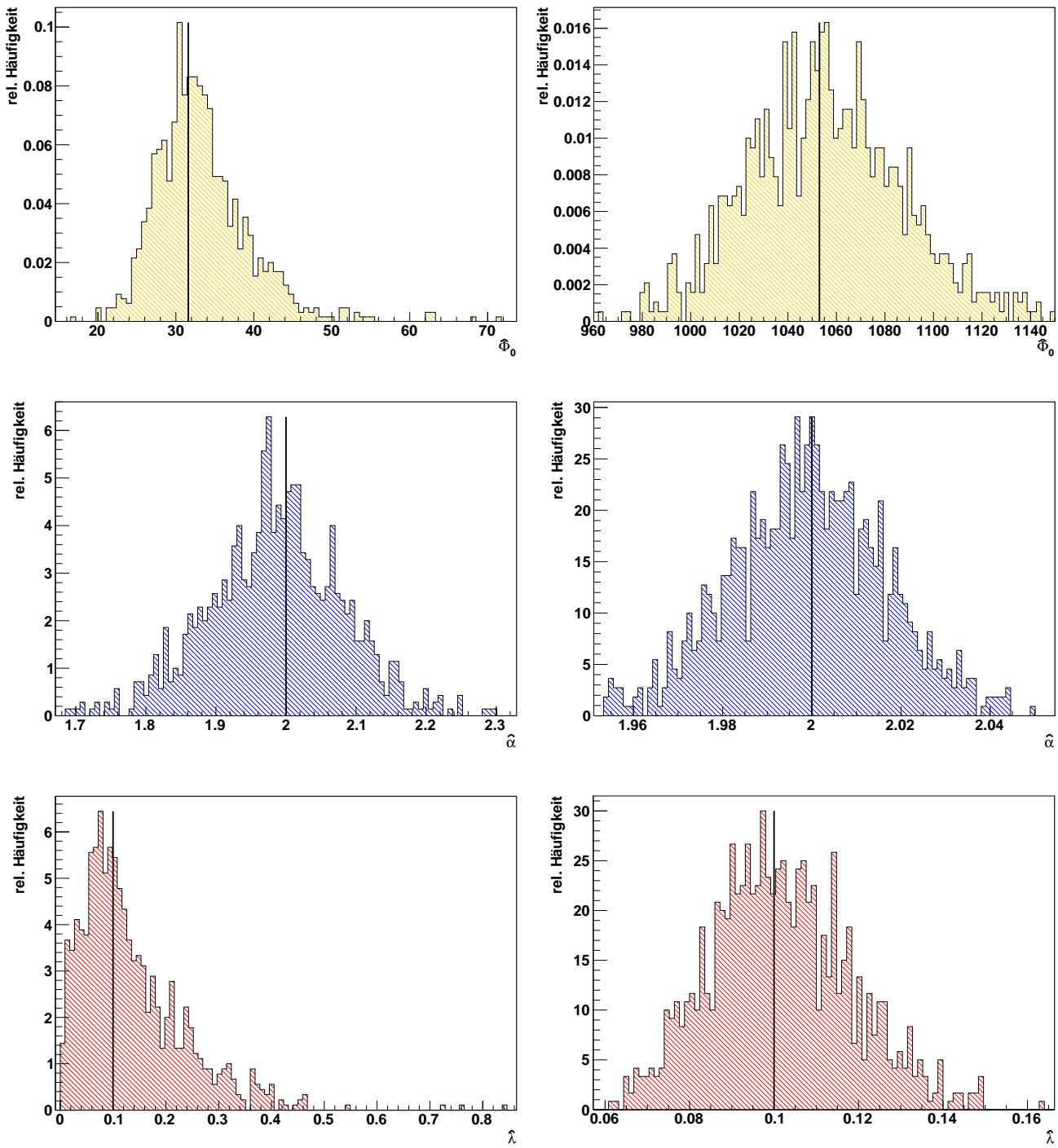


Abbildung 3: Simulierte relative Häufigkeiten der Schätzfunktion aller drei Parameter $\hat{\Phi}_0$, $\hat{\alpha}$ und $\hat{\lambda}$ für $N=1000$. Links jeweils für eine Stichprobengröße von $n=300$, rechts jeweils $n=10000$. Die oberste Reihe zeigt die simulierten geschätzten Werte für $\hat{\Phi}_0$, darunter $\hat{\alpha}$ und schließlich $\hat{\lambda}$. Bereits für eine kleine Stichprobengröße von $n=300$ zeigen die simulierten Verteilungen von $\hat{\Phi}_0$ und $\hat{\alpha}$ eine recht symmetrische Form. Die simulierte Verteilung von $\hat{\lambda}$ zeigt bei dieser Stichproben eine stark asymmetrische, rechtsschiefe Verteilung. Ab einer Stichprobengröße von $n=10000$ lässt sich eine Approximation der simulierten Verteilung mit einer Normalverteilung durchführen. Grund hierfür ist die Symmetrie der simulierten Verteilung. Hier zeigt sich die Eigenschaft der Maximum-Likelihood-Schätzer asymptotisch normalverteilt zu sein. Wie gut die Approximation der Normalverteilung der simulierten Verteilung tatsächlich passt, wurde nicht untersucht. Etwaige Ausreißer wurden in der Auftragung nicht berücksichtigt. Die eingezeichneten Linien entsprechen jeweils dem wahren Parameterwert.

Die Schätzwerte für die Parameter aus Tabelle 1 lassen sich auch gegen die zugehörige Stichprobengröße auftragen, um so das Verhalten der Schätzfunktionen in Abhängigkeit der Stichprobengröße zu veranschaulichen.

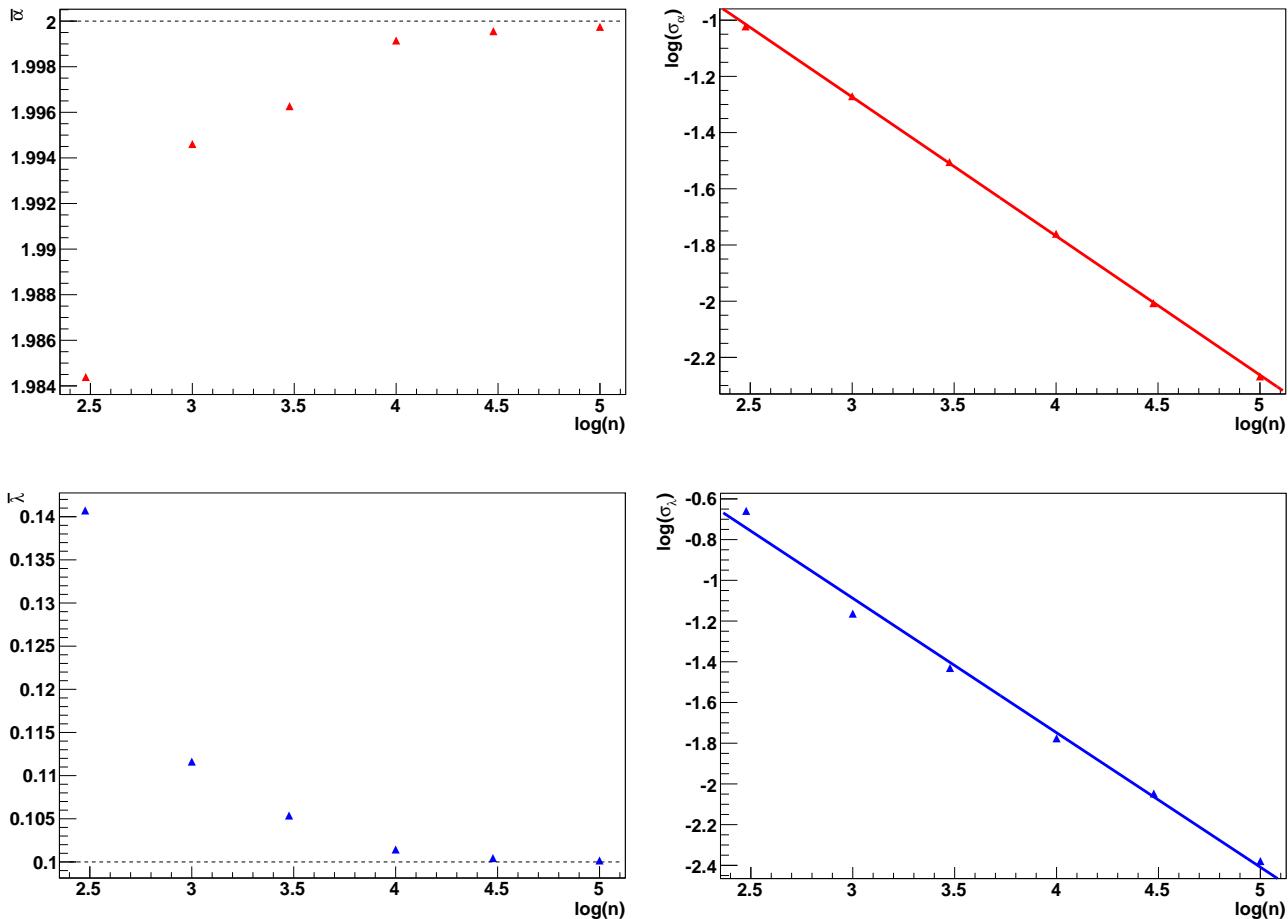


Abbildung 4: Die Plots zeigen die Schätzwerte aus Tabelle 1, aufgetragen gegenüber der Stichprobengröße n , um das Verhalten der Schätzfunktionen in Abhängigkeit von n zu veranschaulichen. Die obere Zeile, in rot dargestellt, zeigt die Werte für $\hat{\alpha}$, links und rechts die Werte von $\sigma_{\hat{\alpha}}$. Die untere Zeile in rot dargestellt zeigt die Werte für $\hat{\lambda}$, links und rechts die Werte von $\sigma_{\hat{\lambda}}$. Die gestrichelten Linien in den Plots bei den arithmetischen Mitteln zeigen die wahren Werte der Parameter $\alpha_0 = 2.0$ und $\lambda_0 = 0.1$. Bei der Auftragung der geschätzten Standardabweichungen von $\hat{\alpha}$ und $\hat{\lambda}$ eignet sich die doppelt-logarithmische Auftragung. Hier zeigt sich, dass die Werte der Standardabweichungen in Abhängigkeit der Stichprobengröße linear mit einer Steigung von ca. 0.5 abfallen. Die Standardabweichung ist für beide Schätzfunktionen damit proportional zu $\frac{1}{\sqrt{n}}$. Bei den arithmetischen Mittelwerten von $\hat{\alpha}$ und $\hat{\lambda}$ zeichnet sich bei der doppelt-logarithmischen Auftragung kein linearer Zusammenhang ab. Die arithmetischen Mittel zeigen für beide Schätzgrößen die Konvergenz gegen die wahren Werte. Die Standardabweichungen zeigen beide die Konvergenz gegen 0.

4 Erweiterte Maximum-Likelihood-Methode: Gebinnte Version

Im vorherigen Kapitel wurde die erweiterte Maximum-Likelihood-Methode auf ungebinnedne Messwerte der Energie der kosmischen Gammastrahlung angewandt. In diesem Kapitel wird die Methode und die sich daraus ergebenden Schätzfunktionen für α und λ auf gebinnte Daten untersucht. Grundsätzliche Idee ist es hier, zu zählen, wie viele Ereignisse in den einzelnen Bins registriert werden. Dies ist das übliche statistische Instrument, mit der in der Gammaastronomie Schätzfunktionen konstruiert werden. Dazu wird wiederum zunächst allgemein der mathematische Formalismus vorgestellt, der sich für gebinnte Daten ergibt. Danach werden wiederum mit verschiedenen Simulationen zu vorgegebenen Paramterwerten von α und λ die Verteilungen der zugehörigen Schätzfunktionen $\hat{\alpha}$ und $\hat{\lambda}$ simuliert, anhand deren Ergebnisse sich die Schätzfunktionen der erweiterte Maximum-Likelihood-Methode in Anwendung auf gebinnte Daten bewerten lassen und mit den Schätzfunktionen bei ungebinnnten Daten vergleichen.

4.1 Prinzip der erweiterten Maximum-Likelihood-Methode bei gebinnten Daten

Angenommen, eine Zufallsvariable X nehme Werte in einer Menge \mathcal{X} an, z.B. $[m, m']$ mit $m, m' \in \mathbb{R}$ und $m' > m$, und es sei eine Stichprobe (X_1, \dots, X_n) vom Umfang n gegeben. Diese Stichprobe liegt gebinnt vor, wenn dieser Wertebereich \mathcal{X} in verschiedene Teilmengen, sogenannte *Bins*, unterteilt wird und letztlich nur die Information vorliegt, in welchen dieser Teilmengen die X_i liegen bzw. in welchem Bin ihr Wert eingetragen ist. Ein Beispiel hierfür ist ein Histogramm. Gebinnte Daten gehen damit automatisch mit einem Informationsverlust einher.

Es lässt sich folgendermaßen eine Zufallsvariable N_i definieren, die das Binning der Zufallsvariablen X beschreibt:

$$N_i = \sum_{j=1}^n \mathbf{1}_{A_i}(X_j), \quad i \in \{1, \dots, K\}$$

A_i ist dabei eine der Teilmengen von \mathcal{X} , die Vereinigung dieser disjunktten Teilmengen ergeben wiederum \mathcal{X} . Die verwendete Funktion $\mathbf{1}_A(X)$ ist wie folgt definiert:

$$\mathbf{1}_A(X) = \begin{cases} 1, & \text{wenn } X \in A \\ 0, & \text{sonst} \end{cases}$$

Da die Gesamtzahl der Ereignisse N poissonverteilt ist, so sind N_i damit eine poisson-verteilte Zufallsvariable mit Werten in der natürlichen Zahlenmenge und zählt wieviele der $X_i, i = 1, \dots, n$, Werte im i-ten von insgesamt K Bins annehmen. Der zugehörige Poisson-Parameter, also die erwartete Anzahl von Einträgen im i-ten Bin, ist l_i . Aus einer Stichprobe von X mit Umfang n , also (X_1, \dots, X_n) , erhält man damit eine Stichprobe der Zufallsvariablen N , (N_1, \dots, N_K) , der Länge K .

Allgemein lassen sich die Teilmengen A_i eines Intervall $\mathcal{X} = [m, m']$ folgendermaßen definieren: $A_i = [m + (i - 1) \cdot b_i, m + i \cdot b_i]$, b_i ist dabei die Länge des i-ten Bins, $i = 1, \dots, K$. Häufig wird \mathcal{X} äquidistant aufgeteilt, so gilt $L_i = L, \forall i, l > 0$. Im Kontext von potenzförmigen Zusammenhängen sind vor allem exponentiell skalierte Intervalllängen nützlich. Hier betrachtet man das Intervall $\mathcal{X}' = [\log(m), \log(m')]$ und unterteilt dieses wiederum in äquidistante Intervalllängen. Vorteil dieser Wahl ist, dass bei hohen Werten der Zufallsvariablen X , die mit geringer Wahr-

scheinlichkeit eintreten, durch die exponentiell anwachsenden Binlängen, es nun auch recht wahrscheinlich ist, dass hier Bins mit Einträgen vorliegen.

Um nun die Poisson-Parameter l_i für das Energiespektrum der kosmischen Gammastrahlung, betrachtet im Energiebereich $[m, m']$, für ein bestimmtes Bin zu berechnen, integriert man die Funktion des Spektrums $g(x) = \Phi_0 x^\alpha e^{-\lambda x}$ im Bereich A_i auf. Die Poisson-Parameter $l_i = l_i(\Phi_0, \alpha, \lambda)$ hängen damit wiederum von den Parameter des Energiespektrums $g(x)$ ab:

$$l_i(\Phi_0, \alpha, \lambda) = \int_{A_i} g(x) dx = \int_{m+(i-1)\cdot b_i}^{m+i\cdot b_i} \Phi_0 x^\alpha e^{-\lambda x} dx, \quad x \in [m, m']$$

$$\text{wobei allgemein gilt: } \int_m^{m'} \Phi_0 x^\alpha e^{-\lambda x} dx = \mathcal{N}, \quad \mathcal{N} \in \mathbb{N}$$

Die zugehörigen Poisson-Parameter l_i , also die erwartete Zahl von Ereignissen in einem Bin i, ergibt sich also aus der Integration über das Energiespektrum über den Bereich der Energien, welche das Bin abdeckt. Die Wahrscheinlichkeit, dass (N_1, \dots, N_K) die Werte (n_1, \dots, n_K) annimmt, falls die Stichprobe von X eine Größe von n hat, $n = \sum_{j=1}^K n_j$, ist folgendermaßen gegeben:

$$\mathcal{P}\left(N_1 = n_1, \dots, N_K = n_K \mid n = \sum_{i=1}^K n_i\right) = \frac{n!}{\prod_{j=1}^K n_j!} \prod_{j=1}^K p_j^{n_j}, \quad \text{wobei } p_j = l_j/\mathcal{N}$$

p_j entspricht dabei dem Quotient aus erwarteter Zahl in einem Bin und erwarteter Gesamtzahl der Ereignisse. Die Idee ist also ausgehend von n Ereignissen, diese auf K verschiedene Bins zu verteilen. Dies entspricht dem Kontext einer Multinomialverteilung.

Berücksichtigt man, wie bei den ungebundenen Messwerten, dass man auch an die Gesamtzahl der Ereignisse anpasst, ergänzt man die Gleichung um die Poisson-Wahrscheinlichkeit insgesamt n Ereignisse zu beobachten.

$$\mathcal{P}(N_1 = n_1, \dots, N_K = n_K) = \mathcal{P}\left(n = \sum_{i=1}^K n_i\right) \cdot \mathcal{P}\left(N_1 = n_1, \dots, N_K = n_K \mid n = \sum_{i=1}^K n_i\right)$$

$$\mathcal{P}(N_1 = n_1, \dots, N_K = n_K) = \frac{\mathcal{N}^n e^{-\mathcal{N}}}{n!} \cdot \frac{n!}{\prod_{j=1}^K n_j!} \prod_{j=1}^K \left(\frac{l_j}{\mathcal{N}}\right)^{n_j}$$

Wobei $n = \sum_{i=1}^K n_i$. Mit $\mathcal{N} = \sum_{j=1}^K l_j$ folgt:

$$\mathcal{P}(N_1 = n_1, \dots, N_K = n_K) = \prod_{j=1}^K \frac{e^{-l_j} l_j^{n_j}}{n_j!}$$

Man sieht letztlich, dass die gemeinsame Verteilung der Zufallsvariablen, die die Anzahl der Einträge in den Bins beschreibt, in K unabhängige poissonverteilte Zufallsvariablen mit Parameter l_j faktorisiert. Andererseits lässt sich sagen, dass dies gleichbedeutend mit der Beschreibung ist, die Gesamtanzahl der Ereignisse Poisson-verteilt zu betrachten und mit der Multinomialverteilung die Verteilung der n Einträge in die Bins zu beschreiben.

Die erweiterte Maximum-Likelihood-Methode für gebinnte Daten mit K Bins besteht also darin, das Produkt aus K unabhängigen Poisson-verteilten Zufallsvariablen in Abhängigkeit der zu schätzenden Parameter Φ_0, α und λ zu maximieren.

4.2 Untersuchung der Schätzfunktionen

Im Folgenden werden die Verteilungen der Schätzfunktionen $\hat{\Phi}_0$, $\hat{\alpha}$ und $\hat{\lambda}$ für diverse Stichprobengrößen jeweils $N = 1000$ simuliert. Die vorgegebenen Werte sind so gewählt, dass sie denen aus dem vorherigen Kapitel entsprechen: $\alpha_0 = 2.0$ und $\lambda_0 = 0.1$. Die entsprechenden Werte für Φ_0 ergeben sich wiederum aus den beiden Parameterwerten und der Stichprobengröße aufgrund der Normierung auf die Stichprobengröße n . Mit den Ergebnissen für die Schätzfunktionen, zusammengefasst in Tabelle 2, lassen sich diese auf ihre Qualität bewerten.

n	$\bar{\hat{\Phi}}_0$	$\hat{\sigma}_{\hat{\Phi}_0}$	$\bar{\hat{\alpha}}$	$\hat{\sigma}_{\hat{\alpha}}$	$\bar{\hat{\lambda}}$	$\hat{\sigma}_{\hat{\lambda}}$
300	33.4	5.9	1.97849	0.10276	0.144452	0.121614
1 000	107.0	10.0	1.99465	0.05377	0.111583	0.052735
3 000	318.6	17.1	1.99627	0.03131	0.105411	0.031493
10 000	1055.3	31.7	1.99916	0.01742	0.101451	0.016605
30 000	3161.8	52.7	1.99958	0.00985	0.100443	0.008969
100 000	10534.7	96.1	1.99976	0.00542	0.100173	0.004789

Tabelle 2: Simulierte Mittelwerte und Standardabweichungen für die drei Parameter Φ_0 , α und λ für verschiedene Stichprobengrößen n .

Aus Tabelle 2 kann man ablesen, dass die geschätzten Werte für α und λ sich mit wachsender Stichprobengröße an die wahren Werte annähern und ihre Streuung abnimmt. Sie erfüllen damit die Bedingung konsistenter Schätzer.

Dabei nähert sich $\hat{\lambda}$ von oben an λ_0 an, die Werte von $\bar{\hat{\alpha}}$ nähern sich hingegen von unten an α_0 an. Mit anderen Worten lässt sich feststellen: $\hat{\lambda}_n > \hat{\lambda}_{n'} > \lambda_0$ und $\bar{\hat{\alpha}}_n < \bar{\hat{\alpha}}_{n'} < \alpha_0$, für verschiedene Stichprobengrößen $n' > n$. Damit zeigt sich auch, dass das Schätzverfahren asymptotisch erwartungstreue Schätzfunktionen für die Parameter α und λ liefert. Die Bewertung des Schätzers für Φ_0 wird hier nicht diskutiert.

Dass die Werte von $\hat{\alpha}$ einerseits auf ihren Mittelwert bezogen für alle Stichprobengrößen n unterhalb des wahren Wertes liegen, die von $\hat{\lambda}$ hingegen oberhalb, spricht analog zu Beobachtungen aus Kapitel 3, für eine Antikorrelation der beiden Schätzfunktionen.

Abbildung 6 zeigt, analog zu den Ergebnissen aus Kapitel 3, innerhalb der Auftragung der Standardabweichungen gegenüber der Stichprobengröße außerdem, dass hier ein linearer Zusammenhang mit Steigung von ca.-0.5 vorliegt, somit gilt wiederum für beide Schätzfunktionen: $\hat{\sigma} \propto \frac{1}{\sqrt{n}}$.

Der Korrelationskoeffizient $c_{\hat{\alpha}, \hat{\lambda}} = -0.72$ liegt damit leicht unter dem bei gebinnten Daten.

Die Abbildung 5 zeigt die simulierten Verteilungen für die Stichprobengrößen $n = 300$ und $n = 3000$. Die simulierten Verteilungen für $\hat{\Phi}_0$ und $\hat{\alpha}$ zeigen für $n = 300$ recht symmetrische Verteilungen, die sich mit der Normalverteilung approximieren lassen. Die simulierte Verteilung für λ zeigt für $n = 300$ dabei eine stark asymmetrische, rechtsschiefe Verteilung, die sich erst ab circa $n = 3000$ als annähernd symmetrisch und normalverteilt behandeln lässt. Wie gut in den einzelnen Fällen die Approximation mit der Normalverteilung ist, wurde nicht im Detail untersucht.

Die Abbildung 6 zeigt außerdem, dass sich der Zusammenhang der mittleren quadratischen

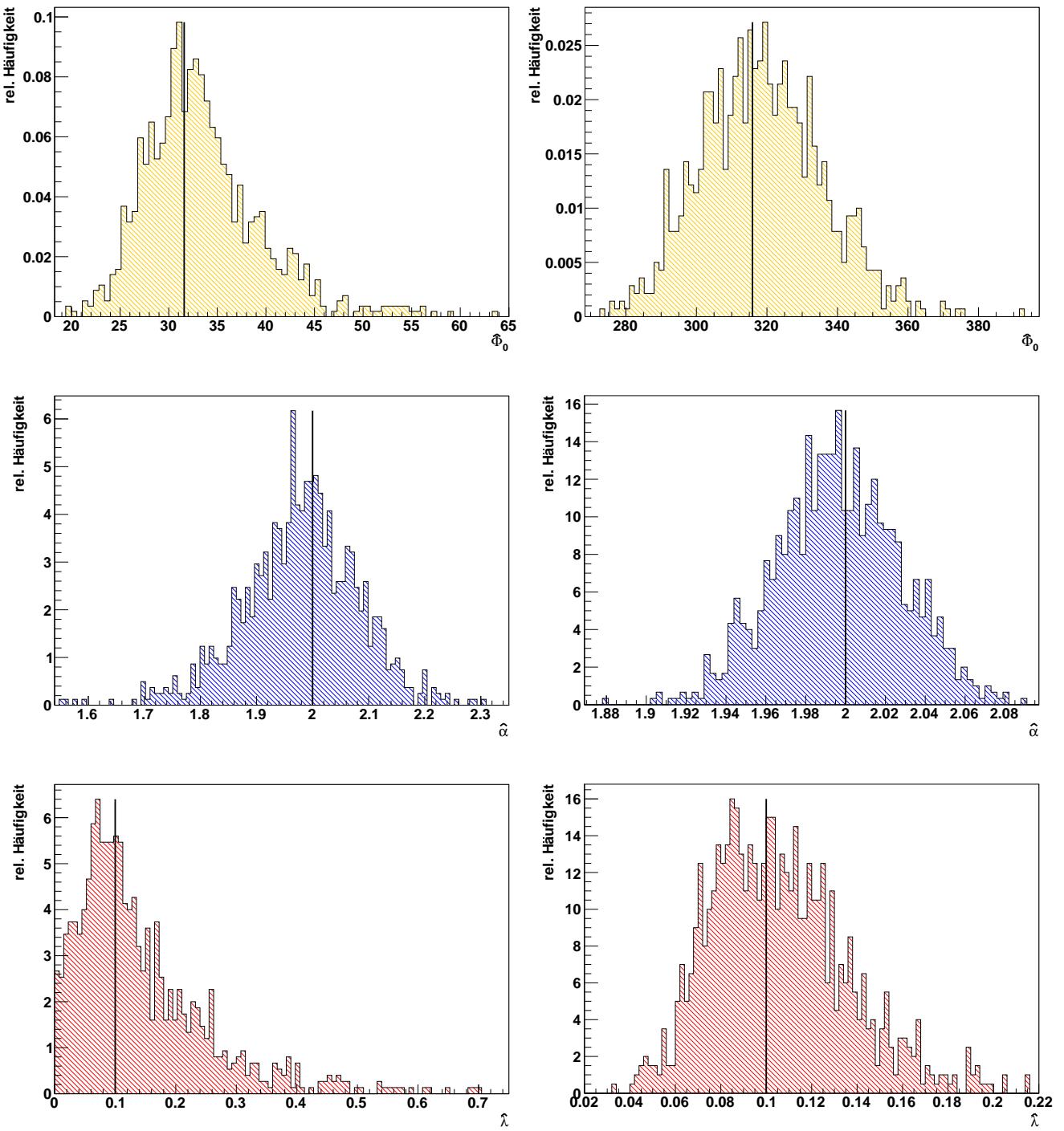


Abbildung 5: Die obigen Plots zeigen die simulierten relativen Häufigkeiten der Schätzfunktionen für Φ_0 (gelb), α (blau) und λ (rot) für $N=1000$. Links sind dabei jeweils die Verteilungen für eine Stichprobengröße von $n=300$ und rechts die für $n=3000$ dargestellt. Ab der Stichprobengröße $n=3000$ lässt sich eine Approximation der simulierten Verteilung mit einer Normalverteilung durchführen.

Fehler und der Stichprobengröße n linear annähern lässt.

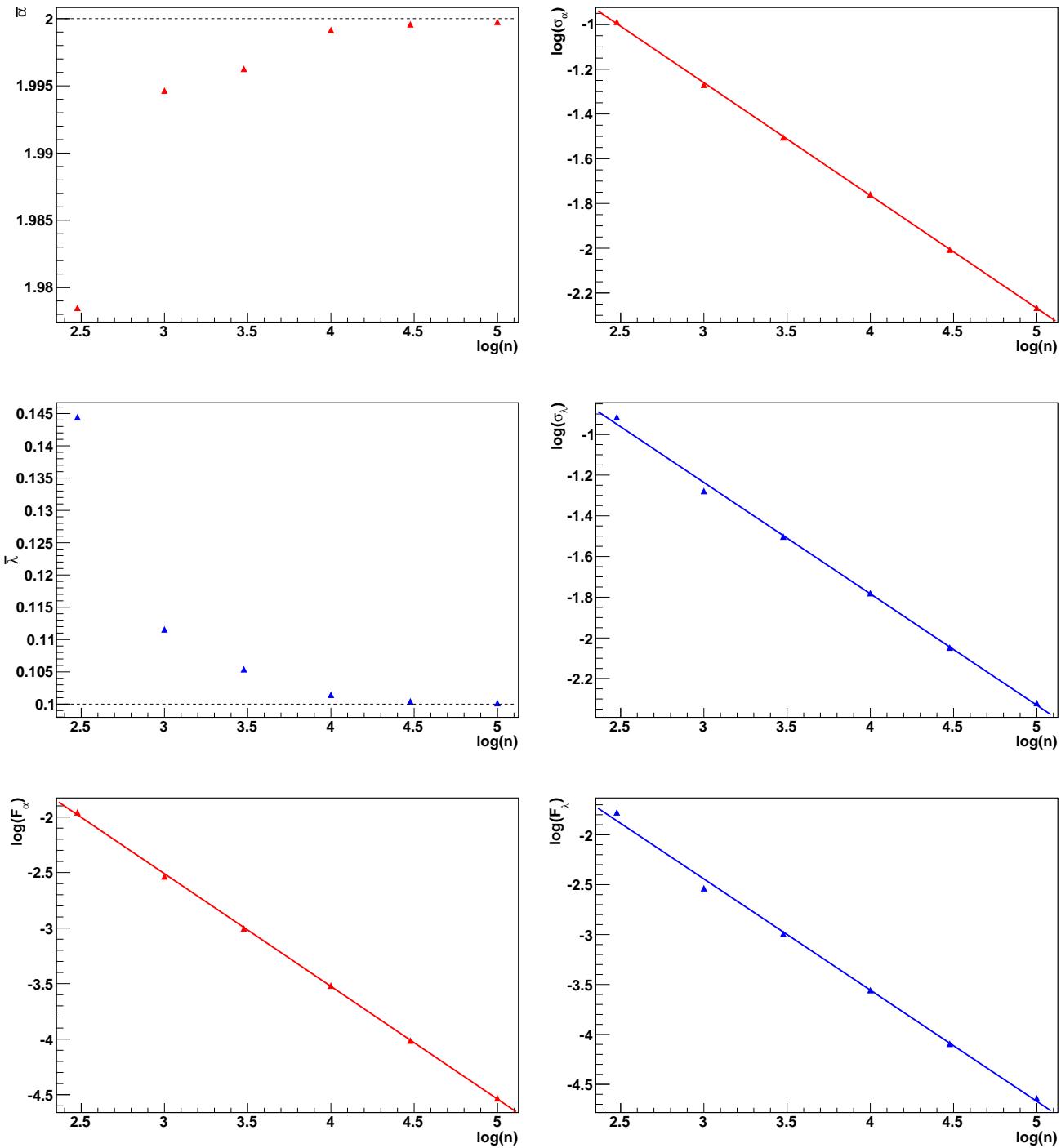


Abbildung 6: Die Plots zeigen die Werte für $\hat{\alpha}$ (rot) und $\hat{\lambda}$ (blau) aus Tabelle 2 zur grafischen Veranschaulichung. Aufgetragen sind auf der x-Achse der Logarithmus der Stichprobengröße n gegenüber dem arithmetischen Mittelwert bzw. Standardabweichung der jeweilige Schätzfunktionen. Die gestrichelten Linien in den beiden Plots von $\bar{\alpha}$ und $\bar{\lambda}$ zeigen jeweils die wahren Werte α_0 und λ_0 an. Die untersten beiden Plots zeigen den mittleren quadratischen Fehler F beider Schätzfunktionen, der sich aus der Summe aus dem quadrierten Bias und geschätzter Varianz der jeweiligen Schätzfunktion ergibt. Die Standardabweichungen und die quadratischen Fehler lassen sich dabei in dieser doppelt-logarithmischen Auftragung als Gerade fitten.

4.3 Vergleich der erweiterten Maximum-Likelihood-Methode hinsichtlich ungebinnter und gebinnter Daten

Zum Abschluss der Kapitel über die erweiterte Maximum-Likelihood-Methode, werden die sich daraus ableitenden Schätzfunktionen anhand ihrer Schätzqualität bei ungebinnten und gebinnten Daten verglichen. Um hier einen sinnvollen Vergleich treffen zu können, wurden bei beiden Versionen, ungebinnt und gebinnt, die Schätzqualität anhand von Simulationen mit den gleichen Parameterwerten und Stichprobengrößen vollzogen.

Wie bereits erwähnt, bedeutet das Binning der gemessenen Daten ein Informationsverlust. Der genaue Energiewert eines Gammaphotons liegt also nicht mehr vor. Damit könnte man für die Schätzmethoden bei ungebinnten Daten erwarten, dass die Werte der Parameter α und λ einen kleineren Bias bei einer bestimmten Stichprobengröße und geringerer zugehöriger Streuung im Vergleich zu den gebinnten Daten aufweisen.

n	$\bar{\hat{\Phi}}_0$	$\hat{\sigma}_{\hat{\Phi}_0}$	$\bar{\hat{\alpha}}$	$\hat{\sigma}_{\hat{\alpha}}$	$\bar{\hat{\lambda}}$	$\hat{\sigma}_{\hat{\lambda}}$
300	44.8	239.6	1.98439	0.09503	0.140725	0.219383
1 000	107.0	10.0	1.99461	0.05365	0.111627	0.052620
3 000	318.6	17.1	1.99627	0.03127	0.105384	0.030106
10 000	1055.3	31.6	1.99915	0.01739	0.101442	0.016565
30 000	3161.9	52.7	1.99956	0.00984	0.100449	0.008970
100 000	10534.8	95.91	1.99975	0.00540	0.100172	0.004782

Tabelle 3: Ungebinnte Daten: Arithmetische Mittelwerte und Standardabweichungen der simulierten Verteilungen der drei Schätzfunktionen $\hat{\Phi}_0$, $\hat{\alpha}$ und $\hat{\lambda}$ für verschiedene Stichprobengrößen n und einer Simulationsgröße $N=1000$ bei den vorgegebenen Parametern $\alpha_0 = 2.0$ und $\lambda_0 = 0.1$.

n	$\bar{\hat{\Phi}}_0$	$\hat{\sigma}_{\hat{\Phi}_0}$	$\bar{\hat{\alpha}}$	$\hat{\sigma}_{\hat{\alpha}}$	$\bar{\hat{\lambda}}$	$\hat{\sigma}_{\hat{\lambda}}$
300	33.4	5.9	1.97849	0.10276	0.144452	0.121614
1 000	107.0	10.0	1.99465	0.05377	0.111583	0.052735
3 000	318.6	17.1	1.99627	0.03131	0.105411	0.031493
10 000	1055.3	31.7	1.99916	0.01742	0.101451	0.016605
30 000	3161.8	52.7	1.99958	0.00985	0.100443	0.008969
100 000	10534.7	96.1	1.99976	0.00542	0.100173	0.004789

Tabelle 4: Gebinnte Daten: Arithmetische Mittelwerte und Standardabweichungen der simulierten Verteilungen der drei Schätzfunktionen $\hat{\Phi}_0$, $\hat{\alpha}$ und $\hat{\lambda}$ für verschiedene Stichprobengrößen n und einer Simulationsgröße $N=1000$ bei den vorgegebenen Parametern $\alpha_0 = 2.0$ und $\lambda_0 = 0.1$.

Vergleicht man die Ergebnisse der Simulationen aus Kapitel 3 und 4, siehe Tabelle 3 und 4, so lässt sich erkennen, dass sich nur für eine Stichprobengröße von $n = 300$ die Ergebnisse, also die arithmetischen Mittelwerte und Standardabweichungen sich voneinander merklich unterscheiden. $\hat{\alpha}$ zeigt bei ungebinnten Daten einen kleineren Bias und kleinere Standardabweichung im Vergleich zu den gebinnten Daten, $\hat{\lambda}$ zeigt bei ungebinnten einen leicht kleineren Bias, aber eine deutlich größere Standardabweichung im Vergleich zur Schätzmethode bei gebinnten Daten. Auf diese beiden Parameter bezogen, gibt es bei $n = 300$ keine Methode, die in allen Kriterien besser ist. Für größere Stichproben zeigen sich nur noch vernachlässigbare Unterschiede. So lässt

sich festhalten, dass die Schätzmethoden $\hat{\alpha}$ und $\hat{\lambda}$ bei gebinnten Daten vor allem bei Stichprobengrößen ab $n = 1000$ trotz des Informationsverlustes qualitativ gleichwertige Schätzfunktionen sind im Vergleich zu denen bei ungebinnten Daten.

5 Ersetzungsmethode

Eine alternative Methode zur erweiterten Maximum-Likelihood-Methode zur Schätzung der Parameter α und λ des Spektrums $g(x) = \Phi_0 x^{-\alpha} e^{-\lambda x}$, $x \in [m, m']$, beruht auf der Ersetzungsmethode. Dazu betrachten wir im Folgenden nicht mehr die Funktion $g(x)$, sondern wechseln zur Dichtefunktion $f(x) = C(\alpha, \lambda) x^{-\alpha} e^{-\lambda x}$, $x \in [m, m']$, siehe Kapitel 3. Außerdem werden wir den Parameter m , also das Minimum von \mathcal{X} , als unbekannt ansehen. Dies würde der Situation im Experiment entsprechen, dass man die niedrigste Energie der Photonen einer Gammastrahlungsquelle nicht kennt. Die obere Grenze, entsprechend der höchsten Energie einer Gammaquelle, setzen wir hier auf $m' = \infty$. Dies entspricht zwar nicht der physikalischen Realität, allerdings hat bei der oberen Grenze der erweiterten Maximum-Likelihood-Methode von 10 TeV, die Funktion $f(x)$ in guter Näherung wie bei ∞ den Wert 0.

So erhält man $f(x) = C(m, \alpha, \lambda) x^{-\alpha} e^{-\lambda x}$, $x \in [m, \infty)$, wobei die Normierungskonstante nun auch vom unbekannten Parameter m abhängt. Die zugehörige Verteilung wird mit $\Gamma_{m, \alpha, \lambda}$ bezeichnet. Das prinzipielle Vorgehen der Schätzung der drei Parameter m , α und λ im Kontext der Ersetzungsmethode liegt darin, dass zwei theoretische Kenngrößen der $\Gamma_{m, \alpha, \lambda}$ -Verteilung durch ihre Pendants aus der Stichprobe ersetzt und dann miteinander in Beziehung gebracht werden, um so aus dem Informationsgehalt einer Stichprobe (x_1, \dots, x_n) , gleichzeitig zwei Parameter, also α und λ zu schätzen. Der dritte Parameter m wird einzeln und getrennt von den anderen Parametern geschätzt.

Um sich diese Methode zunächst an einem einfachen Beispiel klar zu machen, wird das Prinzip an der bekannten Gammaverteilung $\gamma(p, b)$ veranschaulicht.

Die Gammaverteilung $\gamma(p, b)$ zeigt nicht nur große Ähnlichkeit zur $\Gamma_{m, \alpha, \lambda}$ -Verteilung in Bezug auf die Wahrscheinlichkeitsdichte, der Rechtsschiefe und der Tatsache von mehreren Parametern bestimmt zu werden, mit ihr lässt sich die Methode auch analytisch zeigen.

Wir wechseln also in das statistische Modell der Gammaverteilung $\mathcal{M}_{\gamma_{p,b}} = ((0, \infty), \mathcal{B}(0, \infty), \mathcal{P}_{(p,b)} : (p, b) \in \mathbb{R}^+ \times \mathbb{R}^+)$ mit Dichtefunktion: $f_{p,b}(x) = \frac{b^p}{\Gamma(p)} x^{p-1} e^{-bx}$, $x > 0$, bzw. bei gegebener Stichprobe (x_1, \dots, x_n) in das n-fache Produktmodell. Die zugehörige Dichte ist in Abbildung 7 an einem Beispiel für $p=2$ und $b=3$ dargestellt.

Im Folgenden sei eine Zufallsvariable X gammaverteilt mit Parametern (p, b) . Der Erwartungswert ist gegeben mit $\mathbb{E}(X) = \frac{p}{b}$ und die Varianz mit $\mathbb{V}(X) = \frac{p}{b^2}$. Angenommen der Erwartungswert und die Varianz einer gammaverteilten Zufallsvariable seien bekannt. Im Folgenden werden damit Schätzwerte für die Parameterwerte (p, b) bestimmt.

Aus der Definition des Erwartungswertes und der Varianz ergeben sich durch Umstellen für beide Parameter jeweils eine Bestimmungsgleichung:

$$\left. \begin{aligned} \mathbb{E}(X) &= \frac{p}{b} \\ \mathbb{V}(X) &= \frac{p}{b^2} \end{aligned} \right\} \Rightarrow \begin{aligned} b &= \frac{\mathbb{E}(X)}{\mathbb{V}(X)} \\ p &= \frac{\mathbb{E}(X)^2}{\mathbb{V}(X)} \end{aligned}$$

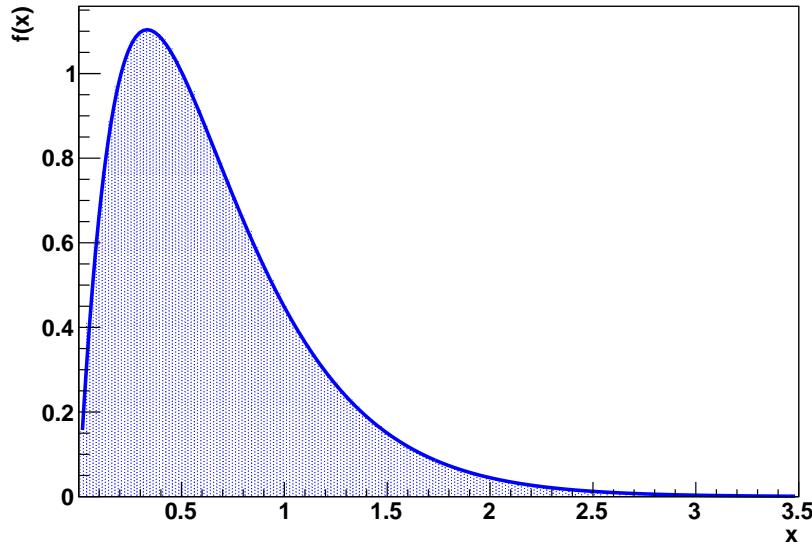


Abbildung 7: Beispiel einer Gammaverteilung mit zugehöriger Wahrscheinlichkeitsdichte für $p=2, b=3$.

Diesen Zusammenhang kann man auch grafisch, siehe Abbildung 8, darstellen. Stellt man zunächst so um, dass man aus der Definition für Erwartungswert und für Varianz Funktionengleichungen für p erhält, also $p(b) = b \cdot \mathbb{E}(X)$ und $p(b) = b^2 \cdot \mathbb{V}(X)$, so ist das wahre Parameterpaar (p, b) genau der Schnittpunkt beider Kurven:

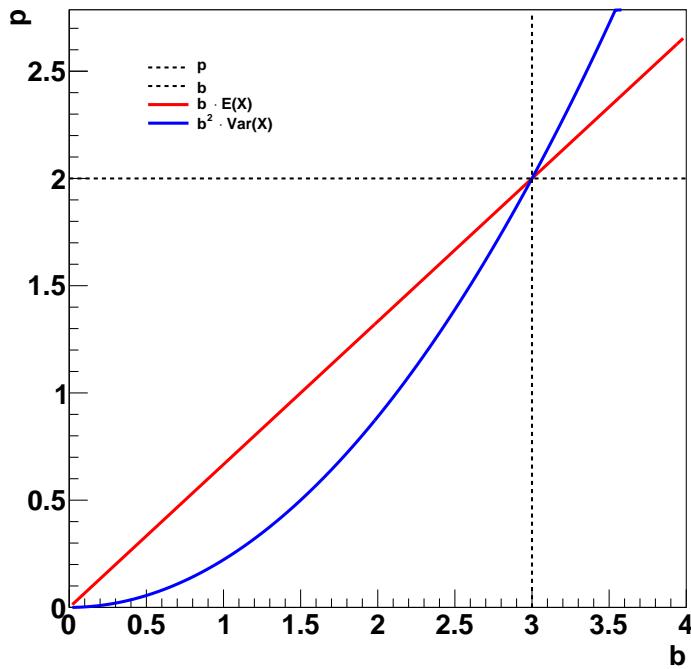


Abbildung 8: Der Schnittpunkt von Gerade und Parabel liegt exakt beim wahren Parameterpaar, hier $p=2, b=3$.

Genau dieser Zusammenhang lässt sich nun mittels der Ersetzungsmethode ausnutzen und so

beide Parameter gleichzeitig anhand einer Stichprobe schätzen.

Ersetzt man nun den Erwartungswert $\mathbb{E}(X)$ mit dem arithmetischen Mittel der konkreten Stichprobe $\hat{\mu} := \bar{x}_n = \frac{1}{n} \sum_{i=1}^n x_i$ und die Varianz $\mathbb{V}(X)$ mit der korrigierten Stichprobenvarianz der konkreten Stichprobe $\hat{\sigma}^2 := s^2(n) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$ erhält man folgende Schätzwerte für die Parameter:

$$\hat{b} = \frac{\hat{\mu}}{\hat{\sigma}^2} \quad \text{und} \quad \hat{p} = \frac{\hat{\mu}^2}{\hat{\sigma}^2}$$

Als nächstes wird die Qualität dieser Schätzer untersucht. Es handelt sich bei dem arithmetischen Mittel und der Stichprobenvarianz um beste Schätzer für den Erwartungswert und die Varianz einer Zufallsvariablen, aber dies gilt nicht automatisch für Funktionen von diesen wie in dieser Situation.

Die Plots aus Abbildung 9 zeigen die simulierten Verteilungen von $\hat{\mu}$, $\hat{\sigma}^2$, \hat{p} , \hat{b} für $N = 1000$ und einer Stichprobengröße von $n = 1000$. Als Parameterwerte wurden wieder $p = 2$ und $b = 3$ gewählt, zugehöriger Erwartungswert lautet $\mathbb{E}(X) = \frac{2}{3}$ und zugehörige Varianz $\mathbb{V}(X) = \frac{2}{9}$.

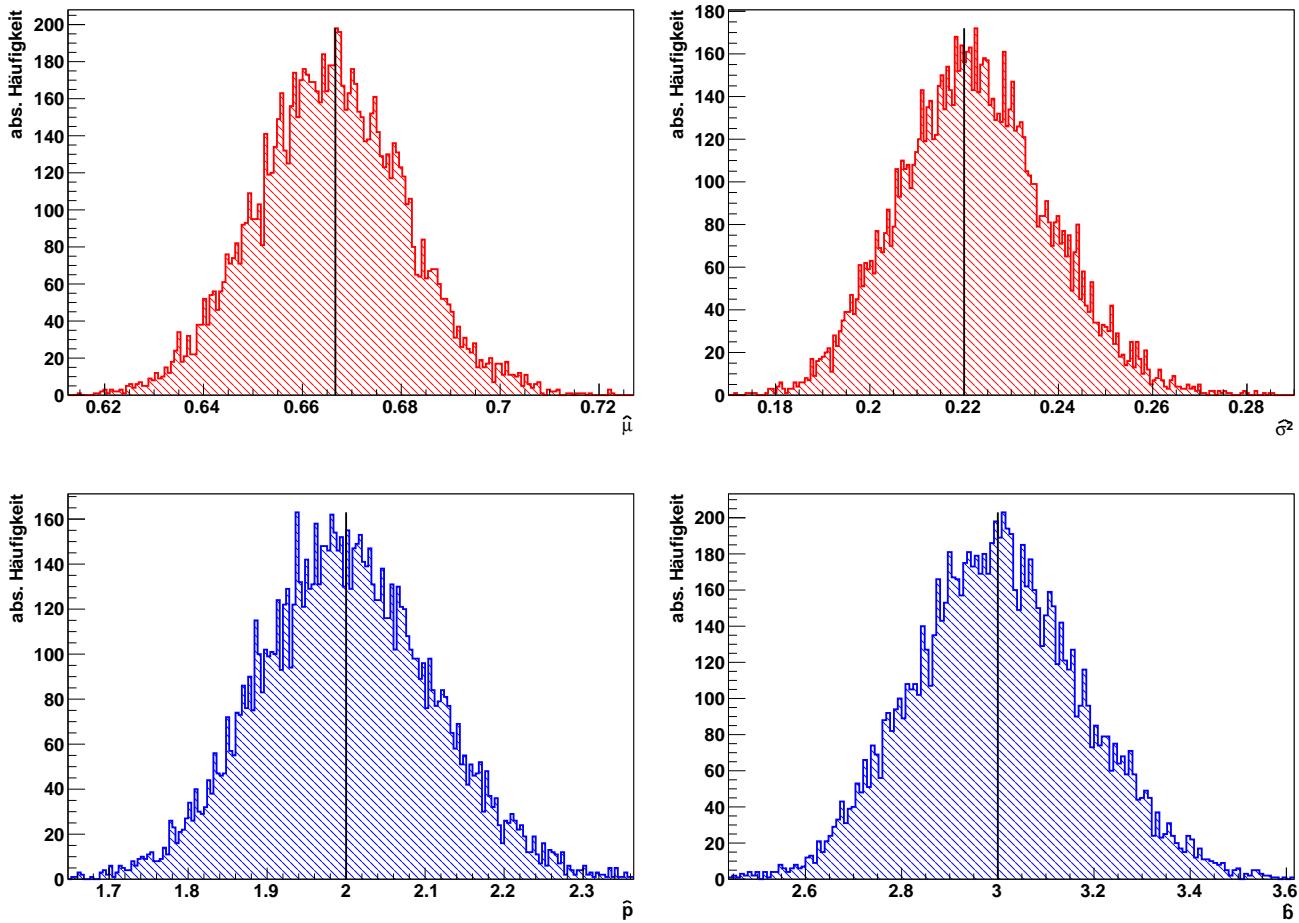


Abbildung 9: Simulierte absolute Häufigkeiten für eine Stichprobengröße $n=1000$. Die obere Reihe (rot) zeigt links die simulierten Verteilungen des arithmetischen Mittelwerts $\hat{\mu}$ und rechts die der Stichprobenvarianz $\hat{\sigma}^2$. In der unteren Reihe (blau) sind die zugehörigen simulierten Verteilungen für \hat{p} links und die für \hat{b} rechts dargestellt. Die simulierten Verteilungen lassen sich aufgrund der Form und der Symmetrie gut mit einer Normalverteilung approximieren. Wichtige Stichprobenkenngroßen und deren Fehler lassen sich aus den nachfolgenden Tabellen 5 und 6 entnehmen.

In der nachfolgenden Tabelle 5 sind die zugehörigen Mittelwerte der simulierten Verteilungen

von $\hat{\mu}$, $\hat{\sigma}^2$, \hat{p} und \hat{b} aufgelistet.

n	$\bar{\hat{\mu}} \pm \sigma_{\bar{\hat{\mu}}}$	$\bar{\hat{\sigma}^2} \pm \sigma_{\bar{\hat{\sigma}^2}}$	$\bar{\hat{p}} \pm \sigma_{\bar{\hat{p}}}$	$\bar{\hat{b}} \pm \sigma_{\bar{\hat{b}}}$
100	0.66382 ± 0.00149	0.21893 ± 0.00158	2.08035 ± 0.00391	3.14952 ± 0.01854
1 000	0.66678 ± 0.00046	0.22265 ± 0.00052	2.00315 ± 0.00040	3.00369 ± 0.00580
10 000	0.66655 ± 0.00015	0.22217 ± 0.00016	2.00122 ± 0.00039	3.00198 ± 0.00182
100 000	0.66671 ± 0.00048	0.22220 ± 0.00005	2.00059 ± 0.000004	3.00069 ± 0.00058

Tabelle 5: Auflistung der Mittelwerte und zugehörige Fehler der simulierten Verteilungen von $\hat{\mu}$, $\hat{\sigma}^2$, \hat{p} und \hat{b} .

Anhand der Werte aus der Tabelle 5 lässt sich feststellen, dass für alle vier Schätzwerte der Abstand zum wahren Wert mit wachsendem n immer kleiner wird. Unter Berücksichtigung der Werte für die Standardabweichungen der simulierten Verteilungen in Tabelle 6, sieht man auch hier, dass die Schätzwerte mit zunehmenden Stichprobenumfang immer weniger um den wahren Wert streuen.

Die Schätzwerte für μ und σ^2 zeigen schon bei n=100 eine geringe Abweichung vom wahren Wert, gemessen an ihren Fehlern. Dies spricht für die Erwartungstreue dieser Schätzfunktionen. Die hohe Abweichung von \hat{p} und \hat{b} von ihren wahren Werten bei n=100 spricht hingegen für keine erwartungstreue Schätzfunktionen, beurteilt nach den zugehörigen Fehlern. So ist der Abstand vom wahren Wert jeweils zu groß um sie mit einer statistischen Fluktuation der Schätzfunktionen zu erklären. \hat{p} und \hat{b} sind daher nicht erwartungstreu.

Mit den Ergebnissen der Simulationen lässt sich jedoch sagen, dass die Schätzfunktionen für p und b asymptotisch erwartungstreu sind.

n	$\sigma_{\hat{\mu}}$	$\sigma_{\hat{\sigma}^2}$	$\sigma_{\hat{p}}$	$\sigma_{\hat{b}}$
100	0.04734	0.04886	0.121573	0.573308
1 000	0.01521	0.01596	0.012427	0.179950
10 000	0.00474	0.00498	0.001196	0.056233
100 000	0.00149	0.00161	0.000127	0.018231

Tabelle 6: Auflistung der zugehörigen Standardabweichungen der simulierten Verteilungen von $\hat{\mu}$, $\hat{\sigma}^2$, \hat{p} und \hat{b} . Die Fehler auf die Mittelwerte der Schätzfunktionen aus Tabelle 5 ergeben sich aus den zugehörigen Werten dieser Tabelle mit der Multiplikation von $\frac{1}{\sqrt{N}}$ und sind in Tabelle 5 gelistet.

Die Ersetzungsmethode bringt hier also konsistente Schätzfunktionen für die beiden Parameter p und b der Gammaverteilung hervor.

Bei der $\Gamma_{m,\alpha,\lambda}$ -Verteilung ergibt sich hingegen keine direkte analytische Lösung für die Schätzfunktionen. Hier gilt es ein numerisches Verfahren zu entwickeln. Doch zunächst zu den mathematischen Vorbereitungen. Das statistische Modell ist folgendermaßen gegeben:

$$(\mathcal{X}, \mathcal{A}, \mathbb{P}) = ([m, \infty), \mathcal{B}([m, \infty)), \mathcal{P}_{m,\lambda,\alpha} : m, \lambda, \alpha \in [0, \infty)).$$

Des Weiteren existiere die Zufallsvariable $X(\omega)$ mit Werten in \mathcal{X} , diese ist auf einem geeigneten Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathcal{P})$ definiert, sodass das Wahrscheinlichkeitsmaß gerade dem

Bildmaß der Zufallsvariablen entspricht. Der zugrunde liegende Wahrscheinlichkeitsraum ist hier nicht interessant. Die Wahrscheinlichkeitsdichte, die sich aus „ $x^{-\alpha} \cdot \exp(-\lambda \cdot x)$ “ ergibt, ist mit $f_X(x) = C(m, \alpha, \lambda) \cdot x^{-\alpha} e^{-\lambda \cdot x}$ für $x \in [m, \infty)$ gegeben. Diese Dichte beschreibt das Maß $\mathcal{P}_{m, \lambda, \alpha}$ eindeutig.

Integriert man diese Dichte über ihren Definitionsbereich so erhält man als Normierungskonstante $C(m, \alpha, \lambda) = \frac{\lambda^{1-\alpha}}{\Gamma(1-\alpha, m \cdot \lambda)}$. Innerhalb des Integrals wurde hier $x \rightarrow x \cdot \lambda$ substituiert, so ergibt sich die unvollständige Gammafunktion, die folgendermaßen definiert ist:

$$\Gamma(x, y) = \int_y^\infty s^{x-1} e^{-s} ds, \quad x \in \mathbb{R}, y \in \mathbb{R}^+$$

Im Folgenden sind die wichtigsten Kenngrößen von X gelistet:

- **Wahrscheinlichkeitsdichte** $f_X(x)$

$$f_X(x) = \mathbf{1}_{[m, \infty)}(x) \frac{\lambda^{1-\alpha}}{\Gamma(1-\alpha, \lambda \cdot m)} \cdot x^{-\alpha} \cdot e^{-\lambda \cdot x}, \quad x \in \mathbb{R}$$

- **Erwartungswert** $\mathbb{E}(X)$

$$\begin{aligned} \mathbb{E}(X) &= \int_{\mathbb{R}} x f_X(x) dx = \frac{\lambda^{1-\alpha}}{\Gamma(1-\alpha, \lambda \cdot m)} \int_m^\infty x \cdot x^{-\alpha} \cdot e^{-\lambda \cdot x} dx = \\ &= \frac{1}{\lambda} \cdot \frac{\Gamma(2-\alpha, m \cdot \lambda)}{\Gamma(1-\alpha, \lambda \cdot m)} = m \cdot \frac{E_{\alpha-1}(m \cdot \lambda)}{E_\alpha(m \cdot \lambda)} \end{aligned}$$

Wobei $E_n(x) = \int_1^\infty e^{-t \cdot x} x^{-n} dt = x^{n-1} \Gamma(1-n, x)$.

- **Verteilungsfunktion** $F_X(t)$

$$F_X(t) = \mathcal{P}(X \leq t) = \int_m^t f_X(x) dx = 1 - \frac{\Gamma(1-\alpha, \lambda \cdot t)}{\Gamma(1-\alpha, \lambda \cdot m)}, \quad t \in (m, \infty)$$

Für den Median t_m der Verteilung ergibt sich der nicht weiter vereinfachbare Ausdruck:

$$\frac{\Gamma(1-\alpha, \lambda \cdot t_m)}{\Gamma(1-\alpha, \lambda \cdot m)} = \frac{1}{2}$$

Der Wert des Medians lässt deshalb nur numerisch bestimmen:

$$\min_{y \in (m, \infty)} \left| F_X(y) - \frac{1}{2} \right| = 0 \Rightarrow y = t_m$$

- **Varianz** $\mathbb{V}(X)$

Als k-te Momente ergeben sich:

$$\mathbb{E}(X^k) = m^k \cdot \frac{E_{\alpha-k}(m \cdot \lambda)}{E_\alpha(m \cdot \lambda)}, \quad k \in \mathbb{N}$$

Daraus folgt für die Varianz:

$$\mathbb{V}(X) = m^2 \cdot \left[\frac{E_{\alpha-2}(m \cdot \lambda)}{E_\alpha(m \cdot \lambda)} - \left(\frac{E_{\alpha-1}(m \cdot \lambda)}{E_\alpha(m \cdot \lambda)} \right)^2 \right]$$

Im Folgenden werden die wahren Werte der Verteilung mit $(m_0, \alpha_0, \lambda_0)$ bezeichnet, anhand derer Stichproben (x_1, \dots, x_n) gezogen werden. Die Schätzwerte benennen wir mit $(\hat{m}, \hat{\alpha}, \hat{\lambda})$. Es ergeben sich für das Paar arithmetisches Mittel und Stichprobenmedian sowie das Paar arithmetisches Mittel und Varianz jeweils das gleiche Schätzverfahren nur mittels anderer empirischer Kenngrößen. Das Schätzverfahren wird im Folgenden detailliert vorgestellt.

1. Ausgehend von der Stichprobe bildet man das arithmetische Mittel \bar{x}_n , die Stichprobenvarianz $s_{(n)}^2$, bestimmt den Median $\tilde{t}_{m,n}$ und das Minimum $x_{(1)}$ der Stichprobe.
2. Man ersetzt den Erwartungswert mit dem Stichprobenmittelwert und gleichzeitig den Median mit dem Median der Stichprobe bzw. die Varianz mit der Stichprobenvarianz:

$$(\mathbb{E}(X), t_m) \rightarrow (\bar{x}_n, \tilde{t}_{m,n})$$

$$(\mathbb{E}(X), \mathbb{V}(X)) \rightarrow (\bar{x}_n, s_{(n)}^2)$$

3. Sowohl Erwartungswert, Median und Varianz können jeweils als Funktion der drei Parameter betrachtet werden. Beschreibt man diese funktionale Abhängigkeit mit den Funktionen f , g bzw. h für die jeweilige Kenngröße, lässt sich nach einem der beiden Parameter λ bzw. α auflösen, sofern der andere als konstant betrachtet wird. Der Parameter, der konstant gehalten wird, steht dabei im Index der Funktionen f , g , h . Der dritte Parameter m wird dabei gleich durch das Stichprobenminimum ersetzt. f^{-1}, g^{-1}, h^{-1} beschreiben das analytische Umformen bzw. hier das numerische Auflösen nach einem der Parameter:

$$\mathbb{E}(x_{(1)}, \alpha, \lambda; X) =: f(x_{(1)}, \alpha, \lambda) \Rightarrow \lambda = f_{x_{(1)}, \alpha}^{-1}(\lambda)$$

$$t_m(x_{(1)}, \alpha, \lambda) =: g(x_{(1)}, \alpha, \lambda) \Rightarrow \lambda = g_{x_{(1)}, \alpha}^{-1}(\lambda)$$

$$\mathbb{V}(x_{(1)}, \alpha, \lambda; X) =: h(x_{(1)}, \alpha, \lambda) \Rightarrow \lambda = h_{x_{(1)}, \alpha}^{-1}(\lambda)$$

Dazu müssen f, g, h , jeweils streng monoton sein, um so für ein Wertepaar von $(x_{(1)}, \alpha)$ einen eindeutigen Wert für λ zu erhalten. Siehe dazu Abbildungen 11-13.

4. Als nächstes legt man in einem geeigneten Bereich des Definitionsbereichs für α ein Raster R_1 an, siehe Abbildung 10, also eine Menge von Werten von α , $R_1 = \{\alpha_1, \dots, \alpha_l\}$, $l \in \mathbb{N}$, und bestimmt jeweils das zugehörige $\lambda_i = f_{x_{(1)}, \alpha_i}^{-1}(\lambda)$, $i = 1, \dots, l$, welches mit $x_{(1)}$ den Wert des Erwartungswert und des Medians bzw. der Varianz ergibt.

$$\min_y \left| \bar{x}_n - \mathbb{E}(x_{(1)}, \alpha_i, y; X) \right| = 0 \Rightarrow y = \lambda_{i, \mathbb{E}}$$

$$\min_u \left| \tilde{t}_n - t_{1/2}(x_{(1)}, \alpha_i, u) \right| = 0 \Rightarrow u = \lambda_{i, t_{1/2}}$$

$$\min_z \left| s_{(n)}^2 - \mathbb{V}(x_{(1)}, \alpha_i, z; X) \right| = 0 \Rightarrow z = \lambda_{i, \mathbb{V}}$$

Was man letztendlich erhält, sind drei Mengen mit 2-er Tupeln (α_i, λ_i) für den Mittelwert, Median und Varianz, also $\{(\alpha_1, \lambda_1), \dots, (\alpha_l, \lambda_l)\}_{\mathbb{E}}, \{(\alpha_1, \lambda_1), \dots, (\alpha_l, \lambda_l)\}_{t_{1/2}}$, und $\{(\alpha_1, \lambda_1), \dots, (\alpha_l, \lambda_l)\}_{\mathbb{V}}$. Betrachtet man nun zwei der Mengen, zum Beispiel die für den

Erwartungswert und den Median, lässt sich daraus ein Schätzwert für das wahre (α_0, λ_0) angeben. Plottet man beide Mengen, so liegt dieser Schätzwert genau da, wo sich die beiden Mengenpaare, analog zum Fall der Gammaverteilung, „schneiden“. Die Idee ist hier, dass ein Wertepaar (α, λ) gefunden werden muss, dass gleichzeitig den arithmetischen Mittelwert und den Stichprobenmedian bzw. gleichzeitig den arithmetischen Mittelwert und die Stichprobenvarianz wiedergibt. Für das Raster R_1 findet man also zunächst die Wertepaare $(\alpha_k, \lambda_{k,\mathbb{E}})$ und $(\alpha_k, \lambda_{k,t_{1/2}})$, für die gilt, dass $(\lambda_{k,\mathbb{E}}, \lambda_{k,t_{1/2}})$ den kleinsten Abstand voneinander haben. Für das arithmetische Mittel und die Stichprobenvarianz gilt dasselbe analog. :

$$\min(|\lambda_{1,\mathbb{E}} - \lambda_{1,t_{1/2}}|, \dots, |\lambda_{l,\mathbb{E}} - \lambda_{l,t_{1/2}}|) = |\lambda_{k,\mathbb{E}} - \lambda_{k,t_{1/2}}|$$

$$\min(|\lambda_{1,\mathbb{E}} - \lambda_{1,\mathbb{V}}|, \dots, |\lambda_{l,\mathbb{E}} - \lambda_{l,\mathbb{V}}|) = |\lambda_{k,\mathbb{E}} - \lambda_{k,\mathbb{V}}|$$

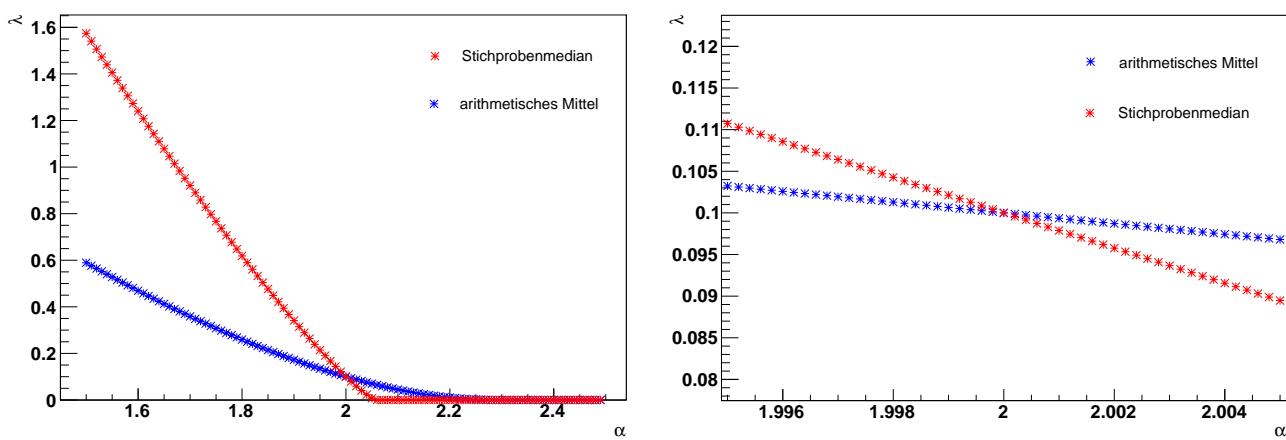


Abbildung 10: Die beiden Plots veranschaulichen Schritt 4 (links) und Schritt 5 (rechts) des Schätzverfahrens. Man wählt zunächst ein Raster auf dem ganzen Definitionsbereich von α , siehe den linken Plot, bzw. auf dem Teil des Definitionsbereichs auf dem man mit geeignetem Vorwissen abschätzen kann, dass hier der Schätzwert für α seine Werte annimmt. Dieses Raster sucht man nun nach einem Minimum des Abstandes beider Punktpaare ab. In rot sind dabei die Punkte dargestellt, die sich aus dem Stichprobenmedian schätzen lassen, in blau die Punkte, die sich aus dem arithmetischen Mittelwert schätzen lassen. Ausgehend von einer geeignet kleinen Region um dieses Minimum, wird im rechten Plot wiederum nach einem Minimum gesucht. Dieses Vorgehen lässt sich bis zur gewünschten numerischen Genauigkeit der Werte für α und λ beliebig oft wiederholen.

5. Ausgehend von dem Ergebnis aus dem Schritt zuvor, gilt es nun erneut ein Raster R_2 analog zu vorher anzulegen um die Genauigkeit an α und λ zu erhöhen.
6. Nach einigen Iterationen lassen sich nun im letzten Durchgang die Schätzwerte so bestimmen, dass man zwischen den Werten $(\lambda_{k,\mathbb{E}})$ und $(\lambda_{k,t_{1/2}})$, die für λ wiederum minimalen Abstand haben, linear interpolieren kann um so den Schätzwert $(\hat{\alpha}, \hat{\lambda})$ erhalten: Insgesamt erhält man:

$$(\hat{m}, \hat{\alpha}, \hat{\lambda}) = \left(x_{(1)}, \alpha_k, \left| \frac{\lambda_{k,\mathbb{E}} - \lambda_{k,t_{1/2}}}{2} \right| \right)$$

Damit die Ersetzungsmethode überhaupt funktionieren kann, müssen sich die Funktionen f , g und h eindeutig nach einem Wert für λ auflösen lassen, sofern der Wert von α vorgegeben wird. Das ist gleichbedeutend mit der Forderung nach der strengen Monotonie der Funktionen. Die Abbildungen 11-13 zeigen dieses Verhalten am Beispiel für einen bestimmten Bereich für α und

für bestimmte Werte für λ . Dies ist auch andersrum möglich. Es wurden also die theoretischen Kenngrößen in Abhängigkeit der Parameter aufgetragen. Diese Auftragungen wurden auch jeweils für zwei Werte von m vollzogen.

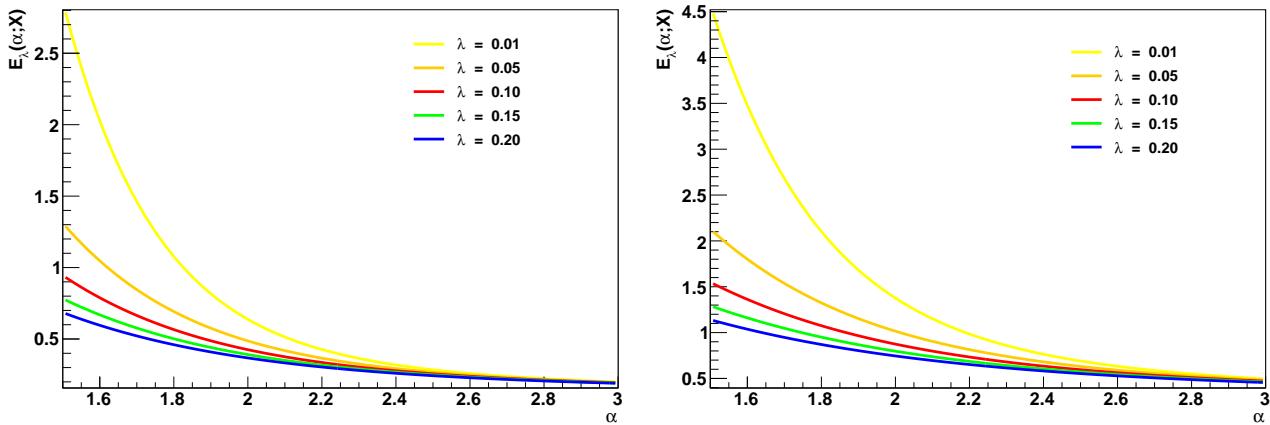


Abbildung 11: Darstellung des Erwartungswertes in Abhängigkeit von α für mehrere Werte von λ . Der Wert für m für die linke Auftragung liegt bei $m = 0.1$, für die rechte Auftragung gilt $m = 0.25$.

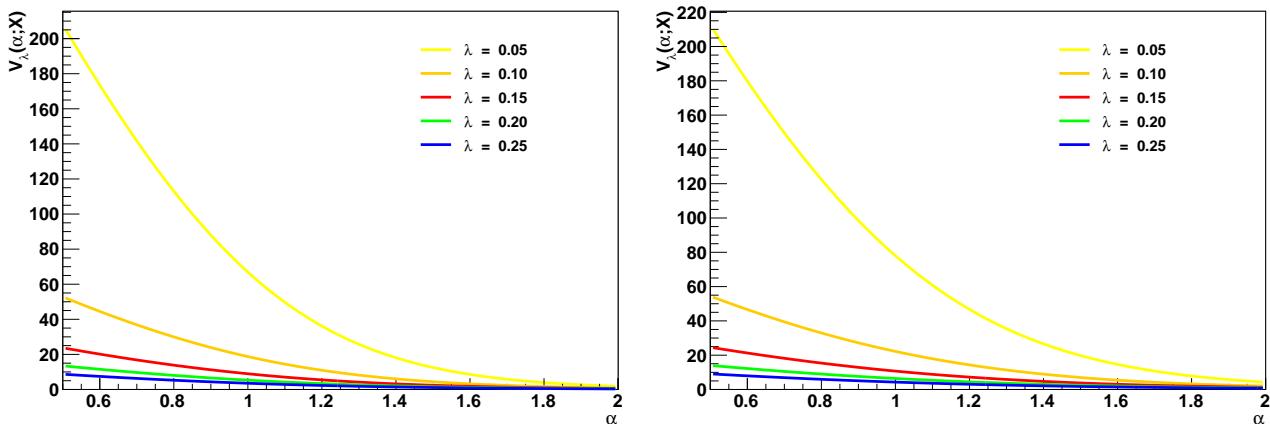


Abbildung 12: Darstellung der Varianz in Abhängigkeit von α für mehrere Werte von λ . Der Wert für m für die linke Auftragung liegt bei $m = 0.1$, für die rechte Auftragung gilt $m = 0.25$.

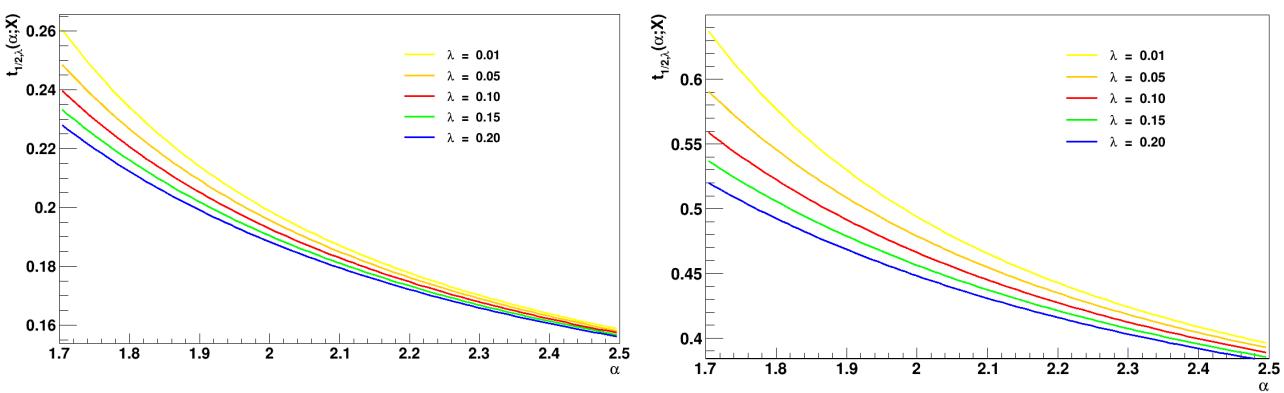


Abbildung 13: Darstellung des Medians in Abhängigkeit von α für mehrere Werte von λ . Der Wert für m für die linke Auftragung liegt bei $m = 0.1$, für die rechte Auftragung gilt $m = 0.25$.

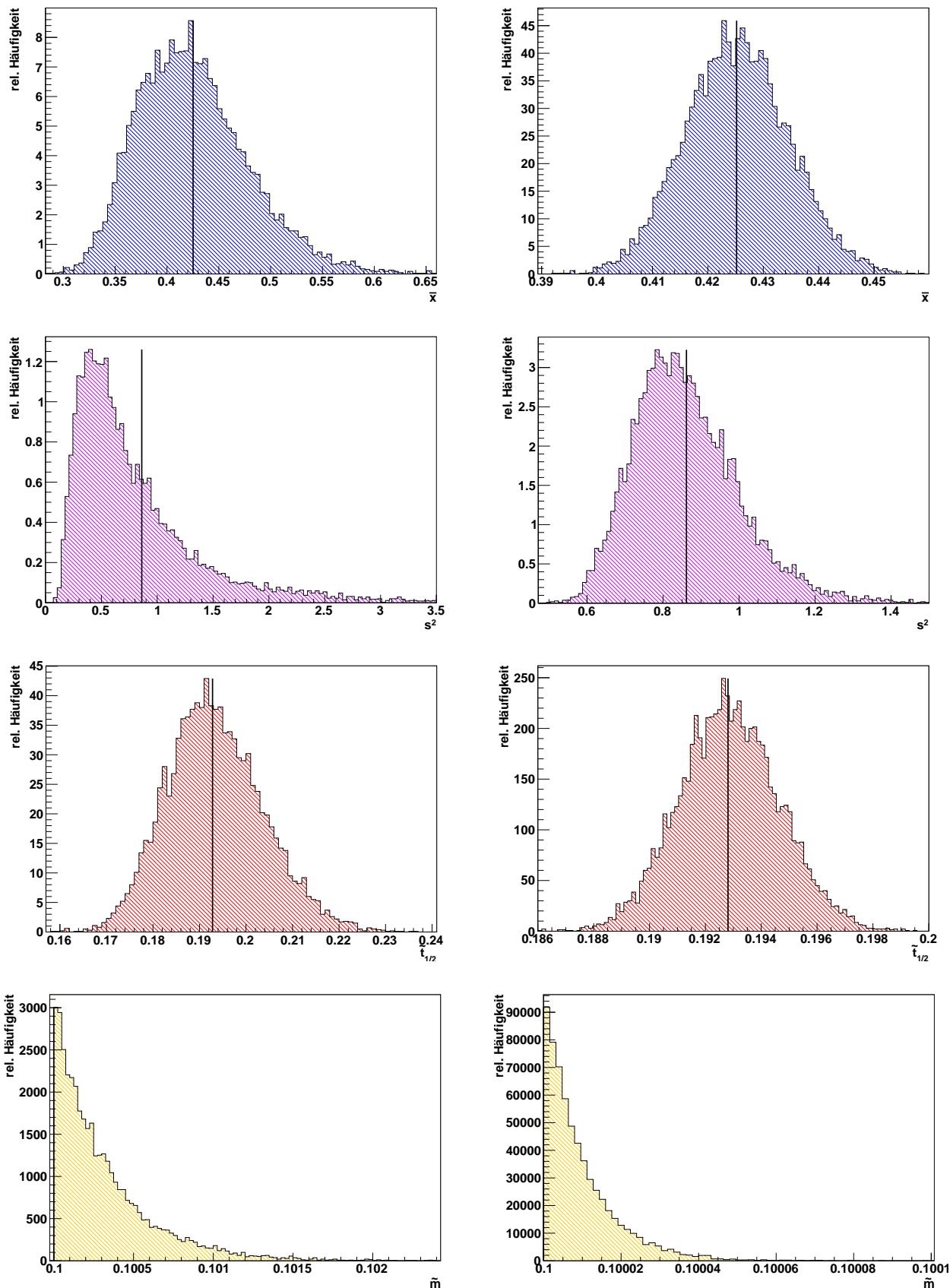


Abbildung 14: Rel. Häufigkeiten des arithm. Mittels (blau), des Stichprobenmedians (rot), der Stichprobenvarianz (violet) und des Stichprobenminimums (orange). Links jeweils eine Stichprobengröße von $n=300$ und rechts $n=10000$. Für das arithmetische Mittel und den Stichprobenmedian zeigt sich jeweils auch für kleine n eine symmetrische Form, die simulierte Verteilung der Varianz zeigt erst für eine Stichprobengröße von $n=10000$ eine Form mit deutlich mehr Symmetrie, die Verteilung des Minimumsschätzers bleibt für alle n asymmetrisch verteilt. Die schwarze Linie zeigt jeweils den vorgegebenen theoretischen Wert der Kenngröße.

Im Folgenden werden die Ergebnisse der Schätzfunktionen $\hat{\alpha}$ und $\hat{\lambda}$ vorgestellt. Dazu wurden für die Werte $\lambda_0 = 0.1$ und $\alpha_0 = 2.0$ für diverse Stichprobengrößen jeweils N=1000 Simulationen durchgeführt um anhand der Ergebnisse wiederum deren Schätzqualität zu beurteilen. Die simulierten Verteilungen der empirischen Kenngrößen $\bar{X}_n, S_{(n)}^2, \tilde{t}_{1/2,n}$ und \tilde{m} sind in Abbildung 14 dargestellt. Die zugehörigen Mittelwerte und Fehler der simulierten Verteilungen sind in Tabelle 7 angegeben. Die Tabelle 8 zeigt die Werte der theoretischen Kenngrößen.

n	$(\bar{x}_n, \sigma_{\bar{x}_n})$	$(\bar{s}_{(n)}^2, \sigma_{s_{(n)}^2})$	$(\tilde{t}_{1/2,n}, \sigma_{\tilde{t}_{1/2,n}})$	$(\tilde{m}, \sigma_{\tilde{m}})$
300	(0.42525 , 0.05349)	(0.85348 , 0.55767)	(0.19339 , 0.01039)	(0.10032 , 0.00032)
1 000	(0.42527 , 0.02941)	(0.85425 , 0.40823)	(0.19296 , 0.00567)	(0.10010 , 0.00010)
3 000	(0.42531 , 0.01698)	(0.86260 , 0.25220)	(0.19288 , 0.00330)	(0.10003 , 0.00003)
10 000	(0.42526 , 0.00943)	(0.86155 , 0.14191)	(0.19285 , 0.00182)	(0.10001 , 0.00001)

Tabelle 7: Mittelwerte der empirischen Kenngrößen und dessen geschätzten Standardabweichungen für N=1000 Simulationen für diverse Stichprobengrößen n. Beim Vergleich mit den theoretischen Kenngrößen, siehe Tabelle 8, nähern sich die Werte weiter an die theoretischen Werte mit wachsendem n an. Die geringen Abweichungen bei \bar{x}_n und bei $\bar{s}_{(n)}^2$ zu ihrem jeweiligen wahren Wert schon bei kleinen Stichprobengrößen bestätigt wiederum, dass es sich hier um erwartungstreue Schätzer handelt. Man sieht hier auch, dass der Schätzer für den Median und der Schätzer für das Minimum auch bei kleinen Stichprobengrößen nur kleine Abweichungen vom wahren Wert aufweisen.

$\mathbb{E}(X)$	$\mathbb{V}(X)$	$t_{1/2}$	m_0
0.42519	0.86173	0.19281	0.1

Tabelle 8: Theoretische Kenngröße für die Parameterwerte $(m_0, \alpha_0, \lambda_0) = (0.1, 2.0, 0.1)$

Da zwar die Form der Wahrscheinlichkeitsdichte, beschrieben durch Potenzgesetz und exponentiellem Cutoff, $x^{-\alpha} \exp(-\lambda x)$, stark rechtsschief ist und ihr Modus, also der Wert mit der größten Wahrscheinlichkeit, sogar bei m liegt, bedeutet das mit anderen Worten, dass Werte nahe m sehr wahrscheinlich sind und selbst innerhalb einer Stichprobe mit nur kleiner Anzahl das Minimum m durch das zugehörige Stichprobenminimum gut geschätzt werden kann. Allerdings bedeutet das auch, dass der Minimumsschätzer $Z := \min(X_1, \dots, X_n)$ nicht erwartungstreue sein kann. Dies lässt sich aber korrigieren, leitet man den Erwartungswert des Minimums her:

$$\begin{aligned} \mathcal{P}(\min(X_1, \dots, X_n) \leq x) &= 1 - \mathcal{P}(\min(X_1, \dots, X_n) > x) = 1 - \prod_{i=1}^n \mathcal{P}(X_i > x) = \\ &= 1 - (1 - F_X(x))^n = 1 - \left(\frac{\Gamma(1 - \alpha, \lambda \cdot x)}{\Gamma(1 - \alpha, \lambda \cdot m)} \right)^n, \quad x \in (m, \infty) \end{aligned}$$

Leitet man diese Verteilung nach x ab, erhält man die zugehörige Dichte der Zufallsvariable Z:

$$f_Z(x) = n \cdot (1 - F_X(x))^{n-1} \cdot f_X(x)$$

Der Erwartungswert des Minimumsschätzers ergibt sich folglich zu:

$$\mathbb{E}(Z) = \int_m^\infty z f_Z(z) dz =: h(n),$$

wobei h eine Funktion ist, die letztlich jeder Stichprobengröße n den Erwartungswert des Minimumsschätzers zuordnet. Allerdings lässt sich hier wiederum kein analytischer Ausdruck finden, andererseits hängt der Erwartungswert von Z weiterhin vom unbekannten Wert von m ab, es gibt somit keinen Korrekturfaktor, der unabhängig von m ist. Daraus folgt, dass ohne das Wissen über m keine Korrektur möglich ist. Im Folgenden lässt sich nun jedoch untersuchen, wie sehr das Schätzergebnis davon beeinflusst wird, wenn man m dazu benutzt, die Korrektur doch auszuführen. Die korrigierte Zufallsvariable $Z' =: m \cdot (1/h(n)) \cdot \min(X_1, \dots, X_n)$ ist also erwartungstreu. Natürlich lässt sich der Erwartungswert von Z auch für eine große Anzahl von Simulationen mit dem sich daraus ergebenden arithmetischen Mittelwert schätzen und sich so rein simulativ die gewünschte Korrektur bestimmen.

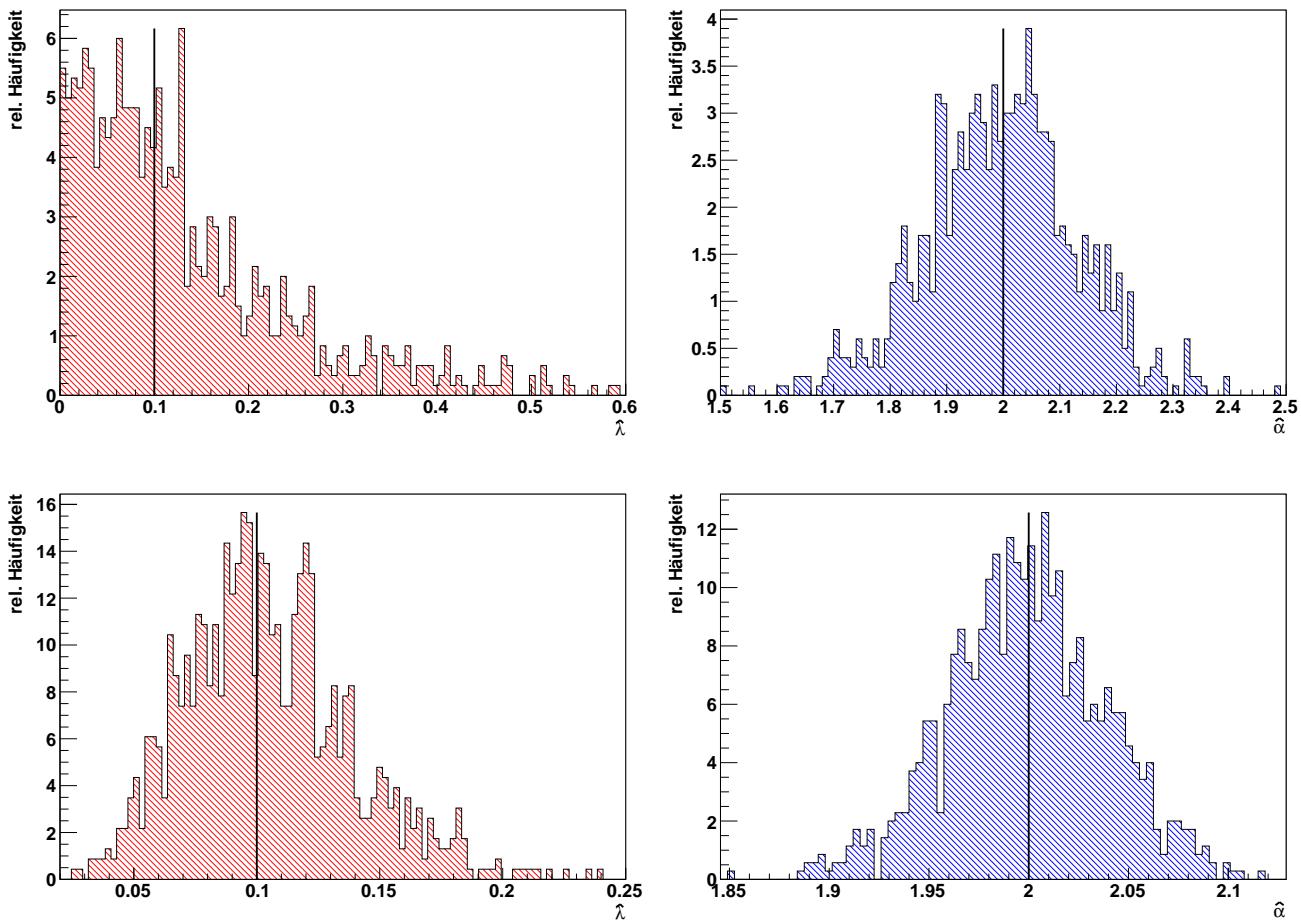


Abbildung 15: Simulierte rel. Häufigkeiten der Schätzfunktionen $\hat{\alpha}$ (blau) und $\hat{\lambda}$ (rot) dargestellt. Links jeweils eine Stichprobengröße von $n = 300$; rechts $n = 10000$. Die simulierte Verteilung von $\hat{\lambda}$ bei $n = 300$ zeigt wiederum eine asymmetrische, rechtsschiefe Form, für $n = 10000$ ist sie jedoch näherungsweise symmetrisch. Die simulierte Verteilung von $\hat{\alpha}$ zeigt hingegen schon für $n = 300$ Symmetrie. Beide Verteilungen weisen bei der größeren Stichprobe n jeweils eine geringere Streuung auf. Die schwarze Linie entspricht jeweils den vorgegebenen theoretischen Wert der Kenngröße.

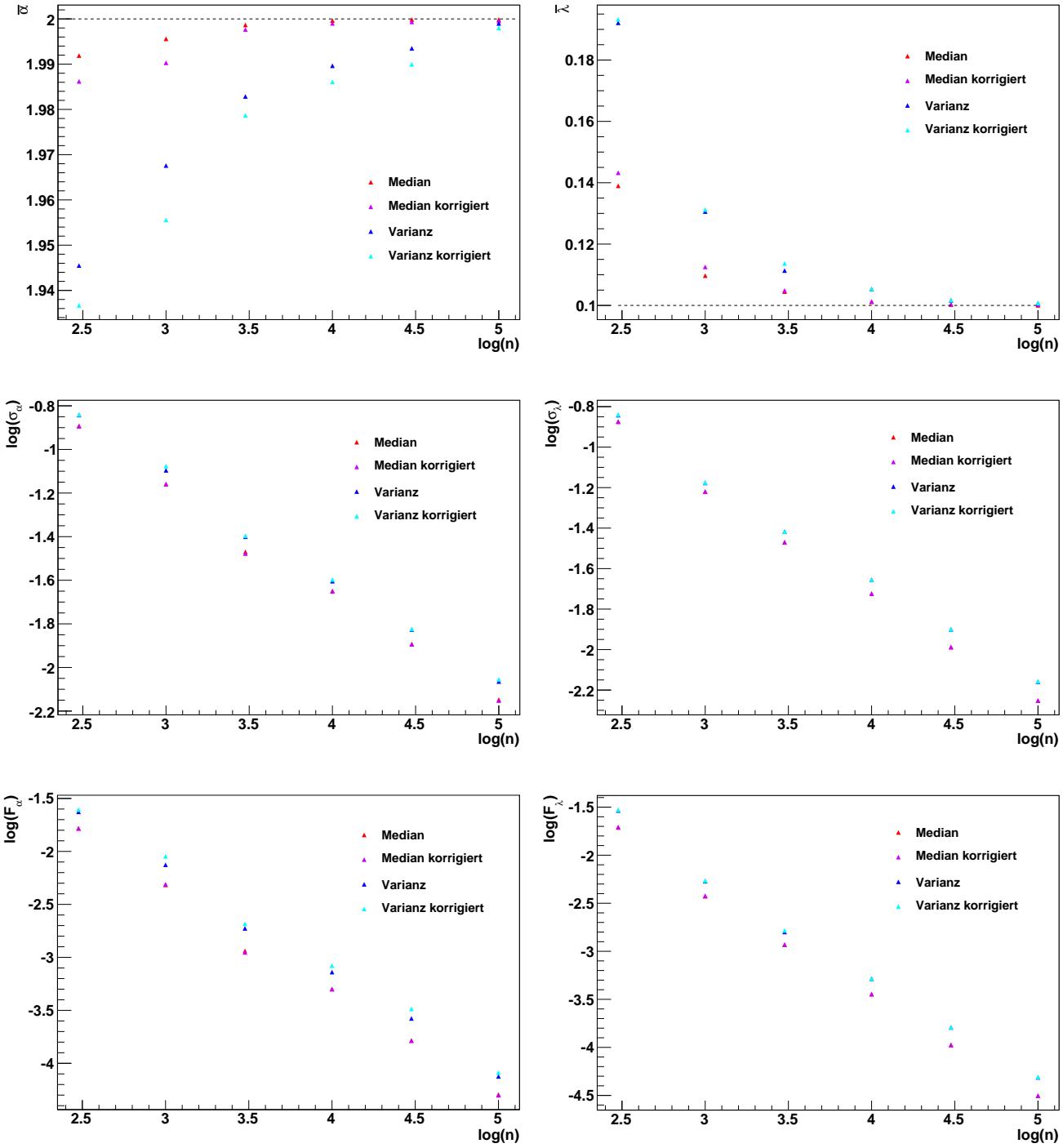


Abbildung 16: Auftragung der Ergebnisse aus den Tabellen 9-10 zur Veranschaulichung des Verhaltens der vier Schätzmethoden, basierend auf einerseits dem Median und der Varianz, andererseits der Korrektur bzw. Nicht-Korrektur des Minimumschätzers. Die oberste Reihe zeigt die Auftragung der Stichprobengröße n gegen den zugehörigen Mittelwert der simulierten Verteilung (links) bzw. gegen die zugehörigen geschätzten Standardabweichungen für $\hat{\alpha}$, die mittlere Reihe zeigt die entsprechenden Ergebnisse für $\hat{\lambda}$. Die untere Reihe zeigt die mittleren quadratischen Fehler F für $\hat{\alpha}$ (links) und für $\hat{\lambda}$ (rechts). Die gestrichelten Linien kennzeichnen die Werte α_0 und λ_0 .

Durch die Korrektur des Minimumschätzers ergeben sich vier verschiedene Möglichkeiten bzw. vier verschiedene Paare, die Erzungsmethode durchzuführen. So werden das arithmetische Mittel mit dem Stichprobenmedian bzw. Varianz kombiniert und sowohl mit dem korrigierten

und dem unkorrigierten Minimumsschätzer untersucht. Ziel ist es nun zu beurteilen, welche der Kombinationen die qualitativ beste Schätzmethode liefert.

Dazu betrachte man die Ergebnisse aus den Tabellen 9-12. Tabelle 9 und 10 listen die Mittelwerte der Schätzfunktionen und Tabelle 11 und 12 listen die zugehörigen geschätzten Standardabweichungen für $\hat{\alpha}$ und $\hat{\lambda}$ auf. Anhand der Tabellenwerte lassen sich für jede der vier Schätzmethoden die gleichen Aussagen, wie im Kapitel zur Beurteilung der Schätzfunktionen im Kontext der erweiterten Maximum-Likelihood-Methode, treffen.

Die Mittelwerte der Schätzfunktionen für beide Parameter nähern sich für wachsende Stichprobengrößen an die wahren Parameterwerte an. Der Schätzer für α nähert sich dabei von unten an α_0 an und der Schätzer für λ nähert sich von größeren Werten an λ_0 an. Mathematisch lässt sich das folgendermaßen festhalten: $\bar{\hat{\lambda}}_n > \bar{\hat{\lambda}}_{n'} > \lambda_0$ und $\bar{\hat{\alpha}}_n < \bar{\hat{\alpha}}_{n'} < \alpha_0$, für verschiedene Stichprobengrößen $n' > n$.

Jede der Schätzfunktionen ist asymptotisch erwartungstreu mit abnehmender Streuung um den jeweiligen Mittelwert bei zunehmender Stichprobengröße. Dies entspricht der Eigenschaft konsistenter Schätzfunktionen.

Dies wird durch die Abbildung 16 bestätigt, die die arithmetischen Mittelwerte der simulierten Verteilungen gegenüber dem Logarithmus der Stichprobengröße aufgetragen zeigt. Man sieht hier wie sich die Mittelwerte der simulierten Schätzfunktionen an die wahren Werte, dargestellt jeweils durch eine gestrichelte Linie, annähern. Die Richtung der Annäherung ist, wie oben angemerkt, gegensätzlich. Man sieht in Abbildung 16 auch, dass sich die geschätzten Standardabweichungen der Schätzfunktionen linear fitten lassen. Dies entspricht einer Abhängigkeit von Standardabweichung und Stichprobengröße, die sich mit einem Potenzgesetz beschreiben lässt. Der Wert der Steigung liegt hier ca. bei -0.5 , was einer Abhängigkeit von $\sigma(n) \propto \frac{1}{\sqrt{n}}$ entspricht. Bei den Mittelwerten liegt eine komplexere Abhängigkeit vor. Auch bei einer doppelt-logarithmisch Auftragung zeichnet sich kein linearer Zusammenhang ab.

Die Abbildung 16 zeigt mit den beiden Auftragungen der untersten Reihe, die den mittleren quadratischen Fehler beider Schätzfunktionen für alle vier Methoden zeigen, eine gute Möglichkeit, die beste Schätzmethode unter den vier präsentierten auszumachen. Grund für die Beurteilung der Schätzer mittels dieser Größe, liegt daran, dass beide Schätzfunktionen für alle vier Methoden nicht erwartungstreu sind. So gilt es nach Kapitel 2.3 den mittleren quadratischen Fehler zu untersuchen.

Betrachtet man zunächst innerhalb dieser Abbildung die Auftragung der Mittelwerte, so lässt sich erkennen, dass die Methode, basierend auf der Stichprobenvarianz, für kleine Stichprobengrößen große Abweichungen vom wahren Wert, im Vergleich zur Methode des Stichprobenmedian, zeigt. Dies gilt für beide Schätzfunktionen $\hat{\alpha}$ und $\hat{\lambda}$.

Im Detail zeigen hier die Methode mit dem Median bei der Schätzung von λ einen relativen Fehler von ca. 40%, die Methode mit der Varianz einen relativen Fehler von ca. 90%. Für große Stichprobenzahlen n , hier $n = 100000$, ergibt sich ein relativer Fehler von ca. 2.6% für die Methode mit dem Median und ca. 7.0% für die Methode, die die Varianz benutzt.

Bei der Schätzung von α für $n = 300$ liegt ein relativer Fehler von ca. 0.4% und für $n = 1000000$ ein relativer Fehler von 0.006% vor, sofern man den Median benutzt, bei der Methode mit der Varianz erhält man einen relativen Fehler von 2.7% für $n = 300$ und für $n = 100000$, erhält man einen relativen Fehler von 0.05%.

Die Unterschiede innerhalb einer Methode, hinsichtlich der Korrektur des Minimumsschätzers, sind im Vergleich zu den Unterschieden in den beiden Methoden, Median und Varianz, nur gering. Dies gilt auch für beide Schätzfunktionen $\hat{\alpha}$ und $\hat{\lambda}$. Die Korrektur hat dabei allenfalls nur bei kleinen Stichprobengrößen Auswirkung, allerdings verschlechtert die Korrektur des Minimumsschätzers den Mittelwert der jeweiligen Schätzfunktionen, in dem Sinne, dass die

Abweichung vom wahren Wert zunimmt. Unterscheidet man zwischen korrigiertem und nicht korrigiertem Minimumsschätzer ergibt sich ein Unterschied für $\hat{\lambda}$ beim Median im relativen Fehler für $n = 300$ von 2.5%, bei $n = 100000$ ein Unterschied im relativen Fehler von 0.005%. Bei der Varianz liegt der Unterschied im relativen Fehler für $n = 300$ bei ca. 1%, für $n=100000$, bei ca. 0.0002%.

Für $\hat{\alpha}$ erhält man einen Unterschied im relativen Fehler beim Median bei $n = 300$ von 0.3%, bei $n = 100000$ einen Unterschied im relativen Fehler von 0.005%. Bei der Varianz liegt der Unterschied im relativen Fehler zunächst bei ca. 1%, bei $n = 100000$ bei ca. 0.0002%.

Untersucht man schließlich noch die geschätzten Standardabweichungen zeigen sich für $\hat{\lambda}$ Unterschiede in den Werten von circa 0.5%, bzw. 0.2% für die Methode basierend auf Median bzw. Varianz, wenn sie untereinander mit dem korrigierten und nicht korrigierten Schätzer verglichen werden. Die Unterschiede zwischen den Methoden, Median und Varianz, liegen bei knapp 7% bei $n = 300$. Was es hier zu bemerken gilt, ist dass die korrigierte Version des Minimumsschätzers eine kleinere Standardabweichung zeigt, für hohe Stichproben lässt sich hier kein statistisch signifikanter Unterschied mehr feststellen.

Insgesamt, hinsichtlich Mittelwerte und Standardabweichungen von $\hat{\alpha}$ und $\hat{\lambda}$, ist die Methode des Stichprobenmedians, die Bessere. Das zeigt sich auch in der untersten Auftragung der Abbildung 16. Dort wird der mittlere quadratische Fehler gegenüber der Stichprobengröße dargestellt, doppelt-logarithmisch aufgetragen. Diese Größe besteht aus dem Quadrat des Bias, also Abstand des Mittelwerts der Schätzfunktion vom wahren Wert, und der zugehörigen Varianz der Verteilung, also die quadrierte Standardabweichung. Aus der Abbildung selbst ist nicht ersichtlich, ob die Korrektur des Minimumsschätzers zu einer besseren Schätzfunktion für beide Parameter führt. Der Unterschied ist am größten bei der kleinsten Stichprobengröße von $n = 300$. Hier ergeben sich Werte des quadratischen Fehlers von 0.019576 bzw. 0.019672 für die unkorrigierte bzw. korrigierte Version des Minimumsschätzers. Dies entspricht, einem relativen Fehler von ca. 0.5%. Insgesamt erweist sich also der unkorrigierte Schätzer, basierend auf dem Median als bester Schätzer, auch wenn der Unterschied nur gering ist. Damit kann man die Problematik, dass man den Minimumsschätzer nicht ohne das Wissen um m korrigieren von keiner Bedeutung, da die Korrektur zu keiner Verbesserung des Schätzverhaltens führt.

Zum Vergleich sei hier noch der mittlere quadratische Fehler bei der Methode der Varianz erwähnt. Man erhält hier einen gemeinsam gerundeten Wert für die unkorrigierte und korrigierte Version, bei einer Stichprobengröße von $n = 300$, von 0.024.

n	$\bar{\hat{\alpha}}_{\tilde{t}_{1/2}}$	$\bar{\hat{\alpha}}_{\tilde{t}_{1/2},korr.}$	$\bar{\hat{\alpha}}_{\mathbb{V}}$	$\bar{\hat{\alpha}}_{\mathbb{V},korr.}$
300	1.99194	1.98625	1.94558	1.93672
1 000	1.99563	1.99031	1.96768	1.95556
3 000	1.99867	1.99767	1.98287	1.97867
10 000	1.99965	1.99902	1.98965	1.98613
30 000	1.99986	1.99934	1.99357	1.98996
100 000	1.99989	1.99955	1.99903	1.99802

Tabelle 9: Zugehörige Fehler der simulierten Verteilungen der Schätzfunktionen $\hat{\alpha}$. Die Indizes $\tilde{t}_{1/2}, \mathbb{V}$ zeigen an, mit welcher Kenngröße das arithmetische Mittel gepaart wurde. Der Zusatz *korr.* bezeichnet, dass der Minimumsschätzer korrigiert wurde.

n	$\bar{\lambda}_{\tilde{t}_{1/2}}$	$\bar{\lambda}_{\tilde{t}_{1/2},korr.}$	$\bar{\lambda}_{\mathbb{V}}$	$\bar{\lambda}_{\mathbb{V},korr.}$
300	0.140897	0.143258	0.192151	0.193249
1 000	0.109739	0.112563	0.130616	0.131244
3 000	0.104495	0.104886	0.111366	0.113695
10 000	0.101169	0.101278	0.105341	0.105361
30 000	0.100319	0.100334	,0.101659	0.101717
100 000	0.100255	0.100265	0.100742	0.100761

Tabelle 10: Zugehörige Fehler der simulierten Verteilungen der Schätzfunktionen $\hat{\lambda}$. Die Indizes $\tilde{t}_{1/2}, \mathbb{V}$ zeigen an, mit welcher Kenngröße das arithmetische Mittel gepaart wurde. Der Zusatz *korr.* bezeichnet, dass der Minimumsschätzer korrigiert wurde.

n	$\sigma_{\hat{\alpha},\tilde{t}_{1/2}}$	$\sigma_{\hat{\alpha},\tilde{t}_{1/2},korr.}$	$\sigma_{\hat{\alpha},\mathbb{V}}$	$\sigma_{\hat{\alpha},\mathbb{V},korr.}$
300	0.128312	0.127625	0.143821	0.144413
1 000	0.069388	0.069314	0.080230	0.083829
3 000	0.033875	0.033375	0.039853	0.040201
10 000	0.022427	0.022353	0.024882	0.025333
30 000	0.012807	0.012776	0.014937	0.015016
100 000	0.007130	0.007054	0.008621	0.008822

Tabelle 11: Zugehörige Fehler der simulierten Verteilungen der Schätzfunktionen $\hat{\alpha}$. Die Indizes $\tilde{t}_{1/2}, \mathbb{V}$ zeigen an, mit welcher Kenngröße das arithmetische Mittel gepaart wurde. Der Zusatz *korr.* bezeichnet, dass der Minimumsschätzer korrigiert wurde.

n	$\sigma_{\hat{\lambda},\tilde{t}_{1/2}}$	$\sigma_{\hat{\lambda},\tilde{t}_{1/2},korr.}$	$\sigma_{\hat{\lambda},\mathbb{V}}$	$\sigma_{\hat{\lambda},\mathbb{V},korr.}$
300	0.133805	0.133408	0.143881	0.144695
1 000	0.060386	0.060212	0.066565	0.066785
3 000	0.033871	0.033927	0.038206	0.038229
10 000	0.018885	0.018897	0.022128	0.022132
30 000	0.010286	0.010287	0.012586	0.012592
100 000	0.005599	0.005601	0.006947	0.006973

Tabelle 12: Zugehörige Fehler der simulierten Verteilungen der Schätzfunktionen $\hat{\lambda}$. Die Indizes $\tilde{t}_{1/2}, \mathbb{V}$ zeigen an mit welcher Kenngröße das arithmetische Mittel gepaart wurde. Der Zusatz *korr.* bezeichnet, dass der Minimumsschätzer korrigiert wurde.

6 Vergleich der erweiterten Likelihood-Methode und der Ersetzungsmethode

In diesem Kapitel werden die beiden vorgestellten Methoden, die erweiterte Maximum-Likelihood-Methode und die Ersetzungsmethode, basierend auf dem arithmetischen Mittel der Stichprobe und dem Stichprobenmedian unter Berücksichtigung eines unkorrigierten Minimumsschätzers diskutiert und die Ergebnisse beider Methoden hinsichtlich der Qualität der zugehörigen Schätzfunktionen verglichen.

Bei der erweiterten Maximum-Likelihood-Methode werden die Schätzwerte der Parameter so an eine Stichprobe angepasst, dass die Wahrscheinlichkeit, diese Stichprobe unter diesen geschätzten Parameterwerten zu erhalten, maximal unter der Bedingung ist, das mittels eines dritten Parameters Φ_0 , der Inhalt der Funktion der beobachteten Stichprobengröße n entspricht.

Hier benutzt man also eine freie Normierungskonstante mit der die zufällige Anzahl der Ereignisse berücksichtigt wird. Der Definitionsbereich der Zufallsvariable X , die gemessenen Energien der Gammastrahlung, ist dabei vorgegeben und entspricht den Werten, die man hier beim H.E.S.S.-Experiment misst.

Bei der Ersetzungsmethode werden der arithmetische Mittelwert und Stichprobenmedian gleich den theoretischen Kenngrößen Erwartungswert und Median gesetzt. Der Wertebereich, auf dem die Zufallsvariable Werte annimmt, wird dabei durch das Minimum der Beobachtung geschätzt. Dies ist also eine allgemeinere und flexiblere Herangehensweise. Die obere Grenze des Wertebereichs ist dabei auf ∞ festgelegt. Da die obere Grenze bei der erweiterten Maximum-Likelihood-Methode den Wert $m' = 10000$ (TeV) hat, hat hier die Funktion des Spektrums schon so stark abgenommen, dass die Beschreibung mit ∞ in sehr guter Näherung die gleichen Ergebnisse bringen würde.

Letzter Unterschied beider Methoden ist, dass bei der Ersetzungsmethode nicht an die Gesamtzahl der Ereignisse, also an die Stichprobengröße n , angepasst wird.

Um die beiden Schätzmethoden anschaulich zu vergleichen werden in Abbildung 17 die arithmetischen Mittelwerte, in Abbildung 18 die Standardabweichungen und in Abbildung 19 die zugehörigen mittleren quadratischen Fehler gegenüber der Stichprobe aufgetragen. Die Größe mittlerer quadratischer Fehler einer Schätzfunktion setzt sich dabei als Summe des quadrierten Bias und der zugehörigen Varianz zusammen.

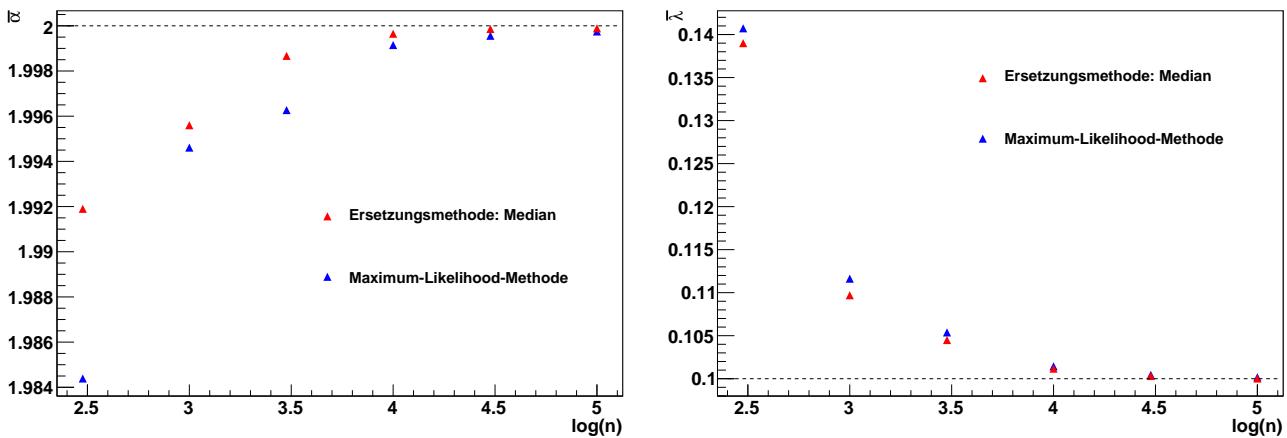


Abbildung 17: Vergleich der erweiterten Maximum-Likelihood-Methode mit der Ersetzungsmethode. Aufgetragen sind die jeweiligen arithmetischen Mittelwerte der beiden Schätzfunktionen gegenüber der jeweiligen Stichprobengröße n . Diese Werte beruhen auf den Werten aus Tabelle 1 und aus den Tabellen 9-12. Man sieht, dass die Ersetzungsmethode über dem ganzen Bereich der Stichprobengröße n bei der Schätzung von λ und α einen kleineren Bias zeigt.

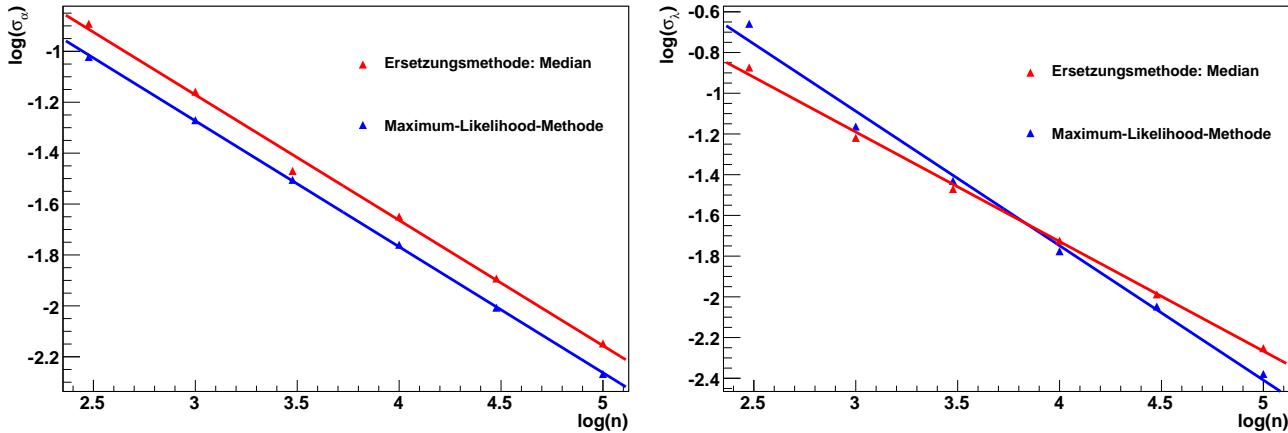


Abbildung 18: Vergleich der erweiterten Maximum-Likelihood-Methode mit der Ersetzungsmethode. Aufgetragen sind die jeweiligen Standardabweichungen gegenüber der jeweiligen Stichprobengröße n . Diese Werte beruhen auf den Werten aus Tabelle 1 und aus den Tabellen 9-12. Man sieht, dass die Ersetzungsmethode über dem ganzen Bereich der Stichprobengröße n bei der Schätzung von α größere Werte der Standardabweichung zeigt. Bei der Schätzung von λ haben die Standardabweichungen der Ersetzungsmethode bis zu einer Stichprobengröße von ca. 5000 kleinere Werte.

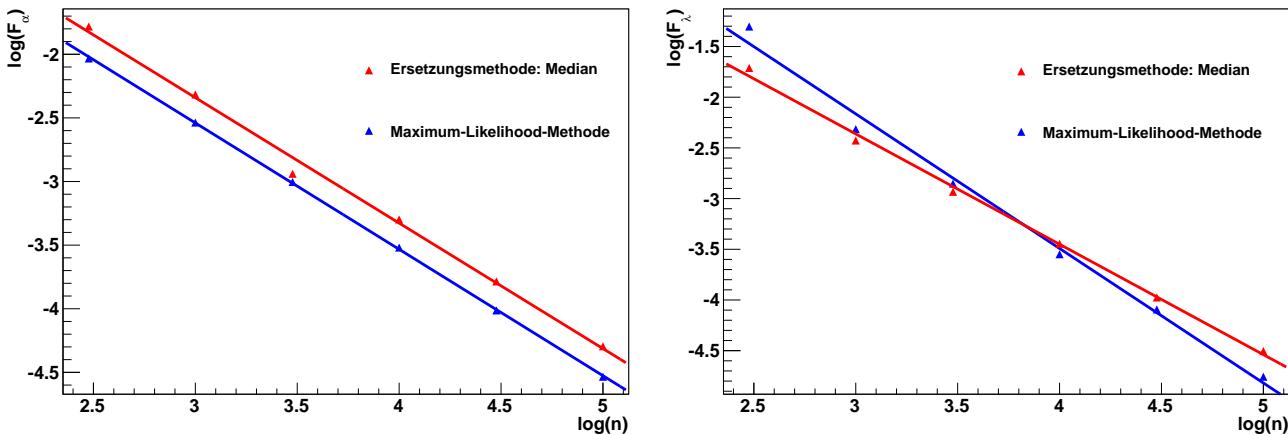


Abbildung 19: Vergleich der erweiterten Maximum-Likelihood-Methode mit der Ersetzungsmethode. Aufgetragen sind die jeweiligen mittleren quadratischen Fehler gegenüber der jeweiligen Stichprobengröße n in doppelt-logarithmischer Auftragung. Diese Werte beruhen auf den Werten aus Tabelle 1 und aus den Tabellen 9-12. Man sieht, dass die Ersetzungsmethode über dem ganzen Bereich der Stichprobengröße n bei der Schätzung von α einen höheren Wert für den mittleren quadratischen Fehler F aufweist. Bei der Schätzung von λ zeigt sich, dass die Ersetzungsmethode bis zu einer Stichprobengröße von ca. 5000 einen kleineren Wert von F aufweist.

Es lässt sich also feststellen, dass die Schätzfunktionen der Ersetzungsmethode für α und λ für alle Stichprobengrößen n einen kleineren Bias aufweisen. Bei den Standardabweichungen und mittleren quadratischen Fehlern von α weist die Ersetzungsmethode für alle Stichprobengrößen n größere Werte auf als die erweiterte Maximum-Likelihood-Methode. Für $\hat{\lambda}$ haben hier die Standardabweichungen und mittleren quadratischen Fehler bei der Ersetzungsmethode kleinere Werte bis zu einer Stichprobengröße von ca. 5000, für größere Stichproben zeigt die erweiterte

Maximum-Likelihood-Methode hier kleinere Werte. Der quadrierte Bias ist bei beiden Schätzfunktionen aber im Vergleich zur Varianz so gering, dass der mittlere quadratische Fehler in guter Näherung gleich der Varianz ist.

Abschließend lässt sich also sagen, möchte man die statistische Genauigkeit vor allem für die Schätzung des Parameters λ bei kleinen Stichprobengrößen erhöhen, empfiehlt sich hier die Ersetzungsmethode. Diese Methode liefert eine Schätzfunktion für λ , die einen kleineren Bias und kleinere Standardabweichung als die der erweiterten Maximum-Likelihood-Methode zeigt. Auch für $\hat{\alpha}$ weist die Ersetzungsmethode einen stets kleineren Bias auf. Für große Stichprobenzahlen hingegen sind die Unterschiede der beiden Methoden vernachlässigbar. Da die Stichprobengrößen im Bereich des H.E.S.S.-Experiments eher im Bereich kleiner Stichprobengrößen von ca. 500-5000 liegen, siehe [Aharonian et al. \(2009\)](#), ist die Ersetzungsmethode eine sehr gute Alternative um den Cutoff-Parameter λ zu schätzen. Im Vergleich zur erweiterten Maximum-Likelihood-Methode angewandt auf gebinnte Daten zeigt die Ersetzungsmethode beim Parameter λ einen stets kleineren Bias, jedoch eine stets größere Standardabweichung.

7 Verbesserung der Schätzfunktion $\hat{\lambda}$

Die Untersuchung der Schätzfunktionen für die Parameter α und λ im Kontext der erweiterten Maximum-Likelihood-Methode, angewandt auf gebinnte Daten, hat ergeben, dass diese Schätzfunktionen konsistent und demnach asymptotisch erwartungstreue sind. Die Frage nach der Effizienz lässt sich allgemein nicht klären, da nicht alle möglichen Schätzverfahren für die beiden Parameter analysiert wurden.

Mit der fehlenden Erwartungstreue der Schätzer geht einher, dass es für kleine Stichproben im Mittel eine Abweichung vom wahren Wert α_0 und λ_0 gibt. Diese Abweichung bezeichnen wir nach Kapitel 2.3 als Bias. Dieser Bias ist von der Stichprobengröße n abhängig und wird für wachsende n kleiner.

Ziel dieses Kapitels ist es die Schätzfunktion für λ so zu korrigieren, dass sie erwartungstreue ist bzw. allgemein einen kleineren Bias aufweist. Dies wird im Folgenden Schritt für Schritt durchgeführt.

Sei zunächst λ_0 bekannt, von α_0 und Φ_0 wird dies nicht gefordert. Simuliert man die Verteilung von $\hat{\lambda}$ für eine beliebige Stichprobengröße n , so erhält man mit $\bar{\hat{\lambda}}$ eine erwartungstreue Schätzfunktion für den Erwartungswert von $\hat{\lambda}$, $\mathbb{E}_n(\hat{\lambda})$. In diesem Fall ließe sich $\hat{\lambda} \rightarrow \hat{\lambda}'$ so korrigieren, dass $\mathbb{E}_n(\hat{\lambda}') = \lambda_0$ ist. Die Schätzfunktionen $\hat{\lambda}$ und $\hat{\lambda}'$ hängen auf folgende Weise zusammen: $\hat{\lambda}' = \frac{\lambda_0}{\mathbb{E}_n(\hat{\lambda})}\hat{\lambda}$.

Da λ_0 unbekannt ist, ist diese Korrektur nicht möglich. Um dieses Problem doch zu lösen, untersuche man zunächst die Abhängigkeit von $\bar{\hat{\lambda}}$ vom wahren Parameter λ_0 .

Dazu werden für $N=1000$ für 25 verschiedene Werte von λ_0 , $\lambda_0^i \in \{0.01, 0.02, 0.03, \dots, 0.25\}$, $i = 1, \dots, 25$, und $\alpha_0 = 2.0$ für verschiedene Stichprobengrößen n die Werte von $\bar{\hat{\lambda}}$ simuliert. Die Abbildung 20 zeigt, dass sich λ_0 und $\bar{\hat{\lambda}}$ approximativ mit einem linearen Zusammenhang beschreiben lassen. Es werden daher für verschiedene Stichprobengrößen, siehe Tabelle 13, Geraden an die Wertepaare $\{(\lambda_0^i, \bar{\hat{\lambda}}^i)\}_{i=1, \dots, 25}$ angepasst. In Tabelle 13 werden die Werte der zugehörigen Fit-Parameter, Steigung c_1 und Achsenabschnitt c_0 der Gerade und reduziertes χ^2 eines Chi-Quadrat-Anpassungstests mit zugehörigem p-Wert angegeben.

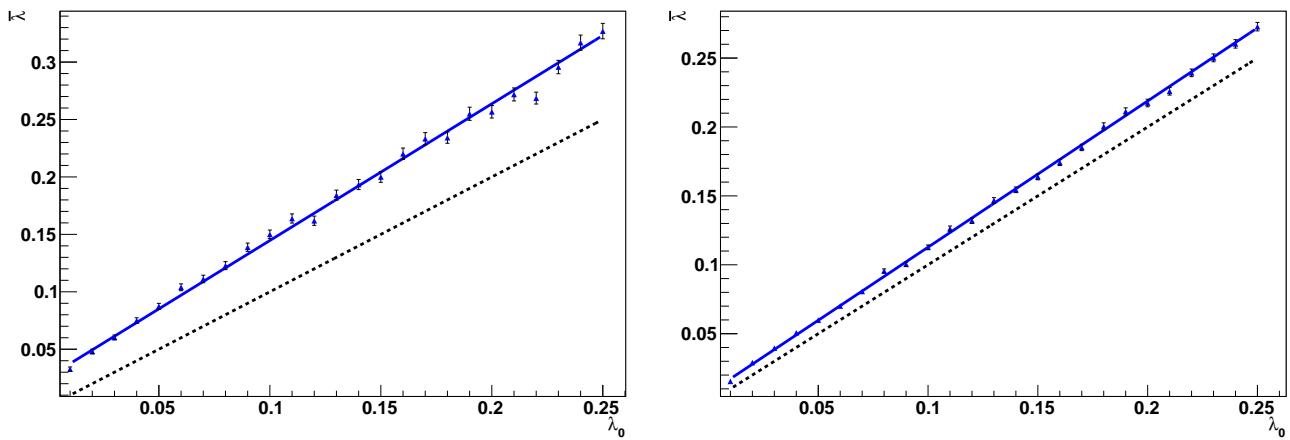


Abbildung 20: Auftragung der geschätzten Werte für λ (blau) in Abhängigkeit des wahren Parameters λ_0 links für $n = 300$ und rechts für $n = 1000$. Die Fehler auf $\hat{\lambda}$, ergeben sich aus den geschätzten Standardabweichungen der simulierten Verteilung von $\hat{\lambda}$ multipliziert mit dem Faktor $\frac{1}{\sqrt{N}}$. Es zeichnet sich ein möglicher linearer Zusammenhang ab, die zugehörige Fit-Gerade ist in blau eingezeichnet. Die gestrichelte Linie stellt die Winkelhalbierende dar.

n	c_0	c_1	p	χ^2/ndf
300	0.0256 ± 0.0011	1.1900 ± 0.012	0.001	47.6/23
1 000	0.0067 ± 0.0004	1.0608 ± 0.0067	0.351	25.0/23
3 000	0.0020 ± 0.0002	1.0243 ± 0.0023	0.973	11.8/23
10 000	0.0063 ± 0.0001	1.0062 ± 0.0015	0.482	22.6/23

Tabelle 13: Auflistung der Fit-Parameter des linearen Modells mit zugehörigen statistischen Größen wie p-Wert, reduziertes χ^2 in Abhängigkeit der Stichprobengröße n. ndf bezeichne die Zahl der Freiheitsgrade.

Aus Tabelle 13 kann man ablesen, dass das lineare Modell ab $n = 1000$ anzunehmen ist, da hier der p-Wert deutlich über dem üblichen gegebenem Signifikanzniveau von 0.05 liegt. Gemessen an den reduzierten χ^2 -Werten, passt mit steigender Stichprobengröße n das Modell immer besser an die Daten. Man erkennt ferner an den Werten für c_1 , die jeweils größer 1 sind, dass ein positiver Bias vorliegt, die geschätzten Mittelwerte $\hat{\lambda}^i$ also stets oberhalb des jeweiligen wahren Wertes λ_0^i liegen. Diese Erkenntnis ist damit eine Verallgemeinerung dessen, was in den Kapiteln 3-5 zum Schätzverhalten von $\hat{\lambda}$, erkannt wurde.

Mit den Werten der Fit-Parameter lässt sich erneut die obige Simulation durchführen, wobei die Schätzfunktion $\hat{\lambda}$ allerdings folgendermaßen angepasst wird: $\hat{\lambda} \rightarrow \hat{\lambda}' = \frac{(\hat{\lambda} - c_0)}{c_1}$. Da $\hat{\lambda}$ und λ_0 über die Fit-Gerade $\hat{\lambda} = c_0 + \lambda_0 c_1$ zusammenhängen, folgt mit dem erwartungstreuen Schätzer $\hat{\lambda}'$ für $\mathbb{E}_n(\hat{\lambda})$ die Erwartungstreue der korrigierten Schätzfunktion $\hat{\lambda}'$ für λ_0 .

$$\mathbb{E}_n(\hat{\lambda}') = \mathbb{E}_n\left(\frac{(\hat{\lambda} - c_0)}{c_1}\right) = \left(\frac{\mathbb{E}_n(\hat{\lambda}) - c_0}{c_1}\right) = \left(\frac{c_0 + \lambda_0 c_1 - c_0}{c_1}\right) = \lambda_0$$

Abbildung 21 zeigt korrigierte und nicht korrigierte Schätzfunktion $\hat{\lambda}$ und $\hat{\lambda}'$. Die Auftragung zeigt anschaulich, dass die Korrektur des Schätzers tatsächlich zu einem erwartungstreuen Schätzer

führt. Man sieht außerdem anhand der Fehlerbalken, dass die geschätzten Standardabweichungen der Verteilung von $\hat{\lambda}'$ kleiner sind als die von $\hat{\lambda}$. Der Grund dafür liegt an der Korrektur. Angenommen die Verteilung der unkorrigierten Schätzfunktion $\hat{\lambda}_n$ hat eine Varianz $V_n(\hat{\lambda})$ für eine bestimmte Stichprobengröße n . So ist die Varianz der korrigierten Schätzfunktion $\hat{\lambda}'$ folgendermaßen gegeben:

$$V_n(\hat{\lambda}') = V_n\left(\frac{(\hat{\lambda} - c_0)}{c_1}\right) = \frac{V_n(\hat{\lambda})}{c_1^2} < V_n(\hat{\lambda})$$

Da der Koeffizient c_1 stets oberhalb von eins liegt, da $\bar{\lambda}$ immer oberhalb von λ_0 liegt, ist die Varianz der korrigierten Schätzfunktion kleiner. Dies gilt entsprechend auch für die geschätzten Standardabweichungen.

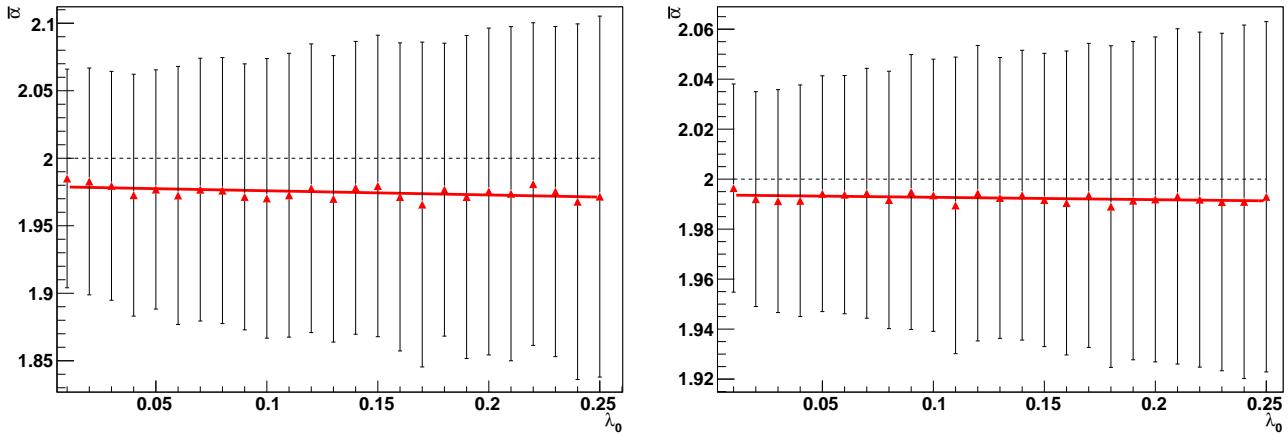


Abbildung 21: Auftragung der Werte von $\hat{\lambda}$ (blau) und $\hat{\lambda}'$ (rot) in Abhängigkeit des wahren Parameters λ_0 , für links für $n = 300$, rechts für $n = 1000$. Die Fehlerbalken beschreiben die geschätzten Standardabweichungen der Schätzfunktionen $\hat{\lambda}$ und $\hat{\lambda}'$. Die jeweils leicht nach rechts gerückten Balken gehören zu $\hat{\lambda}'$. Es ergeben sich mitunter kleinere Fehler. Die gestrichelte Linie stellt die Winkelhalbierende dar.

Ein Problem bei der Korrektur von $\hat{\lambda}$ ist die Bestimmung der Koeffizienten der Fit-Gerade. Um diese zu erhalten, bedarf es der Simulation der Verteilung der $\hat{\lambda}_0^i$. Um diese Simulation durchzuführen, muss man allerdings auch einen Wert für α vorgeben. Damit die Korrektur sinnvoll ist und zum gewünschten Ergebnis führt, muss man das wahre α_0 dafür vorgeben. Dieser Wert ist allerdings wiederum unbekannt.

Weiterhin lässt sich α im Mittel auch nur mit einem Bias schätzen, da $\hat{\alpha}$, wie die Ergebnisse aus Kapitel 5 zeigen, im Mittel unterhalb seines wahren Wertes α_0 liegt.

Wie lässt sich nun trotzdem eine Korrektur von $\hat{\lambda}$ vornehmen?

Um dies zu beantworten, betrachte man zunächst die Verteilung der geschätzten $\hat{\alpha}$ in Abhängigkeit der λ_0^i , bei fixer Stichprobengröße und vorgegebenem Wert α_0 für die Simulation. Die Ergebnisse werden in Abbildung 22 dargestellt. Man sieht, dass die Werte von $\hat{\alpha}$ unterhalb des Wertes α_0 liegen. Dies gilt für alle Werte der vorgegebenen λ_0^i . Interessant ist allerdings, dass die Mittelwerte von $\hat{\alpha}$ dabei offenbar nicht systematisch ansteigen, sondern vielmehr nur leicht um den Mittelwert (gebildet über alle λ_0^i) schwanken. Diese Beobachtung rechtfertigt, eine Konstante an die Werte von $\hat{\alpha}$ zu fitten, um so einen globalen Mittelwert $\bar{\hat{\alpha}}_{glob}$ zu bestimmen. Ferner lässt sich erkennen, dass die zugehörige Standardabweichungen $\hat{\sigma}_\alpha$ dabei nur leicht zu

nehmen. Es lässt sich außerdem aufgrund der nur leichten Zunahme der Standardabweichungen hierfür auch ein globaler Mittelwert bestimmen, der analog als $\bar{\sigma}_{\alpha_{glob.}}$ bezeichnet wird.

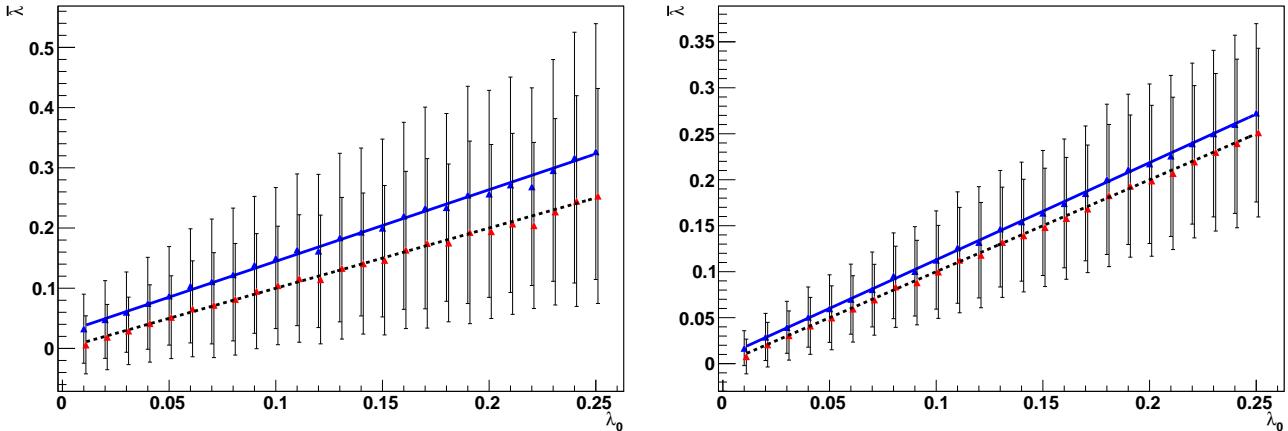


Abbildung 22: Auftragung der geschätzten Mittelwerte von $\hat{\alpha}$ mit zugehörigen geschätzten Standardabweichungen der simulierten Verteilung in Abhängigkeit der λ_0^i für eine Stichprobengröße von $n = 300$ (links) und $n = 1000$ (rechts). Die gestrichelte Linie kennzeichne den wahren Wert α_0 . Die rote Linie stelle die gefittete Konstante an die geschätzten Werte von α dar.

Simuliert man die Verteilung von $\hat{\alpha}$, siehe Kapitel 4, kann man ihr approximativ unterstellen, dass sie einer Normalverteilung folgt. Bei einer normalverteilten Zufallsvariable fallen Erwartungswert und Modus der Verteilung zusammen. Bei der Kenngröße Modus handelt es sich um das Element des Wertebereichs einer Zufallsvariable, an dem die Dichte ihr Maximum annimmt. Aufgrund der Symmetrie sind Werte mit dem gleichen Abstand vom Erwartungswert gleichwahrscheinlich. Aus der vorherigen Überlegung wird im Folgenden untersucht, wie sich die Korrektur von $\hat{\lambda}$ verhält, wenn man für die Korrektur den globalen Mittelwert von $\hat{\alpha}$ aus Abbildung 22 benutzt. Außerdem werden die Werte von $\hat{\alpha}$ betrachtet, die den Abstand einer Standardabweichung haben: $\hat{\alpha}_\pm := \bar{\hat{\alpha}}_{glob.} \pm \bar{\sigma}_{\alpha_{glob.}}$. In dem Intervall $[\hat{\alpha}_-, \hat{\alpha}_+]$ sind aufgrund der Normalverteilungsannahme ca. 68% der Werte von $\hat{\alpha}$ enthalten.

Tabelle 14 listet die gefitteten Geradenparameter für die Stichprobengrößen $n = 300$ und $n = 1000$. Simuliert man mit den Werten $\{\hat{\alpha}_-, \bar{\hat{\alpha}}_{glob.}, \hat{\alpha}_+\}$ und den λ_0^i ergeben sich drei Datensätze, an die sich jeweils eine Fit-Gerade anpassen lässt. Bei $n = 1000$ liegen alle drei reduzierten χ^2 -Werte bei ca. 52/23 und die drei p-Werte bei ca. 0.001, bei $n = 300$ liegen die drei reduzierten χ^2 -Werte bei ca. 35/23 und die drei p-Werte bei ca. 0.05 bzw. 0.07. Damit ist für diese Stichprobengrößen der lineare Fit zwar nicht gerechtfertigt, es lässt sich aber dennoch der systematische Anstieg damit linear nähern.

Für $n = 300$ ergeben sich für die Werte $\{\hat{\alpha}_-, \bar{\hat{\alpha}}_{glob.}, \hat{\alpha}_+\}_{n=300} = \{1.876, 1.976, 2.076\}$ mit geschätzter Standardabweichung von gemittelt ca. 0.1 und für $n = 1000$ ergeben sich $\{\hat{\alpha}_-, \bar{\hat{\alpha}}_{glob.}, \hat{\alpha}_+\}_{n=1000} = \{1.9377, 1.9927, 2.0477\}$ mit geschätzter Standardabweichung von gemittelt ca. 0.05.

Die Werte aus Tabelle 14 zeigen wiederum, dass für die größere Stichprobe die Steigung der gefitteten Gerade abnimmt und der Achsenabschnitt ebenfalls kleiner wird. Man erkennt außerdem, dass die Werte der Steigung mit wachsenden Werten für $\hat{\alpha}$ zunehmen, und wie erwartet liegt der Wert aus Tabelle 14 genau zwischen $\bar{\hat{\alpha}}_{glob.}$ und $\hat{\alpha}_+$.

Mittels der gefitteten Koeffizienten lässt sich die Schätzfunktion $\hat{\lambda}$ in Abhängigkeit der drei Werte von $\hat{\alpha}$ mit $\hat{\lambda} \rightarrow \hat{\lambda}' = \frac{(\hat{\lambda} - c_0)}{c_1}$ korrigieren. Im Vergleich zur obigen Korrektur sind die Koeffizienten so bestimmt, dass berücksichtigt wurde, dass α_0 unbekannt ist. Mit der Korrektur ergeben sich nun für alle drei Werte für α neue Schätzwerte für $\hat{\lambda}'$, die in Abbildung 23 aufgetragen sind. Man sieht, dass durch die Korrektur die geschätzten Werte von $\hat{\lambda}'$ näher um die Winkelhalbierende

	$\hat{\alpha}_-$	$\bar{\hat{\alpha}}_{glob.}$	$\hat{\alpha}_+$	n
c_0	0.0148 ± 0.0007	0.0231 ± 0.0010	0.0345 ± 0.0013	300
c_0	0.0050 ± 0.0004	0.0064 ± 0.0004	0.0082 ± 0.0005	1 000
c_1	1.1701 ± 0.0081	1.1861 ± 0.010	1.2019 ± 0.0123	300
c_1	1.0573 ± 0.0042	1.0607 ± 0.0046	1.0637 ± 0.0051	1 000

Tabelle 14: Auflistung der Parameter der linearen Fits für die Stichprobengrößen n=300 und n=1000 in Abhängigkeit der Werte von $\hat{\alpha}$.

liegen und um diese streuen. Für die Stichprobengröße von $n = 300$ ist der Effekt der Korrektur größer, hier ist die Anpassung mit der Gerade schlechter. Außerdem lässt sich feststellen, dass für $\hat{\lambda}'$ die Werte, die sich aus der Korrektur von $\hat{\alpha}_-$ ergeben, oberhalb der Werte von $\hat{\lambda}'$ durch die Korrektur durch $\hat{\alpha}_+$ sind. Um die Korrektur zu bewerten, wird der relative Fehler f_r für $\hat{\lambda}$ und $f_{r'}$ für $\hat{\lambda}'$ für alle beteiligten Datensätze in Abhängigkeit der Werte von $\hat{\alpha}$ bzw. Tabelle 15 angegeben. Dieser ergibt sich aus der Abweichung des jeweiligen geschätzten Wertes von dem wahren λ_0^i gewichtet mit demselben, gemittelt über die 25 Werte: $f_r = \frac{1}{25} \cdot \sum_{i=1}^{25} \frac{\hat{\lambda}_i - \lambda_0^i}{\lambda_0^i}$ bzw. $f_{r'} = \frac{1}{25} \cdot \sum_{i=1}^{25} \frac{\hat{\lambda}'_i - \lambda_0^i}{\lambda_0^i}$. Das Zeichen „‘“ steht hier dafür, dass die zugehörige korrigierte Schätzfunktion $\hat{\lambda}'$ benutzt wurde, die sich aus der Simulation mit einem der drei Werte für $\hat{\alpha}$ ergibt.

Anhand der Ergebnisse aus Tabelle 15 lässt sich also sagen, dass die Korrektur von $\hat{\lambda}$ mittels der

	$\hat{\alpha}_-$	$\bar{\hat{\alpha}}_{glob.}$	$\hat{\alpha}_+$	n
f_r	0.39	0.53	0.72	300
f'_r	0.15	0.05	0.14	300
f_r	0.13	0.16	0.19	1000
f'_r	0.03	0.02	0.03	1000

Tabelle 15: Auflistung der relativen Fehler der Schätzfunktionen für den Parameter λ für die Stichprobengrößen $n = 300$ und $n = 1000$ in Abhängigkeit der Schätzwerte von $\hat{\alpha}$.

gefitteten Geraden-Parameter c_0 und c_1 ausgehend von Werten aus der Menge $\{\hat{\alpha}_-, \bar{\hat{\alpha}}_{glob.}, \hat{\alpha}_+\}_n$ verringert den relativen Fehler deutlich. Aus Abbildung 23 ist ersichtlich, dass für jedes geschätzte $\hat{\lambda}^{i'}$ gilt näher am Erwartungswert zu sein, als der unkorrigierte Schätzer $\hat{\lambda}^i$. Und dies gilt wiederum für alle $\hat{\alpha}$ innerhalb des 68%-Konfidenzintervalls. Damit wurde gezeigt, dass die Korrektur unter der Berücksichtigung des Unwissens über den wahren Wert von α_0 , zu besseren Ergebnissen im Schätzverhalten des Parameter λ führt.

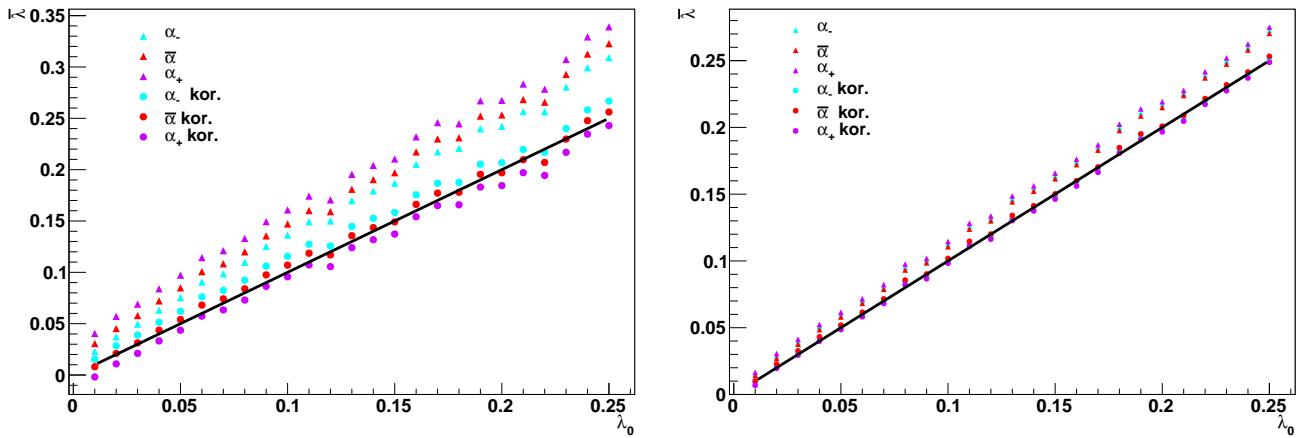


Abbildung 23: Auftragung der geschätzten Mittelwerte von $\hat{\lambda}$ und $\hat{\lambda}'$ in Abhängigkeit der λ_0^i für eine Stichprobengröße von $n = 300$ (links) und $n = 1000$ (rechts). Die eingezeichnete Gerade entspricht der Winkelhalbierenden mit den Wertepaaren $(\lambda_0^i, \hat{\lambda}_0^i)$.

Um nun final zu vergleichen, wie gut die Methode ist, gilt es die korrigierte Version mit der unkorrigierten Version der Schätzung aus Abbildung 18 hinsichtlich ihres mittleren quadratischen Fehlers F zu vergleichen. So wurden jeweils mit $\alpha_0 = 2.0$ und $\lambda_0 = 0.1$ Simulationen durchgeführt und dabei einerseits mit den korrigierten Schätzfunktionen geschätzt. Andererseits kennt man die Ergebnisse oder unkorrigierten Schätzung aus Kapitel 4. So gilt es das Ergebnis der korrigierten Schätzung bei vorgegebenem Wert von λ_0^i zu betrachten. Die Tabelle 16 zeigt dabei für die drei verschiedenen Werte von $\hat{\alpha}$ die Ergebnisse der korrigierten Schätzung bei einem vorgegebenen Wert von $\lambda_0^i = 0.1$. Diese Werte sind auch in der Abbildung 23 enthalten.

Hier die Werte der unkorrigierten Schätzung aus Kapitel 4. Für $n = 300$ hat man $\bar{\hat{\lambda}} = 0.1445$, $\sigma_{\hat{\lambda}} = 0.1216$ und einen zugehörigen mittleren quadratischen Fehler von $F = 0.0150$. Für $n = 1000$ hat man $\bar{\hat{\lambda}} = 0.1116$, $\sigma_{\hat{\lambda}} = 0.0527$ und zugehörigen mittleren quadratischen Abweichungen von $F = 0.0029$. Vergleicht man nun diese Werte, so lässt sich klar erkennen, dass die korrigierte Version jeweils einen kleineren Bias und jeweils eine kleinere Streuung und damit insgesamt einen kleineren mittleren quadratischen Fehler liefert und so als bessere Schätzfunktion anzusehen ist.

n	$\hat{\alpha}_-$	$\bar{\hat{\alpha}}_{glob.}$	$\hat{\alpha}_+$	
300	0.1156	0.1070	0.0957	$\bar{\hat{\lambda}}$
300	0.1001	0.0987	0.0974	$\sigma_{\hat{\lambda}}$
1 000	0.1019	0.1002	0.0983	$\bar{\hat{\lambda}}$
1 000	0.0505	0.0504	0.0502	$\sigma_{\hat{\lambda}}$

Tabelle 16: Auflistung der arithmetischen Mittelwerte und geschätzte Standardabweichungen der simulierten Verteilung von $\hat{\lambda}'$ für Stichprobengrößen $n=300$ und $n=1000$ für die korrigierte Version des Schätzers.

Zum Abschluss noch ein Kommentar zur Betrachtung der Fehler. Alle geschätzten Standardabweichungen, die im Fit benutzt wurden, beruhen auf der Annahme symmetrischer, normalverteilter Fehler. Dies zeigt sich, dass die Fehlerbalken in den Plots in den negativen Bereich reichen. Dies ist vor allem für kleine Werte von λ_0^i und n eine grobe Näherung und bedarf gegebenenfalls einer detaillierten Untersuchung mit asymmetrischen Fehlern beruhend auf Quantilen.

n	$\hat{\alpha}_-$	$\bar{\alpha}_{glob.}$	$\hat{\alpha}_+$	
300	0.0103	0.0098	0.0095	F
1 000	0.0026	0.0025	0.0025	F

Tabelle 17: Auflistung der mittleren quadratischen Abweichungen F der simulierten Verteilung von $\hat{\lambda}'$ für Stichprobengrößen n=300 und n=1000 für die korrigierte Version des Schätzers.

Für einen Großteil der simulierten Werte von λ_0^i ist die Näherung gerechtfertigt. Das ist dann der Fall, wenn die korrigierten Werte für λ inklusive Fehlerbalken im positiven Bereich liegen. Dies gilt also für Werte von $\lambda > 0.05$. Ob die Korrektur anwendbar ist, lässt sich durch die unkorrigierte Schätzfunktion abschätzen und so schließlich $\hat{\lambda}$ korrigieren.

8 Berücksichtigung der effektiven Fläche beim Schätzen der Parameter

Eine weitere Schwierigkeit beim Schätzen der Parameter α und λ ergibt sich aus der *effektiven Fläche*. Mit diesem Begriff wird eine Größe beschrieben, die von mehreren Parametern, wie dem Zenitwinkel der Beobachtung, der Position der Quelle in der Kamera, den verwendeten Analyseschnitten und den Energien der Gammaereignisse abhängt [Aharonian et al., 2006]. Die Größe der effektiven Fläche beschreibt dabei die effektive Nachweisfläche des Instruments. So hat diese Größe auch die Einheit einer Fläche. Multipliziert man diese Größe mit dem Fluss der Gammaquanten, der als Anzahl der Ereignisse pro Zeiteinheit und Flächeneinheit definiert ist, ergibt sich so eine bestimmte Anzahl von Ereignissen pro Zeiteinheit. Im Folgenden wird aber nur ihre Energieabhängigkeit betrachtet, wobei die anderen Größen, von denen die effektive Fläche abhängt, als konstant gelten sollen, bezogen auf eine bestimmte Zeit und eine bestimmte Fläche. Unter diesen Gegebenheiten lässt sich die effektive Fläche mit einer Effizienz vergleichen, mit der Ereignisse abhängig von ihrer Energie registriert bzw. detektiert werden. Diese nimmt einheitlos Werte in $[0, 1]$ an. Dies bedeutet, dass jedem Ereignis eine energieabhängige Wahrscheinlichkeit zukommt, registriert zu werden. Mittels gegebener Monte Carlo Simulationen gibt das Paper Aharonian et al. (2006) Aufschluss darauf, mit welcher mathematischen Funktion sich diese effektive Fläche modellieren lässt. Je größer die Energie eines Ereignisses ist, desto wahrscheinlicher lässt sich das Ereignis registrieren - für hohe Werte der Energie ist die Wahrscheinlichkeit detektiert zu werden 1 - lässt sich dies mit einer streng monoton steigenden Funktion modellieren, die die Limesbedingung berücksichtigt. Beschreibe man die Energiewerte nun wieder mit der Zufallsvariablen X mit konkreten Werten x , eignet sich die Funktion $h(x) = (1 - \exp(-s \cdot x))$, $x \in [m, m']$, $s > 0$ zur Beschreibung. Je größer der Wert des Parameters s , desto mehr verschiebt sich die Sättigung der Wahrscheinlichkeitsdichte, zu kleineren Werten von x , siehe Abbildung 24 rechts. Mit dem Parameter s lässt sich also ansatzweise der Zenitwinkel, siehe Abbildung 24 links, beschreiben. Da der Parameter s positiv ist, gilt so: $h(m') \approx 1$, $h(m) = (1 - \exp(-s \cdot m))$.

Ausgehend von der Funktion des Spektrums $g(x)$ ergibt sich:

$$g^*(x) := \Phi_0 x^{-\alpha} e^{-\lambda x} \cdot (1 - e^{-sx})$$

Die Funktion g^* ist also das Spektrum der Photonenraten, wenn man die effektive Fläche berücksichtigt. Einen guten anschaulichen Vergleich von einem beispielhaften Spektrum g und g^* zeigt die Abbildung 25. Man erkennt hier wie g durch die effektive Fläche, hier im Beispiel von $s=1$, verformt wird und so die Ereignisse für kleine Energien nicht mehr einem Potenzgesetz folgen.

Die hier gewählte Funktion zur Beschreibung der effektiven Fläche ist allerdings keine passende Beschreibung einer effektiven Fläche, sie ist allerdings geeignet um die Schätzmethode unter Berücksichtigung einer effektiven Fläche zu entwickeln. Angenommen, man misst eine Stichprobe vom Umfang n von Ereignissen aus der zugehörigen Verteilung von g^* , wobei der Wert für den Parameter s aus Vergleichssimulationen bekannt ist [Aharonian et al., 2006], so zeigt die Tabelle 18 die zugehörigen Schätzwerte für die Parameter α und λ . Diese sind wiederum durch die arithmetischen Mittelwerte $\bar{\alpha}$ und $\bar{\lambda}$ und die zugehörigen Fehler auf die jeweiligen Mittelwerte $\sigma_{\bar{\alpha}}$ und $\sigma_{\bar{\lambda}}$ dargestellt. Die Fehler auf den Mittelwert einer Schätzfunktion ergeben sich aus dem Fehler der simulierten Verteilung der Schätzfunktion $\sigma_{\hat{\alpha}}$ und $\sigma_{\hat{\lambda}}$ durch die Multiplikation mit dem Faktor $\frac{1}{\sqrt{N}}$.

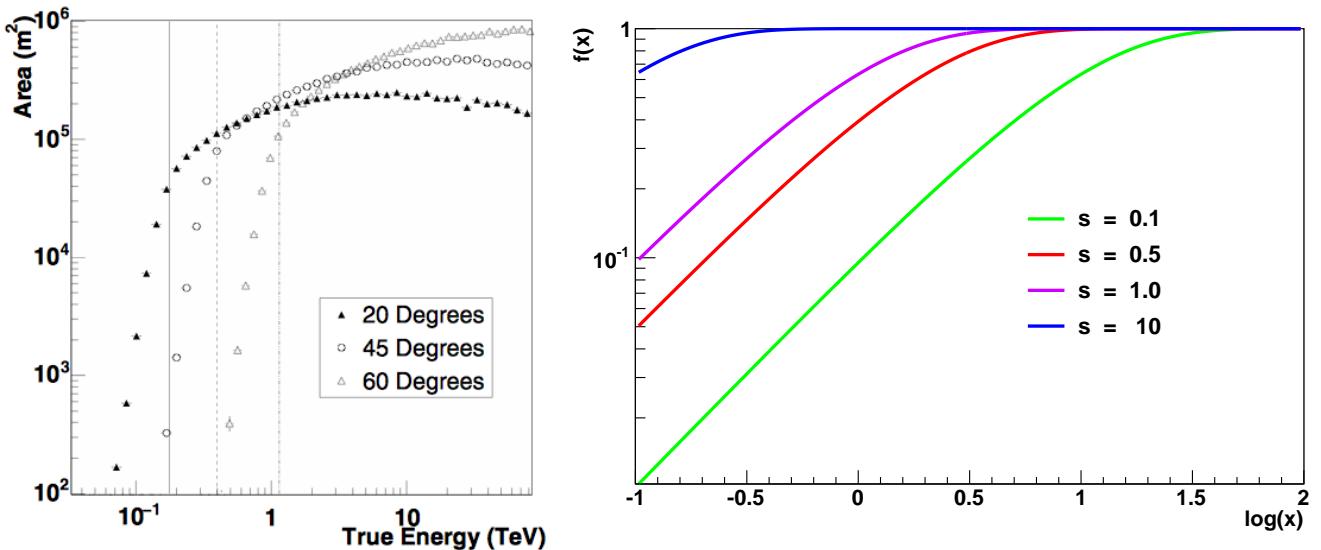


Abbildung 24: Der linke Plot zeigt die simulierten effektiven Flächen des H.E.S.S. für verschiedene Zenitwinkel anhand einiger Datenpunkte [Aharonian et al., 2006]. Der rechte Plot zeigt die Funktion zur Modellierung der effektiven Fläche mittels einiger geeigneter Werte für den Parameter s . Man erkennt, dass die Wahl der Parametrisierung für die effektive Fläche diese nicht treffend beschreibt.

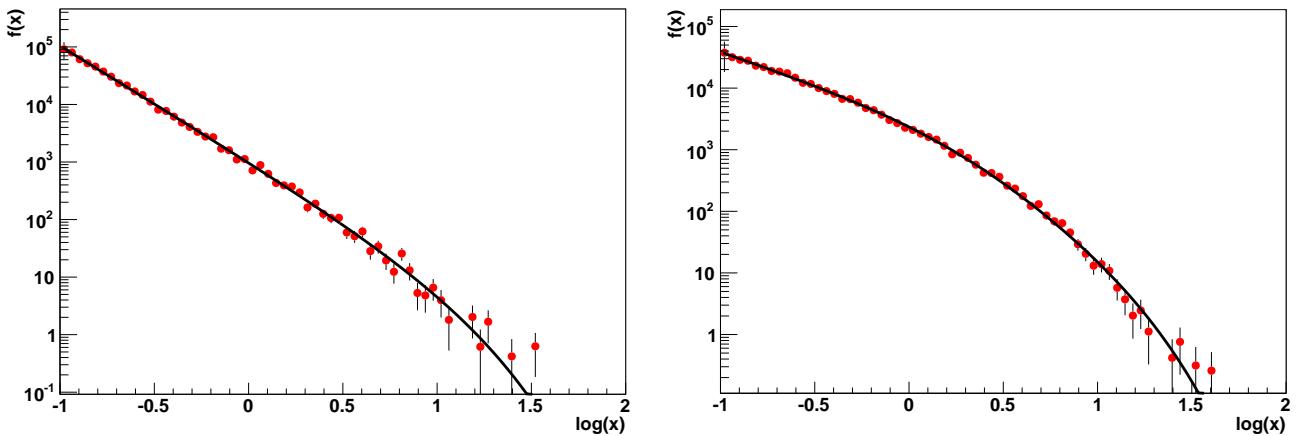


Abbildung 25: Links: Zählspektrum $g(x)$ mit Datenpunkten (rot) und theoretischem Verlauf (schwarz) für eine Stichprobengröße von $n = 10.000$ und Parameterwerten $\alpha_0 = 2.0$ und $\lambda_0 = 0.1$. Rechts: Zählspektrum $g^*(x)$ ebenfalls mit Datenpunkten (rot) und theoretischem Verlauf (schwarz), mit gleichen vorgegebenen Parametern und gleicher Stichprobengröße, aber zusätzlicher Modifikation durch eine beispielhafte effektive Fläche, modelliert durch die Funktion $h(x)$ mit dem Parameterwert $s=1$. Man kann erkennen, dass $g^*(x)$ nun weniger steil abfällt und in der Mitte eine leichte Wölbung zeigt. Für hohe Werte von X zeigt sich keine Veränderung der Form des Spektrums g .

n	$\bar{\alpha} \pm \sigma_{\bar{\alpha}}$	$\bar{\lambda} \pm \sigma_{\bar{\lambda}}$
300	1.9879 ± 0.0026	0.11544 ± 0.00183
1 000	1.9974 ± 0.0014	0.10382 ± 0.00091
3 000	1.9978 ± 0.0008	0.10189 ± 0.00054
10 000	1.9997 ± 0.0005	0.10040 ± 0.00003

Tabelle 18: Auflistung der Mittelwerte der simulierten $\hat{\alpha}$ und $\hat{\lambda}$ und zugehörige Fehler.

Es lässt sich erkennen, dass die Werte für $\hat{\alpha}$ und $\hat{\lambda}$ deutlich näher an den wahren Werte $\alpha_0 = 2.0$ und $\lambda_0 = 0.1$ liegen, verglichen mit den Werten aus Kapitel 4, Tabelle 2. Der Grund dafür, dass bei gleicher Stichprobengröße n bessere Ergebnisse bei der Schätzung der Parameter erzielt werden, liegt daran, dass durch die effektive Fläche, die Wahrscheinlichkeit detektiert zu werden, für Ereignisse mit höheren Werten von X höher ist. Es sammeln sich innerhalb der Stichprobe also anteilmäßig mehr Ereignisse mit höheren Werten von X an, als ohne effektive Fläche bei gleicher gegebener Stichprobengröße. Da nur für hohe Werte von X das Spektrum von der Potenzgesetzhängigkeit merklich abweicht, ist $\hat{\lambda}$ vor allem in diesem Bereich sensitiv. Durch die höhere Anzahl von Ereignissen in diesem Bereich hat man also hierfür eine größere Statistik und das Ergebnis der Schätzung für λ liegt näher am wahren Wert λ_0 , entsprechend den Ergebnissen aus Kapitel 5.

Um die Ergebnisse also sinnvoll miteinander vergleichen zu können, lässt sich für beide Fälle nicht die gleiche Stichprobengrößenanzahl n vorgeben. Um dafür eine geeignete Stichprobengröße n^* zu finden, die man für die Simulation von Zufallszahlen aus $g^*(x)$ vorgibt, kann man sich Folgendes überlegen. Die Modifikation des Spektrums $g(x)$ kommt mathematisch dadurch zustande, dass man zunächst das ursprüngliche Spektrum $g(x)$ betrachtet und nach der zugehörigen Verteilung Zufallszahlen zieht. Als nächstes zieht man eine Bernoulli-verteilte Zufallsvariable Y , entsprechend einem Münzwurf, wobei gilt:

$$Y = \begin{cases} 1, & \text{mit Wahrscheinlichkeit } p(X) = (1 - e^{-s \cdot X}) \\ 0, & \text{mit Wahrscheinlichkeit } q(X) = 1 - p(X) \end{cases}$$

Die Wahrscheinlichkeit ein Ereignis mit einem Wert x zu registrieren, entsprechend $Y = 1$, ist gleich dem Wert $p = (1 - e^{-s \cdot x})$. Simuliert man also wiederum für $N = 1000$ jeweils eine Stichprobe vom Unfang n und untersucht, wie oft in Abhängigkeit der Stichprobe $Y = 1$ ist, lässt sich so ein Wert für n^* angeben. $y_{i,j}$ bezeichne dabei das Ergebnis der Zufallsvariablen Y in der j -ten Simulation beim i -ten Ereignis.

$$n^* = n \cdot \frac{\int g^* dx}{\int g dx}$$

Simuliert man also mit den Werten $\alpha_0 = 2.0$, $\lambda_0 = 0.1$ und $s = 1$ für diverse Stichprobengrößen die Verteilung der Zufallsvariablen Y , lässt sich nach obiger Gleichung jeweils ein Schätzwert für n^* angeben. Tabelle 19 stellt jeweils n und geschätztes n^* gegenüber.

n	n^*
300	126
1 000	417
3 000	1253
10 000	4178

Tabelle 19: Auflistung der Werte der Stichprobengrößen und zugehörige Stichprobengrößen n^* , die die Zahl der n Ereignisse beschreibt, die registriert werden. Der Quotient der beiden Stichprobengrößen liegt jeweils ca. 0.42. Dieser Wert hängt lediglich von s ab. Die Werte von n^* sind dabei gerundete Werte.

Nutzt man diese Werte von n^* und simuliert anhand dieser Werte aus der Verteilung g^* , ergeben sich die Werte, die in Tabelle 20 zusammengestellt werden. Wir verwenden hier analog zur obigen Tabelle 18 die Mittelwerte der Schätzfunktionen und die zugehörigen Fehlern auf diese.

Die untere Tabelle 21 zeigt die analogen Kenngrößen der Schätzfunktionen aus Kapitel 4.

n^*	$\bar{\alpha} \pm \sigma_{\bar{\alpha}}$	$\hat{\lambda} \pm \sigma_{\hat{\lambda}}$
126	1.9652 ± 0.0041	0.13905 ± 0.0032
417	1.9909 ± 0.0021	0.11117 ± 0.0035
1253	1.9970 ± 0.0012	0.10375 ± 0.0033
4178	1.9988 ± 0.0007	0.10111 ± 0.0004

Tabelle 20: Auflistung der Mittelwerte und Fehler der beiden Schätzfunktionen für $g^*(x)$.

n	$\bar{\alpha} \pm \sigma_{\bar{\alpha}}$	$\hat{\lambda} \pm \sigma_{\hat{\lambda}}$
300	1.9785 ± 0.0032	0.14445 ± 0.0038
1 000	1.9947 ± 0.0017	0.11158 ± 0.0016
3 000	1.9963 ± 0.0010	0.10541 ± 0.0010
10 000	1.9992 ± 0.0006	0.10145 ± 0.0005

Tabelle 21: Auflistung der Mittelwerte und Fehler der beiden Schätzfunktionen für $g(x)$.

Daraus lässt sich nun erkennen, dass die Stichprobengröße n^* dazu führt, dass die jeweiligen Mittelwerte in dem Fehlerbereich des jeweiligen anderen liegen und es statistisch dafür spricht, dass beide Methoden zu gleichen Ergebnisse der Schätzung von λ führen. Mit anderen Worten heißt das, dass die Simulation mit einer Stichprobe n^* mit Zufallszahlen verteilt gemäß $g^*(x)$ im Mittel die gleichen Ergebnisse für die Schätzfunktionen bringen wird, wie die Simulation mit n Zufallszahlen verteilt gemäß $g(x)$.

Die Werte für λ aus der Simulation mittels $g^*(x)$ liegen dabei stets leicht unterhalb der Werte von $g(x)$. Hier zeigt sich, dass n^* die Anzahl der Ereignisse ist, die bei einer Anzahl von n Ereignissen detektiert werden. Der Wert n^* ist dabei also die Anzahl der Ereignisse auf dem ganzen Wertebereich von X . Die Form der effektiven Fläche sagt aber nun aus, dass anteilmäßig innerhalb der n^* mehr Ereignisse enthalten sind, die höhere Werte aufweisen, als in der Stichprobe von n unter $g(x)$. Große Abweichung zeigt sich hier bei der Schätzung von α für eine Stichprobe mit $n = 300$ bzgl. $g^*(x)$. Hier macht sich die fehlende Statistik bei kleineren Werten von x^* bemerkbar, sodass der Abstand zu α_0 größer ausfällt.

Interessant zu erwähnen ist, dass die Werte von n^* und n ein konstantes Verhältnis von ca. 0.42 haben, siehe Tabelle 19. Da der Parameter s aus Simulationen bekannt ist, lässt sich aus diesem Wissen auf die tatsächliche Stichprobengrößen n schließen, wenn n^* gemessen wurden, in dem man den Mittelwert von $(1 - \exp(-s \cdot X))$ simuliert der dann für eine hohe Simulationsanzahl gleich dem Quotient von n und n^* ist.

Abschließend lässt sich also sagen, dass die Zahl der tatsächlich detektierten Ereignisse nicht der Verteilung $g(x)$ folgt, also nicht mehr die Form eines Potenzgesetzes hat. Dies könnte zu neuem Schätzverhalten, beispielsweise zu Unterschieden im Bias der beiden Schätzfunktionen $\hat{\alpha}$ und $\hat{\lambda}$ führen. Vergleicht man aber mittels der Stichprobenzahlen n und n^* bezüglich $g(x)$ und $g^*(x)$, so zeigt sich, dass sie das gleiche Schätzverhalten haben.

9 Erweiterte Maximum-Likelihood-Methode unter Berücksichtigung von Untergrundereignissen

Bei der tatsächlichen Analyse von Gammastrahlungsquellen befindet man sich in der Situation, dass dem Signal der Gammaquanten noch ein Untergrundsignal überlagert ist. Verantwortlich für das Untergrundsignal ist die kosmische Strahlung der geladenen Teilchen, die zwar direkt keine Gammaquanten enthält, aber hinsichtlich der Analyse eines Gammaereignisses durch das Messinstrument gelgentlich als Gammaquant fehlidentifiziert werden. Der Grund hierfür liegt in der bodengebundenen Detektion von Gammaereignissen. Diese können nicht direkt gemessen werden, sondern nur über die Wechselwirkung mit der Erdatmosphäre nachgewiesen werden. Es erfolgt also ein Nachweis mittels Sekundärteilchen. Diese Sekundärteilchen können auch durch Nicht-Gamma-Ereignisse entstehen, die unter Umständen auch einem Gamma-Ereigniss zugeordnet werden, sodass sich so letztlich ein Untergrundsignal beimischt.

Im Folgenden soll nur die geladene kosmische Strahlung als Untergrund modelliert werden. Das zugehörige Energiespektrum folgt im Energiebereich der Gammaereignisse einem Potenzgesetz mit spektralem Index von $\Gamma \approx 2.7$. Diese Strahlung selbst setzt sich dabei hauptsächlich aus Protonen und zu einem geringen Teil aus schweren geladenen Kernen zusammen [Grupen \(2000\)](#). Mathematisch lässt sich das Spektrum der kosmischen Strahlung g_k folgendermaßen beschreiben:

$$g_k(x) = \Phi \cdot x^{-\Gamma}, \quad x \in [m, m']$$

Wobei Φ die Flussnormierung ist. Das Spektrum nur wird dabei im gleichen Bereich betrachtet, in auch das Energiespektrum der Gammastrahlung innerhalb dieser Arbeit betrachtet wird. Diee Abbildung 26 zeigt das Spektrum der kosmischen Strahlung im Energiebereich von 1 GeV bis zu ca. 10^5 PeV .

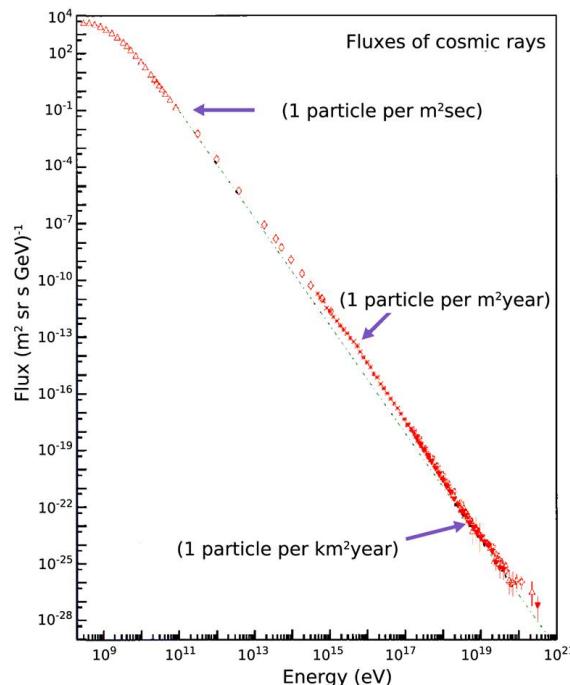


Abbildung 26: [\[Cronin et al., 1997\]](#) Spektrum der kosmischen Strahlung in doppelt-logarithmischer Darstellung. Im Energiebereich von 100 GeV bis 10 PeV erhält man ein Potenzgesetz mit Index von circa 2.7. Bei höheren Energien nimmt der Index dann leicht zu.

Was man also letztlich misst, ist nicht mehr nur das Gammastrahlungssignal, sondern ein Signal, welches sich aus dem Gammasignal und dem Untergrundsignal zusammensetzt. Liegen die Daten wieder gebinnt vor, betrachtet man also die zufällige Anzahl der Ereignisse pro Bin, so ist diese für Quell- und Untergrundsignal, modelliert durch geeignete Zufallsvariablen S_i und B_i jeweils Poisson-verteilt mit verschiedenen zugehörigen Parametern l_i^S und l_i^B . Diese Zufallsvariablen sind einerseits unabhängig voneinander, andererseits misst man letztlich nur die Gesamtanzahl der Ereignisse in einem Bin und kann diese nicht mehr in ihre beiden Anteile trennen. Man erhält also als Beobachtung die Summe der beiden Zufallsvariablen $K_i := S_i + B_i$, die aufgrund der Unabhängigkeit der S_i und B_i Poisson-verteilt mit Parameter $l_i^K = l_i^S + l_i^B$, $i = 1, \dots, K$, ist.

Um nun trotz Untergrundsignal weiterhin die Parameter sinnvoll schätzen zu können, die das Quellsignal beschreiben, beobachtet man auch Regionen, sogenannte Off-Regionen, die nur Untergrundsignal enthalten. Die Zähkstatistik in jeder Region werden durch voneinander unabhängigen Zufallsvariablen $B_{i,j}$ beschrieben, wobei der Index j die verschiedenen Off-Regionen unterscheidet. $j \in \{1, \dots, N_B\}$, wobei $N_B \in \{1, 2, 4, 8\}$ die Anzahl der gewählten Off-Regionen beschreibt. Die Anzahl der Off-Regionen N_B wurde hier so gewählt, um eine entsprechende symmetrische Einteilung des Himmels in Off-Regionen zu verwirklichen, wie in Abbildung 27 dargestellt.

Man nimmt nun vereinfacht an, dass diese Off-Regionen und der Untergrundanteil der Quellregion sich durch das gleiche Spektrum beschreiben lassen, hier das Spektrum der kosmischen Strahlung $g_k(x)$. Außerdem nimmt man an, dass die Zeiten der Beobachtungen für Quellregion und Off-Regionen gleich sind. Man maximiert letztlich wieder die Likelihood-Funktion in Abhängigkeit

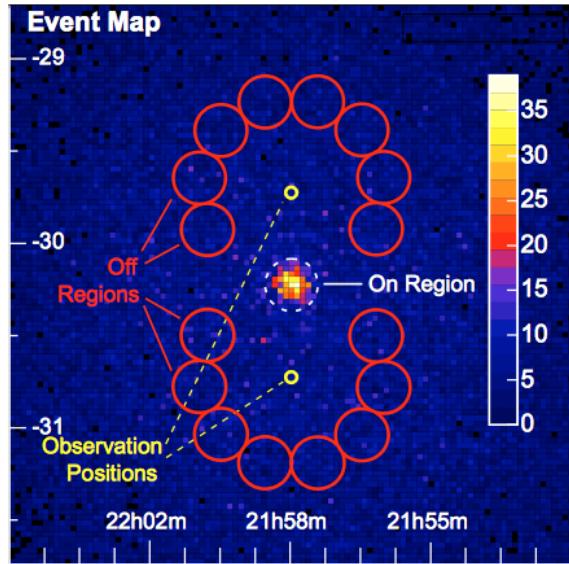


Abbildung 27: Einteilung des Himmels in On- und Off-Regionen [Berge et al., 2007]

der Parameter Φ_0 , α und λ , die die Wahrscheinlichkeit der gemachten Beobachtung beschreibt. Dies ist die gemeinsame Wahrscheinlichkeit $P(K_i = k_i, B_i = b_i) = P(K_i = k_i | B_i = b_i) \cdot P(B_i = b_i)$, also in der Quellregion k_i zu beobachten, wenn man (wie hier im Beispiel von $N_B = 1$) in einer Off-Region b_i Untergrundereignisse registriert. Wegen der Unabhängigkeit der Zufallsvariablen der Quellregion und die der Off-Regionen vereinfacht, sich die bedingte Wahrscheinlichkeit zu: $P(K_i = k_i | B_i = b_i) = P(K_i = k_i)$.

Allgemein erhält man die folgende Likelihood-Funktion:

$$L = \prod_{i=1}^K \left(P_{l_i^S + l_i^B}(K_i = k_i) \cdot \prod_{j=1}^{N_B} \left(P_{l_i^B}(B_{i,j} = b_{i,j}) \right) \right) = \prod_{i=1}^K \left(\frac{e^{-(l_i^S + l_i^B)} (l_i^S + l_i^B)^{k_i}}{k_i!} \prod_{j=1}^{N_B} \left(\frac{e^{-l_i^B} l_i^B^{b_{i,j}}}{b_{i,j}!} \right) \right)$$

Im Folgenden wurden für Stichprobengrößen von $n \in \{300, 1000, 3000, 10000, 30000\}$ jeweils $N = 1000$ Simulationen durchgeführt. Die Stichprobengröße bezieht sich dabei auf die Anzahl der reinen Gammaereignisse der Quellregion.

Die Anzahl der Untergrundereignisse n_b für die Quellregion und für jede der Off-Regionen wurde damit von 10% bis zu 1000% der jeweiligen Stichprobengröße n mit jeweils 100 gleich großen Zwischenstufen gewählt. Als Parameter zur Simulation wurden weiterhin $\lambda_0 = 0.1$ und $\alpha_0 = 2.0$ gewählt. Die Flussnormierung Φ_0 ergibt sich wiederum aus der jeweiligen Stichprobengröße n . Für die gleichzeitige Simulation der Untergrundereignisse wurde als Wert $\Gamma = 2.7$ vorgegeben. Die Flussnormierung ergibt sich hier ebenfalls aus dem vorgegebenen Parameter n_b . Gemäß der Likelihood-Funktion, siehe oben, wurden mittels der Parameter Φ_0 , α , λ und Φ angepasst. Der Parameter Γ ist im Fit ein fixierter Parameter, den man aufgrund des Wissens des spektralen Index der kosmischen Strahlung als bekannt ansehen kann.

Abbildung 28 zeigt die Ergebnisse der geschätzten Werte für λ_{n_b} in Abhängigkeit der Anzahl der Untergrundereignisse n_b am Beispiel von einer Stichprobengröße $n = 1000$. Die geschätzten Werte für λ_{n_b} sind hier mit ihrem Mittelwert als Punkt dargestellt, die Fehlerbalken zeigen die geschätzten Standardabweichungen $\sigma_{\lambda_{n_b}}$ auf die Verteilung der λ_{n_b} , gewichtet mit dem Kehrwert der Anzahl der Simulationen.

- Für jede Anzahl der Off-Regionen N_B lässt sich zunächst feststellen, dass die Werte für die Standardabweichungen $\sigma_{\lambda_{n_b}}$ bzw. $\sigma_{\bar{\lambda}_{n_b}}$ mit zunehmender Untergrundereignisanzahl n_b zunehmen. Die minimalen Werte von $\sigma_{\lambda_{n_b}} = 0.053$ decken sich mit den Werten, die man bei Abwesenheit von Untergrund erhält, und sind dabei unabhängig von der Zahl der Off-Regionen gleich.

Die maximalen Werte von $\sigma_{\lambda_{n_b}}$ sind abhängig von der Zahl der Off-Regionen in Tabelle 22 dargestellt.

N_B	$\sigma_{\lambda_{n_b},max}$
1	0.1599
2	0.1314
4	0.1193
8	0.1057

Tabelle 22: Auflistung der maximalen Werte der Standardabweichungen der simulierten Verteilungen von $\hat{\lambda}$ für jede der Anzahl der Off-Regionen.

So zeigt sich, dass die Berücksichtigung von mehreren Off-Regionen zu kleineren Werten von $\sigma_{\lambda_{n_b}}$ führt und die geschätzten Werte für λ bei einem fixen n_b weniger streuen, je mehr Off-Regionen man berücksichtigt.

- Man kann außerdem feststellen, dass die Spannbreite, in dem sich die Werte für $\bar{\lambda}_{n_b}$ bewegen, für eine steigende Anzahl von Off-Regionen kleiner wird. Des Weiteren bewegt sich dieser Bereich gleichzeitig immer näher an den wahren Wert von $\lambda_0 = 0.1$.

Hierzu seien die minimalen und maximalen Werte von $\bar{\lambda}_{n_b}$ für alle Werte von n_b in Tabelle 23 aufgetragen.

N_B	$\bar{\lambda}_{n_b,min}$	$\bar{\lambda}_{n_b,max}$
1	0.11153	0.14594
2	0.11058	0.13686
4	0.10739	0.13054
8	0.09716	0.12148

Tabelle 23: Auflistung der maximalen und minimalen Werte der arithmetischen Mittelwerte der simulierten Verteilungen von $\hat{\lambda}$ für jede der Anzahl der Off-Regionen.

- Für jeden der vier Plots aus Abbildung 28 lässt sich die x-Achse in zwei Bereiche einteilen. Für kleinere Zahlen der Unterrgrundereignisse n_b steigen die Werte der geschätzten Werte für λ_{n_b} bis zu einem gewissen Punkt an. Für höhere Werte von n_b nehmen die geschätzten Werte λ_{n_b} hingegen ab. Hier lassen sich jeweils Geraden $\lambda_{n_b} = c_1 \cdot n_b + c_0$ an die Werte fitten, die Bereiche die für eine positive Steigung geeignet sind, werden mit zunehmender Anzahl der Off-Regionen kleiner.

Die sich ergebende Werte für den Achsenabschnitt und Steigung der Fit-Geraden c_0 , c_1 , jeweilige p-Werte und reduzierte χ^2/ndf von χ^2 -Anpassungstests, sind in Tabelle 24 dargestellt.

N_B	c_0	c_1	p	χ^2/ndf	Fit-Bereich
1	0.11480 ± 0.0004821	$(2.9812 \pm 0.131) \cdot 10^{-6}$	0.14	94.01/80	[0, 8100]
2	0.11401 ± 0.0005005	$(2.7183 \pm 0.151) \cdot 10^{-6}$	0.12	83.31/69	[0, 7000]
4	0.11445 ± 0.0005391	$(2.6029 \pm 0.196) \cdot 10^{-6}$	0.06	71.97/55	[0, 5500]
8	0.11477 ± 0.0005734	$(1.7371 \pm 0.250) \cdot 10^{-6}$	0.05	61.45/45	[0, 4600]

Tabelle 24: Auflistung der Fit-Parameter des linearen Modells mit positiver Steigung mit zugehörigen statistischen Größen wie p-Wert, reduziertes χ^2 und dem Fit-Bereich auf der n_b -Achse.

N_B	c_0	c_1	p	χ^2/ndf	Fit-Bereich
1	0.1356 ± 0.0148	$(-6.957 \pm 1.634) \cdot 10^{-7}$	0.51	17.16/18	[8100, 10000]
2	0.1617 ± 0.0063	$(-4.882 \pm 0.075) \cdot 10^{-6}$	0.09	39.41/29	[7000, 10000]
4	0.1295 ± 0.0029	$(-1.868 \pm 0.381) \cdot 10^{-6}$	0.53	43.64/45	[5500, 10000]
8	0.1328 ± 0.0019	$(-3.190 \pm 0.264) \cdot 10^{-6}$	0.72	47.67/54	[4600, 10000]

Tabelle 25: Auflistung der Fit-Parameter des linearen Modells mit negativer Steigung mit zugehörigen statistischen Größen wie p-Wert, reduziertes χ^2 und dem Fit-Bereich auf der n_b -Achse.

Die gefitteten Achsenabschnitte im linken Bereich der n_b -Achse liegen alle im Fehlerbereich des jeweils anderen, anhand dieses Ergebnisses ist der Koeffizient c_0 also für N_B bis auf statistische Schwankungen gleich. Die zugehörigen Steigungen des Geradenfits nehmen dabei mit steigender Zahl N_B ab, ebenso wie die Streuung der einzelnen Werte für λ_{n_b} um

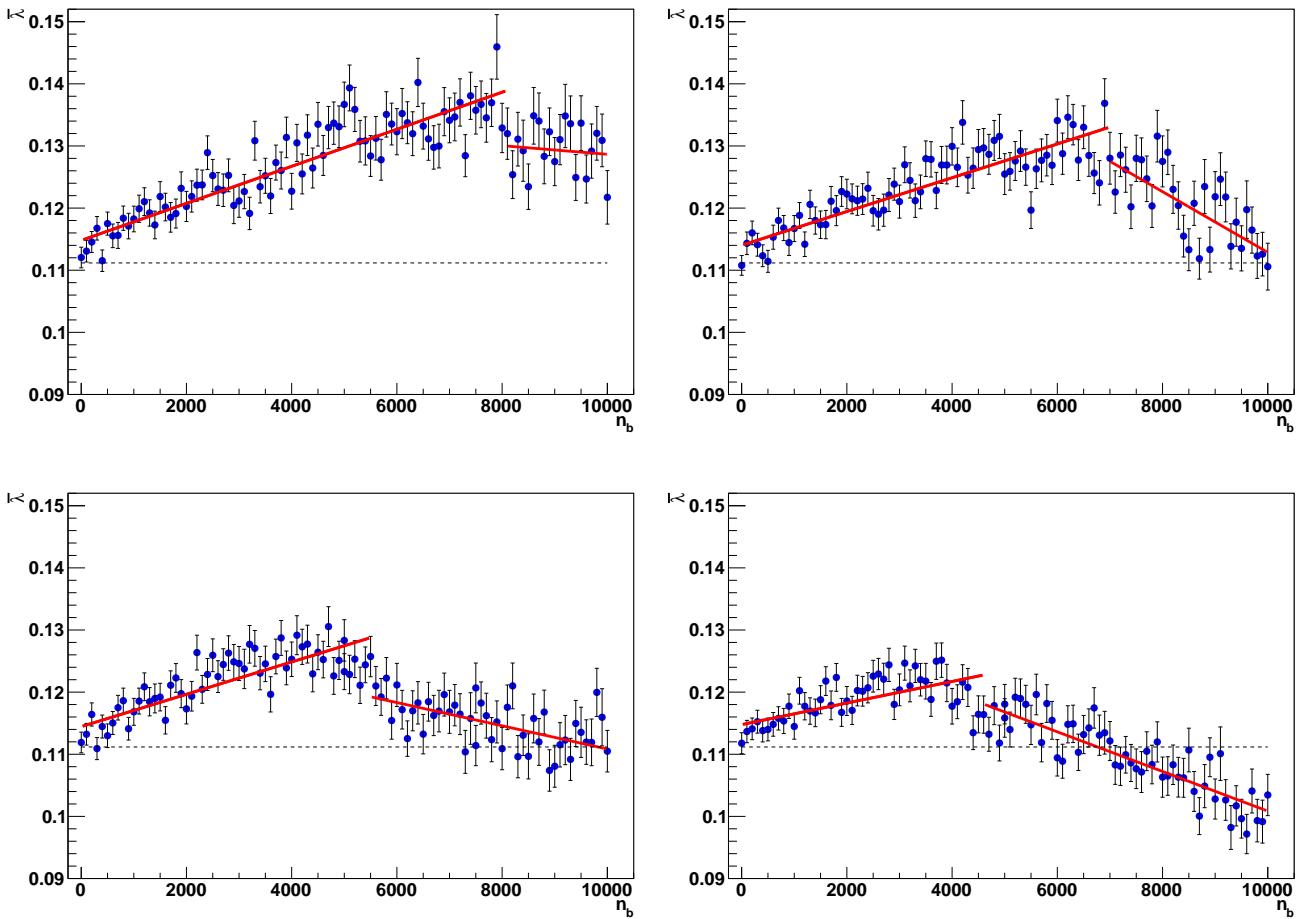


Abbildung 28: Die Plots zeigen die geschätzten Werte für $\hat{\lambda}$ mit zugehörigem Fehler in Abhängigkeit der simulierten Anzahl der Untergrundereignisse von n_b . Der Plot links oben zeigt die Auswertung für $N_B = 1$, der rechts daneben die Auswertung für $N_B = 2$. In der Zeile darunter folgen die jeweiligen Auswertungen für $N_B = 4$ und $N_B = 8$. Die gefitteten Geraden sind jeweils in rot dargestellt.

die Gerade. Über den Fit im zweiten Bereich für hohe n_b lässt sich dabei nur erkennen, dass bei nur einer Off-Region im Vergleich zu den anderen Werten bei mehreren Off-Regionen, die Werte von $\hat{\lambda}$ nahezu konstant bleiben und sie in den anderen Fällen jedoch linear abnehmen, die Streuung um diese Werte wird dabei mit steigender Zahl der Off-Regionen kleiner.

Mögliche Gründe für das Ansteigen der Werte des geschätzten Cutoff-Parameters und dem Absinken ab einem bestimmten Wert für n_b , kann man sich folgendermaßen qualitativ am Beispiel der Abbildung 29 plausibel machen.

Wichtig für das Schätzen von λ sind die Bins mit hohen Werten für X , da die Wahrscheinlichkeitsdichte für kleinere Werte in guter Näherung nur durch den Potenzgesetz-Anteil beschrieben werden. Nur für hohe Werte von X ist λ daher sensitiv.

Vergleicht man die spektralen Funktionen $g(x) = \Phi_0 x^{-\alpha} e^{-\lambda x}$ des Gammasignals und die des Untergrunds $g_k(x) = \Phi x^{-\Gamma}$, so dominiert die Kurve des Untergrunds für hohe Werte von X gegenüber der des Signals. Mit Dominanz ist hierbei gemeint, dass die Zahl zu erwartender Ereignisse in dem Bereich für den Untergrund maßgeblich höher ist.

Anfangen bei nur einem Untergrundereignis dominiert die Kurve des Gammasignals noch über dem ganzen Wertebereich von X , in dem man tatsächlich auch mit nicht verschwindender Wahrscheinlichkeit Ereignisse erwarten kann.

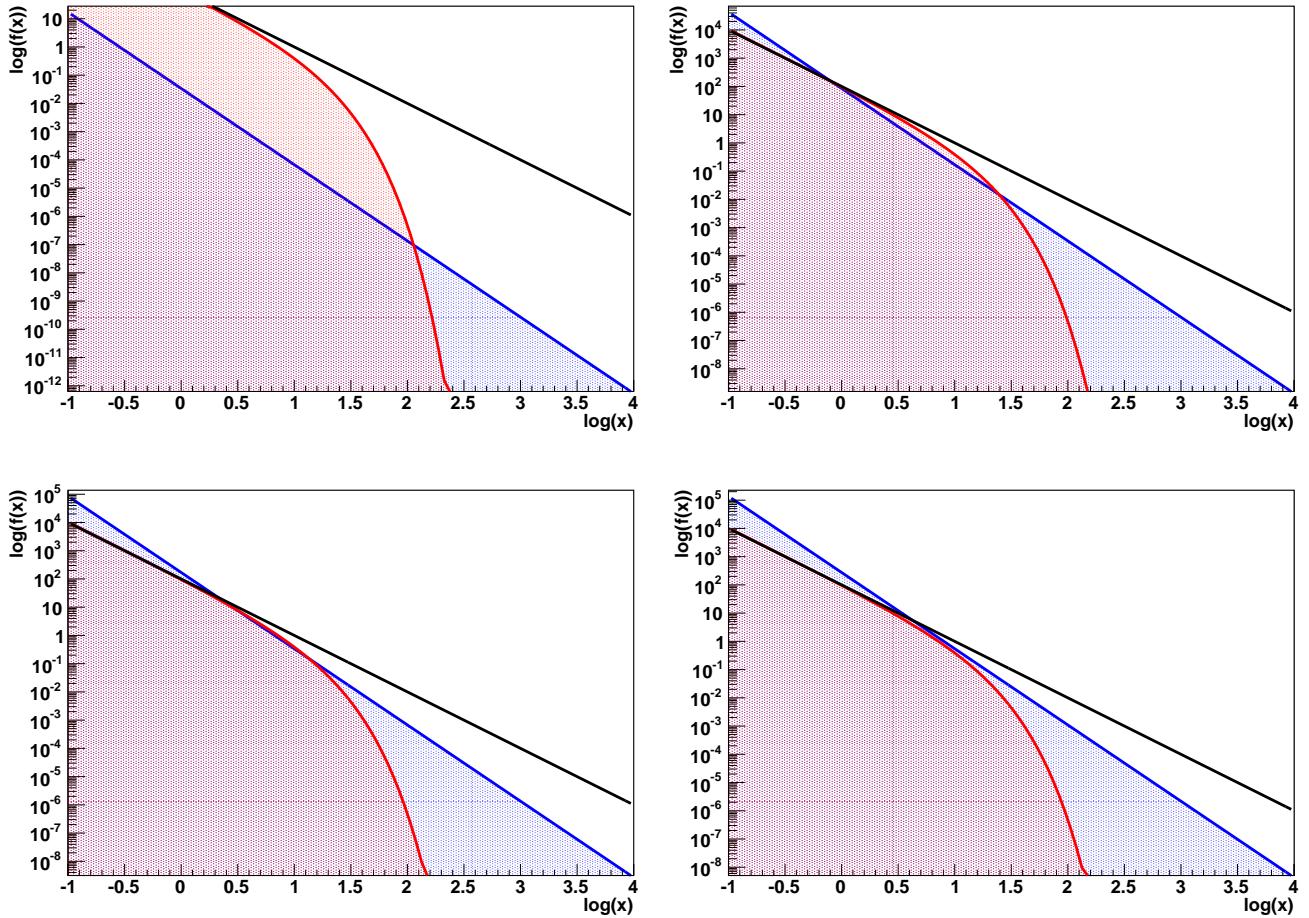


Abbildung 29: Die 4 Plots zeigen die Spektren $g(x) = \Phi_0 x^\alpha e^{-\lambda x}$ (rot) und $g_k(x) = \Phi x^{-\Gamma}$ (blau) für verschiedene n_b bei einer konstanten Stichprobengröße $n = 1000$ in doppelt-logarithmischer Darstellung. n_b nimmt hier die Werte 1, 2500, 5000 und 8100 an. Diese Werte wurden so gewählt, um die Entwicklung der beiden Spektren zueinander gut zu veranschaulichen. Die Reihenfolge entspricht hier von der oberen Zeile von links nach rechts und in der unteren Zeile analog. Das Untergrundsspektrum zeigt sich hier als Gerade mit Steigung von -2.7 . Die schwarze Gerade entspricht dem Potenzgesetzanteil des Signalspektrums mit einer Steigung von -2.0 . Der ausgefüllte Flächeninhalt der beiden Kurven entspricht damit der Anzahl der erwarteten Ereignisse auf dem gesamten Wertebereichs von X .

Erhöht man die Anzahl der Untergrundereignisse, wandert der Bereich, wo das Signal den Untergrund überwiegt, nach links zu kleineren Werten von X . Für ansteigende Werte der Anzahl der Untergrundereignisse bleibt nur noch ein immer kleiner werdender Bereich, in dem das Signal bei mittleren Werten dominiert, bis es schließlich ab circa $n_b = 8100$ im gesamten Bereich nicht mehr signifikant beiträgt.

Dominiert der Untergrund das Gammasignal, kann man also in guter Näherung das Signal vernachlässigen, dann ist die Schätzung von λ abhängig von der Schwankung der Untergrundsignale in den Off-Regionen im Vergleich zum Untergrund in der Signalregion. So kann es auch dazu kommen, dass in einer Off-Region mehr Ereignisse in einem Bin registriert werden, als in der Quellregion insgesamt in diesem Bin. So führt es trotzdem dazu, dass der Mittelwert für ein Bin für das Quellsignal kleiner geschätzt wird, was in einem höheren Schätzwert für λ resultiert. Je größer die Zahl der Untergrundereignisse insgesamt ist, desto mehr Bins wird es auch geben, in denen der Untergrund zahlenmäßig dominiert, sodass $\hat{\lambda}$ immer größer ausfällt. Erreicht man Werte für n_b , sodass der Untergrund dem

Signal in jedem Bin überwiegt, liegt die Schätzung von λ letztlich an den Unterschieden des simulierten Untergrunds in Off und Quellregion. Diese Schwankungen nehmen mit n_b zu, da pro Bin so mehr Ereignisse erwartet werden, und es damit poisson-gemäß auch eine höhere Varianz ergibt. Dies erklärt warum für eine Off-Region ab einem Wert von $n_b = 8100$ die Mittelwerte der geschätzten Werte für λ stark schwanken und keine Systematik, also zum Beispiel für kleinere Werte von n_b ein systematischer linearer Anstieg, mehr erkennbar ist.

Die Tabellen 26-28 beinhalten die entsprechenden Ergebnisse für die anderen Stichprobengrößen n . Dabei wurde der Fit im Bereich mit großen Werten von X wegen der fehlenden Interpretationsmöglichkeiten nicht berücksichtigt. Die p-Werte und χ^2 -Werte des Fits sind auch nicht aufgeführt. Die Idee der Fits war es, den Fit-Bereich für n_b wiederum so zu wählen, dass das lineare Modell an die Daten passt, und den Fit-Bereich dabei möglichst groß zu wählen. Das macht die Angabe der statistischen Größen überflüssig.

N_B	c_0	c_1	Fit-Bereich
1	0.14763 ± 0.00136	$(5.142 \pm 0.174) \cdot 10^{-5}$	[0, 1930]
2	0.14738 ± 0.00161	$(4.713 \pm 0.323) \cdot 10^{-5}$	[0, 1100]
4	0.14971 ± 0.00123	$(3.637 \pm 0.156) \cdot 10^{-5}$	[0, 1950]
8	0.14928 ± 0.00149	$(3.501 \pm 0.232) \cdot 10^{-5}$	[0, 1230]

Tabelle 26: Auflistung der Fit-Parameter im linearen Modell und Fit-Bereich bei einer Stichprobengröße von $n=300$.

N_B	c_0	c_1	Fit-Bereich
1	0.10422 ± 0.00029	$(4.015 \pm 0.357) \cdot 10^{-7}$	[0, 16600]
2	0.10409 ± 0.00031	$(3.321 \pm 0.443) \cdot 10^{-7}$	[0, 14000]
4	0.10470 ± 0.00033	$(2.391 \pm 0.563) \cdot 10^{-7}$	[0, 11500]
8	0.10458 ± 0.00041	$(2.152 \pm 1.155) \cdot 10^{-7}$	[0, 6600]

Tabelle 27: Auflistung der Fit-Parameter im linearen Modell und Fit-Bereich bei einer Stichprobengröße von $n=3000$.

N_B	c_0	c_1	Fit-Bereich
1	0.10138 ± 0.00018	$(3.053 \pm 0.867) \cdot 10^{-8}$	[0, 39000]
2	0.10144 ± 0.00018	$(2.864 \pm 0.906) \cdot 10^{-8}$	[0, 38000]
4	0.10123 ± 0.00019	$(1.689 \pm 1.052) \cdot 10^{-8}$	[0, 35000]
8	0.10154 ± 0.00023	$(1.915 \pm 1.978) \cdot 10^{-8}$	[0, 21000]

Tabelle 28: Auflistung der Fit-Parameter im linearen Modell und Fit-Bereich bei einer Stichprobengröße von $n=10000$.

Im Folgenden wird quantifiziert wie gut die Schätzmethode ist. Dazu werden für die verwendeten Stichprobengrößen die maximalen und minimalen Werte von $\bar{\lambda}_{n_b}$ und $\hat{\sigma}_{\lambda_{n_b}}$ in den Tabellen 29 und 30 angegeben um so zu beurteilen, wie sich die Schätzfunktionen in Abhängigkeit von n_b verhalten.

$n \setminus N_B$	1	2	4	8
300	(0.14455 , 0.28125)	(0.14638 , 0.26600)	(0.14439 , 0.23196)	(0.13696 , 0.20602)
1 000	(0.11153 , 0.14594)	(0.11058 , 0.13686)	(0.10739 , 0.13055)	(0.09716 , 0.12515)
3 000	(0.10132 , 0.11337)	(0.10082 , 0.11174)	(0.09740 , 0.10963)	(0.09037 , 0.10807)
10 000	(0.09849 , 0.10689)	(0.09770 , 0.10345)	(0.09011 , 0.10462)	(0.09658 , 0.10565)
30 000	(0.09621 , 0.10202)	(0.09666 , 0.10191)	(0.09209 , 0.10113)	(0.09049 , 0.10188)

Tabelle 29: Auflistung der minimalen und maximalen Werte von $\bar{\lambda}_{n_b}$.

$n \setminus N_B$	1	2	4	8
300	(0.11520 , 0.53252)	(0.11825 , 0.52745)	(0.11768 , 0.45381)	(0.11601 , 0.35062)
1 000	(0.05247 , 0.16000)	(0.05024 , 0.13144)	(0.05253 , 0.11932)	(0.05332 , 0.10578)
3 000	(0.02965 , 0.06240)	(0.02979 , 0.06422)	(0.02937 , 0.05596)	(0.03022 , 0.05635)
10 000	(0.01607 , 0.01612)	(0.01612 , 0.03592)	(0.01643 , 0.03481)	(0.01623 , 0.03528)
30 000	(0.00926 , 0.01908)	(0.00918 , 0.01919)	(0.00918 , 0.02018)	(0.00916 , 0.01783)

Tabelle 30: Auflistung der minimalen und maximalen Werte von $\hat{\sigma}_{\lambda_{n_b}}$.

Die maximalen und minimalen Werte für $\bar{\lambda}_{n_b}$ und $\hat{\sigma}_{\lambda_{n_b}}$ wurden dabei schon bei der Stichprobe von $n = 1000$ untersucht und bewertet. Wir bewerten damit die Schätzfunktion für λ nun global auf dem gesamten Raum der n_b und der N_B .

Die Aussage soll also auf verschiedene Stichproben ausgeweitet werden. Wir sehen, dass die Spanne der $(\bar{\lambda}_{n_b,min}, \bar{\lambda}_{n_b,max})$ über alle n_b für steigende n immer kleiner wird. Die Werte für den Fehler auf $\bar{\lambda}_{n_b}$ nehmen außerdem für steigende n ab, allerdings bringt es hier keinen Vorteil mehrere Off-Regionen zu benutzen, da die Standardabweichungen davon nicht statistisch signifikant beeinflusst bzw. verkleinert werden.

Wichtig für eine geeignete Schätzfunktion ist, dass sie mit wachsender Stichprobengröße, sich der wahren Größe immer weiter annähert und dabei weniger streut. Dies entspricht hier der immer kleiner werdende Spanne der minimalen und maximalen Werte und dass diese Spanne ab einem bestimmten n auch den wahren Wert von λ_0 beinhaltet. Wir sehen ferner, dass die Berücksichtigung von mehreren Off-Regionen diese Spanne weiter verkürzt. Alles in allem lässt sich trotz Untergrund allein mit der erweiterten-Maximum-Likelihood-Methode unter dessen Berücksichtigung Aussagen bzw. Schätzungen des Parameters λ machen.

Der linke Plot aus Abbildung 30 zeigt die gefitteten Werte für die positiven Steigungen in Abhängigkeit von n am Beispiel von $N_B = 8$. Hier zeigt ein Fit auf der doppelt-logarithmischen Darstellung eine Verträglichkeit von $p = 0.994$. Die Beschreibung durch einen linearen Fit ist auch für $N_B = 4$ geeignet, für kleinere N_B nicht, in diesem Falle lässt sich der Verlauf aber linear approximieren. Aus diesem Fit lässt sich für beliebige Stichprobengrößen n die passende Steigung für die Gerade an $\bar{\lambda}_n$ bestimmen.

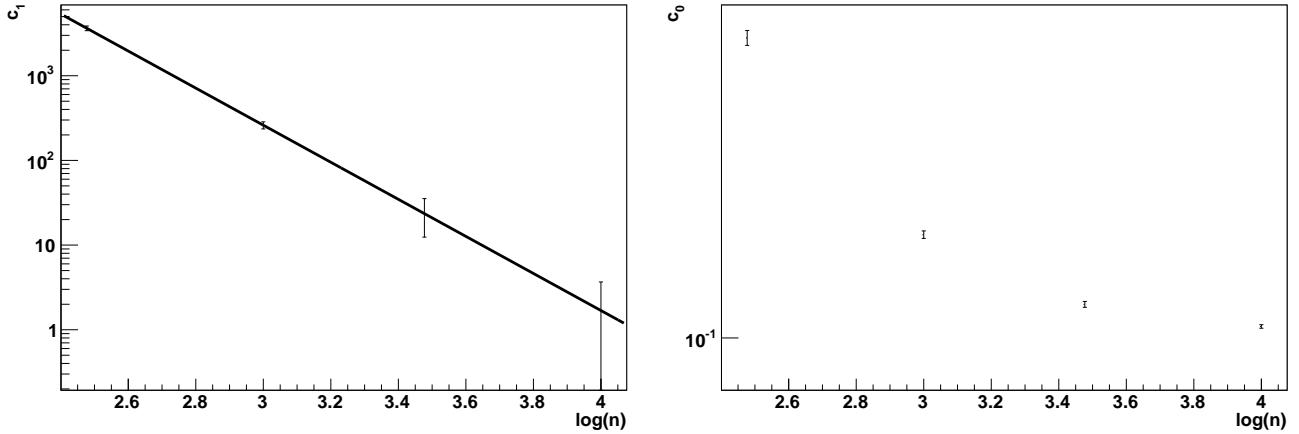


Abbildung 30: Die Plots zeigen die Auftragung des logarithmischen Werts der Stichprobengröße n und den zugehörigen Wert für die Koeffizienten c_1 (mit 10^{-8} multipliziert) bzw. c_0 mit den zugehörigen Fehlern $N_B = 8$. In doppelt-logarithmischen Darstellung wird hier ein lineares Modell an die Werte von c_1 gefittet. Die Anpassung hat dabei einen p-Wert von $p = 0.994$. Mit dem linearen Modell lassen sich also die Werte von c_1 in Abhängigkeit von n beschreiben. Für die Werte von c_0 ist die Beschreibung mit einer geraden in der Auftragung nicht passend, hier gilt es mehrere Datenpaare von (n, c_0) zu sammeln um einen komplexeren Fit an die Daten vornehmen zu können.

Abschließend hier noch einige Gedanken zur Verbesserung des Schätzverhaltens bei Untergrund. Der rechte Plot aus Abbildung 30 zeigt die Abhängigkeit des Fit-Geraden-Parameters c_0 von der Stichprobengröße. Diese Abhängigkeit erweist sich aber in der doppelt-logarithmischen nicht als linear, sodass keine Potenzgesetzabhängigkeit der beiden Größen vorliegt. Stehen einem mehrere Daten bzw. Wertepaare von der Stichprobengröße und Koeffizienten c_0 zur Verfügung lässt sich eine komplexere Kurve an die Werte fitten, um auch hier aussagekräftige Prognosen über andere Stichprobengrößen machen zu können. Hat man nun eine Stichprobe der Quellregion und Off-Regionen vorliegen und bestimmt anhand der erweiterten Maximum-Likelihood-Methode unter Berücksichtigung des Untergrunds die Parameter von $g(x)$ kann man dieses Vorgehen auch damit ersetzen, dass man λ durch den Wert von c_0 für am Besten nur eine Off-Region übernimmt. Dazu benutzt man die Anzahl der Ereignisse der Off-Region als Schätzwert für den Untergrundanteil des Signals aus der Quellregion und schätzt so letztlich die Zahl der Stichprobe n . Für ein n lässt sich dann aus anderen simulierten Werten von c_0 für diverse n und eine daran angepasste Kurve der gesuchte Wert von c_0 bestimmen. Man erhält so einen Schätzwert für λ , welches sich, siehe Kapitel 3-5, als Schätzwert des wahren λ_0 mit einem Bias erweist. Allerdings hat man hier wiederum das Problem, dass dies nur funktioniert, wenn λ_0 bekannt ist. Ähnlich zu Kapitel 7 gilt es hier eine Verbesserung des Schätzers von λ anzustreben. Dazu sind allerdings weitere Untersuchungen mit verschiedenen Werten von λ_0^i nötig.

10 Hypothesentests

In diesem Kapitel wird zunächst der Frage nachgegangen, wie man statistisch entscheiden kann, ob sich Daten hinreichend mit einem Potenzgesetz beschreiben lassen oder sie die Einführung eines zusätzlichen exponentiellen Cutoffs notwendig machen. Im nächsten Schritt wird man dies auf eine allgemeinere Verteilung als das Potenzgesetz mit exponentiellem Cutoff erweitern.

10.1 Test auf reines Potenzgesetz gegen Potenzgesetz mit exponentiellem Cutoff

Wie lässt sich eine Quelle dahingehend untersuchen, ob die zugehörige Energieverteilung einen exponentiellen Cutoff besitzt oder sich hinreichend durch ein reines Potenzgesetz beschreiben lässt? Um diese Frage zu beantworten, konstruiere man einen *Likelihood-Ratio-Test*. Es ergeben sich die beiden Hypothesen mit den Parametermengen:

$$H_0 : f_\alpha(x) = \Phi_0 x^{-\alpha}, \quad (\Phi_0, \alpha) \in \Theta_0 = (\mathbb{R}^+, \mathbb{R}^+), \quad x \in [m, m']$$

$$H_1 : f_{\alpha, \lambda}(x) = \Phi_0 x^{-\alpha} e^{-\lambda x}, \quad (\Phi_0, \alpha, \lambda) \in \Theta_1 = (\mathbb{R}^+, \mathbb{R}^+, \mathbb{R}^+), \quad x \in [m, m']$$

H_0 ist damit der Spezialfall von H_1 für $\lambda = 0$.

Wenn der Test auf gebinnte Daten mit K Bins angewandt wird, sind die zugehörigen Wahrscheinlichkeitsdichten im Zähler und Nenner jeweils Poisson-Zähldichten mit Parametern $l_i(\alpha) = \int_{A_i} f_\alpha(x) dx$ für das reine Potenzgesetz bzw. $l_i(\alpha, \lambda) = \int_{A_i} f_{\alpha, \lambda}(x) dx$ für das Potenzgesetz mit exponentiellem Cutoff. Die Wahrscheinlichkeit im i-ten Bin n_i Ereignisse zu erhalten, modelliert durch die Zufallsvariable N_i , ist also:

$$\mathcal{P}(N_i = n_i) = \frac{e^{-l_i} \cdot l_i^{n_i}}{n_i!}, \quad \text{für } l_i \in \{l_i(\alpha), l_i(\alpha, \lambda)\}$$

i ist dabei der Index für das i-te Bin, mit $i = 1, \dots, K$, und A_i ist das zugehörige Interval des Bins auf der x-Achse. Als Poisson-Parameter für das i-te Bin ist l_i somit gleich der Anzahl zu erwartender Ereignisse und gleichzeitig dessen Varianz. Es ergibt sich hier angewandt auf die Situation folgende Teststatistik Λ :

$$\Lambda(N_1, \dots, N_K) = -2 \log \left(\frac{\sup_{(\Phi_0, \alpha) \in \Theta_0} \prod_{i=1}^K \frac{e^{-l_i(\alpha)} l_i(\alpha)^{N_i}}{N_i!}}{\sup_{(\Phi_0, \alpha, \lambda) \in \Theta_1} \prod_{i=1}^K \frac{e^{-l_i(\alpha, \lambda)} l_i(\alpha, \lambda)^{N_i}}{N_i!}} \right)$$

Nach Kapitel 2.4 gilt für die Verteilung von Λ asymptotisch die χ^2 -Verteilung mit einem Freiheitsgrad. Die Frage ist hier nun, wie gut die Verteilung der Teststatistik $\Lambda(N_1, \dots, N_K)$ in Abhängigkeit von der Stichprobengröße n mit der χ^2 -Verteilung übereinstimmt, sofern H_0 wahr ist, also ein reines Potenzgesetz vorliegt, und man so deren Quantile benutzen kann oder man für kleine n die Verteilung der Teststatistik besser simulieren sollte.

Die Ergebnisse für $N = 10\,000$ Simulationen für $K = 72$ sind in der Tabelle 31 aufgelistetet. Dazu gehören die arithmetischen Mittelwerte $\bar{\Lambda}_n$, empirische Standardabweichungen σ_Λ , die empirische Quantile $\tilde{q}_{1-\epsilon}$ (für $\epsilon = 0.05$) in Abhängigkeit der Stichprobengröße n bezogen auf die Zufallsvariable Λ . Außerdem enthält die Tabelle die aufsummierte Teststatistik eines χ^2 -Anpassungstests, $\sum_{i=1}^M \left(\frac{k_i - N \cdot p_i}{\sqrt{N \cdot p_i}} \right)^2$, anhand der entschieden wird, ob die simulierte Verteilung der Λ mit der χ^2 -Verteilung verträglich ist.

Die Tabelle 32 fasst die theoretischen Kenngrößen der χ^2 -Verteilung zusammen.

n	$\bar{\Lambda}$	σ_{Λ}	$\sum_{i=1}^M \left(\frac{k_i - N \cdot p_i}{\sqrt{N \cdot p_i}} \right)^2$	$\tilde{q}_{1-\epsilon}$
300	0.943	1.375	93.81	3.769
1 000	0.946	1.406	87.02	3.809
3 000	0.995	1.442	83.04	3.875
10 000	0.982	1.419	61.43	3.859

Tabelle 31: Auflistung der Kenngrößen der simulierten Verteilung von Λ wie Mittelwert, Standardabweichung, aufsummierte Teststatistik des Anpassungstests und empirisches 95%-Quantil.

$E(X)$	$V(X)$	q_{ϵ}
1.0	1.414	3.841

Tabelle 32: Auflistung der theoretischen Werte der χ^2 -Verteilung, wie Erwartungswert und Standardabweichung.

Die zugehörigen Verteilungen der Zufallsvariable Λ sind in Abbildung 31 dargestellt.

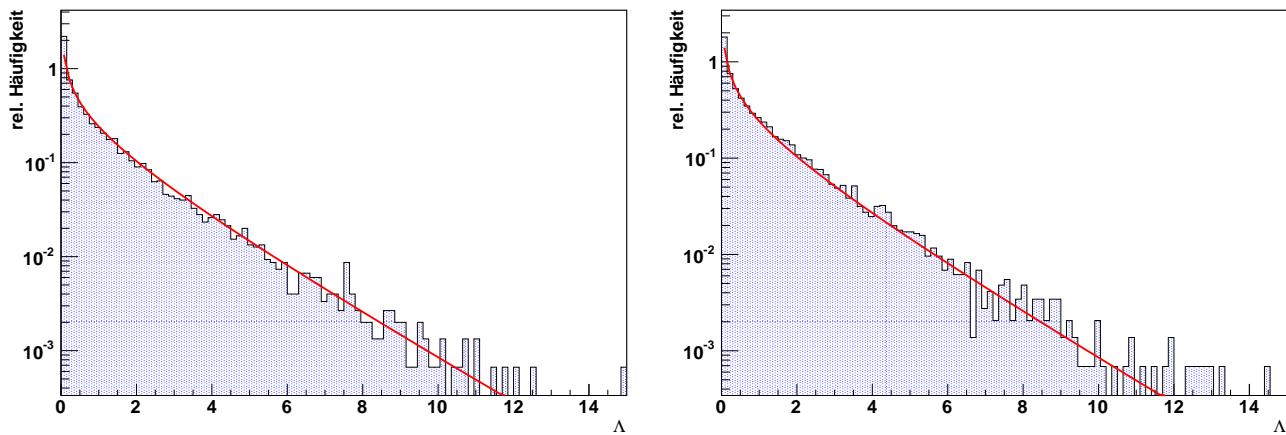


Abbildung 31: Simulierte Verteilungen von Λ in blau, theoretische χ^2 -Dichtefunktion in rot. Die x-Achse kennzeichne die Werte von Λ , die y-Achse den zugehörigen Wert der Verteilung. Links für $n = 300$ und rechts für $n = 10000$. Jeweils logarithmische Darstellung der y-Achse.

Man sieht in den obigen Tabellen 31 und 32, wie sich die empirischen Kenngrößen den wahren Kenngrößen der χ^2 -Verteilung annähern, mit der Ausnahme für das arithmetische Mittel bei $n=10000$. Allerdings unterliegt dieses natürlich auch einem Fehler der sich aus der Standardabweichung durch die Wurzel von N ergibt, hier ca. 0.014, das könnte dafür sprechen, dass der kleinere Abstand zum Mittelwert rein statistisch begründet ist.

Letztlich ist es allerdings so, dass nur für $n = 10 000$ statistisch gesehen davon ausgegangen werden kann, dass eine χ^2 -Verteilung von Λ vorliegt.

Entscheidend für den Hypothesentest ist allerdings nur das $\tilde{q}_{1-\epsilon}$ -Quantil, selbst bei $n = 10000$

macht man hier noch einen kleinen Fehler mit der Abweichung von 0.01.

Um nun wiederum statistisch untermauert entscheiden zu können, wie gut die χ^2 -Verteilung mit der simulierten Verteilung von Λ übereinstimmt, testet man die simulierte Verteilung mittels eines χ^2 -Anpassungstests auf die χ^2 -Verteilung.

Dazu untersucht man die simulierte Verteilung von Λ , die wiederum gebinnt in einem Histogramm vorliegt. Man summiert nun die quadrierten Abweichungen der Anzahl der Ereignisse k_i in einem Bin von der erwarteten Anzahl von Ereignissen $N \cdot p_i$ in diesem Bin und gewichtet sie mit dem zugehörigen Fehler, hier aufgrund der angenommenen Poissonstatistik, gleich der Wurzel der Zahl erwarteter Ereignisse. Man erhält also die Teststatsitik: $\sum_{i=1}^M \left(\frac{(k_i - N \cdot p_i)}{\sqrt{N \cdot p_i}} \right)^2$, wobei M die Anzahl der Bins ist, $p_i = \int_{A_i} f_{\chi^2}(x) dx$, die Wahrscheinlichkeit innerhalb der χ^2 -Verteilung (f_{χ^2} ist die zugehörige Dichtefunktion) einen Wert in dem i-ten Bin zu erhalten. Dieser Quotient ist für eine große Anzahl von Einträgen χ^2_{M-1} -verteilt, hier beträgt der Freiheitsgrad $M - 1$. Da sich im Kontext der Poisson-Statistik die Anzahl der Einträge k_i in einem Bin i für große Stichprobenzahlen gut mit der Normalverteilung approximieren lassen und der Quotient in der Teststatistik so eine standardnormal-verteilte Zufallsvariable im Quadrat ist und diese wiederum χ^2 -verteilt ist, ergibt sich in der M-fachen Summe eine χ^2_{M-1} -verteilte Größe. Einen Freiheitsgrad verliert man aufgrund der Zwangsbedingung, dass sich in der Summe der Bins wiederum die Gesamtzahl der Einträge N ergibt.

Die Testentscheidung ergibt sich dann analog: Für die Nullhypothese - die simulierte Verteilung folgt der χ^2 -Verteilung - muss die Summe der Quotienten kleiner als das $(1 - \epsilon)$ -Quantil der χ^2 -Verteilung sein: Also $\sum_{i=1}^M \left(\frac{(k_i - N \cdot p_i)}{\sqrt{N \cdot p_i}} \right)^2 < \chi^2_{1-\epsilon; M-1} = 66.35$, andernfalls kann man annehmen, dass es sich um eine andere Verteilung handeln wird. Dieser Test wurde hier auf den Bereich von $[0, 8]$ auf der x-Achse angewandt, weil nur in dem Bereich die wichtige Voraussetzung für den Test erfüllt ist, diese lautet $N \cdot p_i > 5, \forall i \in \{1, \dots, M\}$.

Mit den Ergebnissen aus Tabelle 31 ergibt sich, dass in guter Näherung für Stichprobenzahlen ab 10000 die Teststatistik Λ χ^2 -verteilt ist, bzw. dessen Quantile benutzen werden können. Benutzt man allerdings andere Signifikanzniveaus, hier $\epsilon = 0.05$, ergeben sich andere Quantile als Entscheidungskriterium, diese gilt es wiederum auf ihre Verträglichkeit zu überprüfen.

Man kann also mit dem vorgestellten Test statistisch begründen, ob sich eine Stichprobe mit einem reinen Potenzgesetz beschreiben lässt oder ob man diesen Zusammenhang auf den Fall mit dem exponentiellem Cutoff erweitern muss.

Wie gut dabei diese Verteilung passt, wird nicht beantwortet, möglicherweise zeigt sich auch für Stichproben, die aus anderen Verteilungen mit einer ähnlichen Assymetrie in ihrer Dichtefunktion gezogen wurden, dass die Situation hinreichend mit einem Potenzgesetz beschrieben wird.

Dieser Test ist deshalb nur sinnvoll für die Anwendung auf die Situation, dass man Potenzgesetz mit Potenzgesetz und exponentiellem Cutoff vergleicht.

Des Weiteren kann man sich noch eine Frage stellen: Wie ist die Verteilung der geschätzten λ , die sich aus dem Maximum-Likelihood-Verfahren ergeben, angewandt auf eine Stichprobe, die nur aus einem Potenzgesetz gezogen wurde? Damit sind jene Werte von λ gemeint, die in der Teststatistik Λ im Nenner geschätzt werden.

Die obigen Simulationen, deren Ergebnisse in den Tabellen 31 und 32 zusammengefasst sind, wird nun also auf die geschätzten Werte von λ , also den jeweiligen geschätzten exponentiellen Cutoffs untersucht. Die Abbildung 32 zeigt die Verteilung der λ .

Interessant ist hier wieder das $\tilde{q}_{0.95}$ -Quantil, dessen Aussage es ist, dass 95% der Werte von λ unterhalb dieses Wertes liegen, dies entspricht einer oberen Grenze für den Wert von λ . Das bedeutet, dass anhand von Daten, die nach einem Potenzgesetz verteilt sind, daraus Werte für λ rekonstruiert werden, von denen 95% davon unter diesem Wert liegen. Das heißt, erhielte man ein λ mit einem Wert unterhalb dieser Grenze, spricht es statistisch gesehen dafür, dass

die zugehörige Verteilung sich hinreichend durch ein reines Potenzgesetz beschreiben lässt. Wie zu erwarten, sinkt diese obere Grenze mit höhereren Stichprobenzahlen n . Dies zeigt die Tabelle 33. Aus den enthaltenen Werten lässt sich kuryum sagen, dass die simulierte Verteilung mit wachsender Stichprobengröße immer näher um 0 zentriert ist, entsprechend dem, dass die geschätzten Werte im Nenner von Λ wiederum für wachsende n besser geschätzt werden, im Sinne von weniger Abweichung und weniger Streuung um den wahren Wert $\lambda = 0$.

Diese Verteilung der λ lässt sich auch simulieren, wenn man die genaue Verteilung nicht kennt und nur die Information einer Stichprobe hat. Man passt an diese Stichprobe die Parameter des Potenzgesetzes an und zieht dann gemäß der ausgewerteten Stichprobe erneut Zufallszahlen und simuliert die Verteilung der λ .

n	$\bar{\lambda}$	σ_{λ}	\tilde{q}_{ϵ}
300	0.2011	0.0825	0.0134
1 000	0.0089	0.0224	0.0049
3 000	0.0027	0.0067	0.0015
10 000	0.0007	0.0019	0.0004

Tabelle 33: Auflistung der arithmetischen Mittelwerte, Standardabweichung und \tilde{q}_{ϵ} von λ , die sich aus der Simulation von Λ ergeben, für verschiedene Stichprobengrößen n .

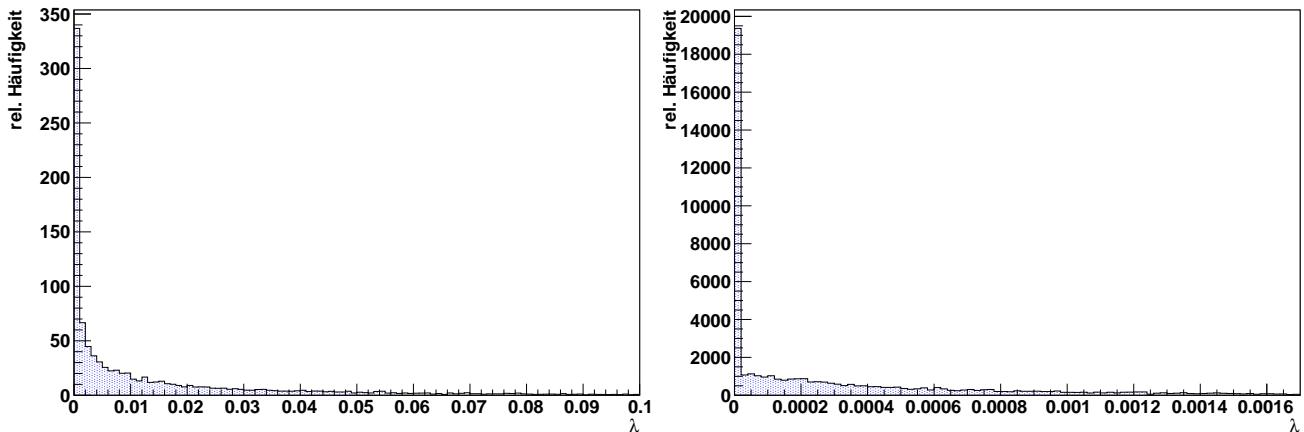


Abbildung 32: Simulierte Verteilung von λ in blau, links für $n = 300$ rechts $n = 10.000$. Die x-Achse kennzeichne die Werte von λ , die y-Achse die Verteilung.

10.2 Test auf Potenzgesetz gegen beliebige Verteilung

Möchte man die Hypothese des reinen Potenzgesetzes gegenüber einer beliebigen Verteilung testen, so ergibt sich für die Alternativhypothese H_1 die Annahme, dass die Zahl der Ereignisse N_i in jedem Bin poisson-verteilt ist mit Parameter l_i für das i-te Bin. Diese Poisson-Parameter werden dabei unabhängig voneinander an die Zahl der Ereignisse im jeweiligen Bin angepasst. Es ergeben sich so im Allgemeinen für jedes Bin verschiedene Poisson-Parameter.

$$H_0 : f_\alpha(x) = \Phi_0 x^{-\alpha}, \quad (\Phi_0, \alpha) \in \Theta_0 = (\mathbb{R}^+, \mathbb{R}^+)$$

$$H_1 : f_{l_1, \dots, l_K}(x) = \prod_{i=1}^K \frac{e^{-l_i} \cdot l_i^{n_i}}{n_i!}, \quad (l_1, \dots, l_K) \in \Theta_1 = \mathbb{R}^{+K}$$

Für den Maximum-Likelihood-Quotienten ergibt sich folgender Ausdruck.

Im Zähler erhält man hier gemäß der Poisson-Statistik also, dass der jeweilige Poisson-Parameter eines Bins dem jeweiligen Wert der Zufallsvariablen entspricht. Im Nenner gilt es nach wie vor mittels der freien Parameter die Wahrscheinlichkeit numerisch zu maximieren.

$$\Lambda(N_1, \dots, N_K) = -2 \log \left(\frac{\sup_{(\Phi_0, \alpha) \in \Theta_0} \prod_{i=1}^K \frac{e^{-l_i(\alpha)} l_i(\alpha)^{N_i}}{N_i!}}{\sup_{(l_1, \dots, l_K) \in \Theta_1} \prod_{i=1}^K \frac{e^{-l_i} l_i^{N_i}}{l_i!}} \right) = -2 \log \left(\frac{\sup_{(\Phi_0, \alpha) \in \Theta_0} \prod_{i=1}^K \frac{e^{-l_i(\alpha)} l_i(\alpha)^{N_i}}{N_i!}}{\prod_{i=1}^K \frac{e^{-N_i} N_i^{N_i}}{N_i!}} \right)$$

Nach dem Theorem von Wilk gilt für die Verteilung von Λ wiederum im Grenzfall mit $n \rightarrow \infty$ die χ^2 -Verteilung, hier allerdings mit dem Freiheitsgrad m , mit $m = K - r$, wobei r die Anzahl der freien Parameter der Nullhypothese ist, hier also $r = 2$. Die Zahl der Bins liegt nach wie vor bei $K = 72$, der Freiheitsgrad der angenäherten χ^2 -Verteilung ist also $m = 70$.

Wie die Abbildung 33 mit den simulierten Verteilungen von Λ und die Tabelle mit den zugehörigen Mittelwerten, der Standardabweichungen und den empirischem 95%-Quantil (hier $\epsilon = 0.05$) zeigen, lässt sich diese erst für extrem hohe Stichprobenzahlen mit der χ^2 -Verteilung approximieren.

Hier gilt es also die Verteilung zu simulieren, um so jeweils die Quantile zu bestimmen, die die Testentscheidung festlegen.

n	$\bar{\Lambda}$	σ_Λ	$\tilde{q}_{1-\epsilon}$
1 000	34.82	8.45	49.74
10 000	48.59	10.06	66.45
100 000	60.97	11.08	80.59
5 000 000	70.13	11.78	90.58

Tabelle 34: Auflistung der arithmetischen Mittelwerte, Standardabweichungen und empirisches 95%-Quantil der simulierten Verteilung von Λ in Abhängigkeit der Stichprobengröße n .

$\mathbb{E}(X)$	$\mathbb{V}(X)$	$q_{1-\epsilon}$
70.00	11.83	90.54

Tabelle 35: Theoretische Kennröhren der χ^2 -Verteilung mit Freiheitsgrad $m = 70$ wie Erwartungswert, Standardabweichung und Wert des 95%-Quantils.

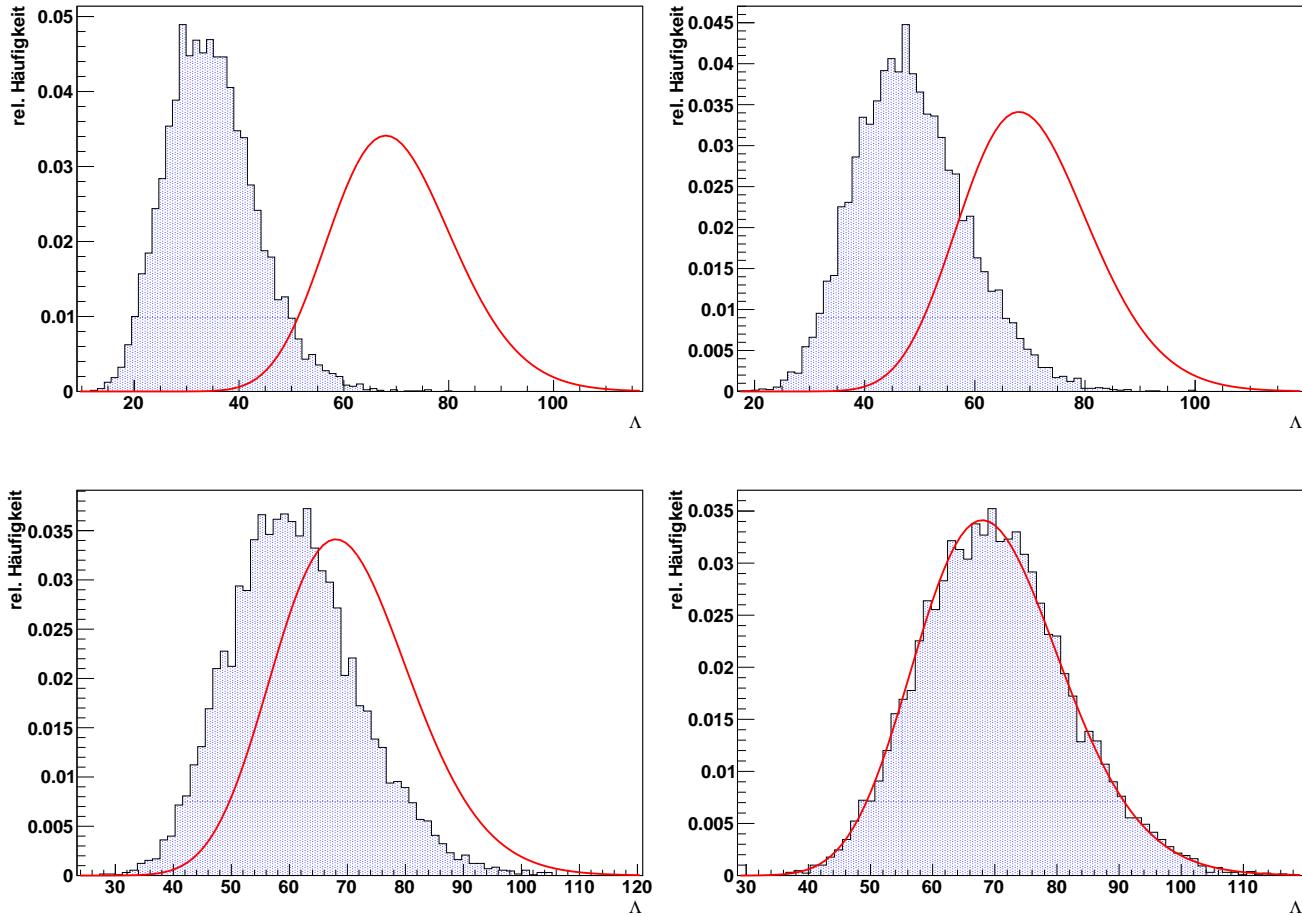


Abbildung 33: Simulierte Verteilungen von Λ (blau). Oben links für $n = 1000$ und rechts für $n = 10000$, und unten links für $n = 100000$ und rechts für $n = 5000000$. In Rot ist jeweils die Dichte der theoretischen χ^2 -Verteilung mit Freiheitsgrad 70 eingezeichnet.

Man sieht im Vergleich zum vorherigen Test, dass erst für eine hohe Stichprobenzahl von 5 Millionen die simulierte Verteilung mit der χ^2 -Verteilung mit Freiheitsgrad m verträglich ist. Testet man die empirische Verteilung bei dieser Stichprobenzahl auf die χ^2 -Verteilung bei einem Freiheitsgrad von 70, ergibt sich bei einem χ^2 -Test auf diese Verteilung ein Wert von 55.14 bei 50 bins. Bei kleineren Stichprobenzahlen gilt es also hier die Verteilung von Λ zu simulieren und anhand deren das notwendige Quantil zur Entscheidungsfindung zu bestimmen. Ein Vorteil ist es, dass für größere Binzahlen, ab $K = 100$ unabhängig von der Stichprobengröße, die Verteilung gut durch eine Normalverteilung approximiert wird.

Insgesamt lässt sich festhalten, dass sich mithilfe von konstruierten Likelihood-Ratio-Tests statistisch entscheiden lässt, ob eine gemessene Stichprobe bzw. Daten einem Potenzgesetz folgen oder wenn sie sich damit nicht hinreichend erklären lassen, davon ausgehen kann, dass die Daten tatsächlich einem Potenzgesetz mit exponentiellem Cutoff oder einer allgemeinen Poisson-Verteilung folgen.

Allerdings muss bei der allgemeinen Poisson-Verteilung die Verteilung derjenigen Teststatistik in Abhängigkeit der Stichprobe erst simuliert werden, um so den Test zu konstruieren. Bei dem Spezialfall mit dem Potenzgesetz mit exponentiellem Cutoff lässt sich ab Stichprobengrößen von $n=10000$ die Quantile der χ^2 -Verteilung benutzen, benutzt man diese schon bei kleineren Stichprobengrößen macht man nur einen kleinen Fehler.

11 Zusammenfassung und Ausblick

Einige gemessene Energiespektren des H.E.S.S-Experiments haben gezeigt, dass sich diese nicht mehr hinreichend durch ein Potenzgesetz beschreiben lassen. Diese Spektren zeigen hier bei hohen Energien einen steileren Abfall. Dies macht die Erweiterung des Potenzgesetzes, $\Phi_0 E^{-\alpha}$, mit einem exponentiellen Cutoff der Form $\exp(-\lambda \cdot E)$ bzw. $\exp(-\frac{E}{E_c})$ mit der Wahl eines positiven Cutoff-Parameters λ bzw. E_c notwendig. Das Wissen über diesen Cutoff-Parameter lässt Rückschlüsse auf den Produktionsprozess und die zugrundeliegenden Teilchenpopulation zu. Das wiederum bedarf einer Bestimmung der Parameter mit nur wenig systematischer Abweichung vom jeweiligen wahren Wert. Die Schwierigkeit beim statistischen Schätzen des Cutoff-Parameters ist es, dass Ereignisse mit hohen Energien in dem Bereich, wo das gemessene Spektrum von der Beschreibung durch ein Potenzgesetz abweicht, nur in geringer Zahl zu erwarten sind. So erhält man nur geringe Statistik, was im Allgemeinen eine Schätzung des Parameters mit wenig Abweichung vom wahren Wert schwierig macht.

Ziel dieser Arbeit war die Untersuchung der Qualität der Schätzfunktionen für die Parameter α und λ der Energieverteilung der kosmischen Gammastrahlung. Diese untersuchten Schätzfunktionen ergeben sich aus der erweiterten Maximum-Likelihood-Methode und der Ersetzungsmethode. Dabei wurden diese Schätzmethoden zunächst auf die Messung eines reinen Signals von Photonen der Gammastrahlung angewandt.

Um die Schätzmethoden zu bewerten, wurden jeweils Simulationen durchgeführt, die alle mittels der gleichen Parameter, $\alpha_0 = 2.0$ und $\lambda_0 = 0.1$ für bestimmte Stichprobengrößen durchgeführt wurden. Diese sind an tatsächlich gemessene Werte dieser Parameter in der Gammaastronomie angelehnt. Um die Schätzfunktionen auch untereinander zu vergleichen wird hier jeweils der mittlere quadratische Fehler bestimmt, der die quadrierte systematische Abweichung des geschätzten Mittelwerts zum wahren Wert angibt und zusätzlich die Streuung der Schätzfunktion berücksichtigt. Dies ist die allgemeinste Form um Schätzfunktionen zu beurteilen.

Zunächst wurde die erweiterte Maximum-Likelihood-Methode, die in der Gammaastronomie das übliche statistische Instrument zur Bestimmung der Parameter ist, untersucht. Der große Unterschied zwischen der erweiterten Maximum-Likelihood-Methode zur gewöhnlichen Maximum-Likelihood-Methode liegt in der zusätzlichen Anpassung an die Gesamtanzahl der gemessenen Ereignisse. Durch Ausnutzung dieser Information wird so das Schätzergebnis verbessert. Diese Methode wurde auf ungebinnedne und gebinnedne Daten angewandt. Hier zeigte sich, dass die gebinnedne Version bessere Ergebnisse für die Schätzung des Parameters λ für kleine Stichproben bis $n = 1000$ erzielt. Hier zeigt sich, dass die Abweichung vom wahren Wert zwar leicht größer ist, die Streuung aber deutlich kleiner, sodass sich hier ein wiederum deutlich kleinerer mittlerer quadratischer Fehler ergibt, der im Falle der ungebinnednen Version circa drei mal größer ist. Für größere Stichprobenzahlen zeigen sich hier nur noch vernachlässigbare Unterschiede. Bei der Schätzung des Parameters α zeigen sich hier vernachlässigbare Unterschiede für alle untersuchten Stichprobengrößen. Der Informationsverlust, der mit dem Binning der Daten automatisch einhergeht, führt also nicht zu schlechteren Ergebnissen.

Für ungebinnedne Daten wurde außerdem eine alternative Methode entwickelt, die auf der Ersetzungsmethode basiert und anhand der Gammaverteilung analytisch vorgestellt wurde. Die Idee ist es hier, anhand einer Messung gleichzeitig zwei Parameter zu schätzen, indem der Erwartungswert mit dem arithmetischen Mittelwert und der Median der Verteilung mit dem Stichprobenmedian ersetzt wird. Diese Methode wurde auch mit der Kombination aus Erwartungswert und Varianz durchgeführt, bei der man also wiederum den Erwartungswert der Verteilung mit dem arithmetischen Mittelwert und die Varianz der Verteilung mit der Stichprobenvarianz ersetzt. Außerdem wurde bei der Ersetzungsmethode der Bereich der Energien, auf dem die Verteilung definiert ist, durch das Stichprobenminimum geschätzt, die obere Grenze wurde zuvor auf unendlich festgesetzt. Bei der erweiterten Maximum-Likelihood-Methode wurde

hier auf einem bereits vorher definierten Energiebereich die Anpassung vollzogen. Dieser entspricht dem Bereich, in dem abhängig vom jeweiligen Experiment, Gammastrahlung detektiert wird. Da sich diese Arbeit an der Analyse und Messungen des H.E.S.S.-Experiments orientiert, liegt der untersuchte Energiebereich, auf dem die Zufallszahlen generiert wurden bei 100 GeV bis 100 TeV. Der Vergleich der Methoden unterschieden durch die Nutzung von Stichprobenvarianz und Stichprobenmedian hat ergeben, dass die Methode mit dem Median, gemessen an den mittleren quadratischen Fehlern, hier die Beste ist. Für eine Stichprobengröße von $n = 300$ zeigt sich hier ein mittlerer quadratischer Fehler der weniger als halb so groß ist wie bei der Methode, die die Varianz benutzt.

Alle drei Methoden, also die erweiterte Maximum-Likelihood-Methode für ungebinnedne und gebinnte Daten und die Ersetzungsmethode haben dabei gezeigt, dass die Schätzgrößen für die Parameter α und λ asymptotisch erwartungstreu sind. Dies ist gleichbedeutend mit der Eigenschaft der Konsistenz der Schätzer, bei der die Schätzfunktionen für wachsende Stichprobengrößen immer näher an ihrem wahren Wert liegen und weniger um diesen streuen. Somit ergibt sich für kleine Stichprobengrößen ein Bias, also ein systematischer Abstand vom wahren Wert. Die Schätzfunktion für α zeigt dabei immer das Verhalten, für alle Stichprobengrößen unter dem wahren Wert zu liegen, die Schätzfunktion für λ liegt hingegen für alle Stichprobengrößen über dem wahren Wert. Ein Vergleich der beiden ungebinnednen Methoden, also die erweiterte Maximum-Likelihood-Methode und die Ersetzungsmethode basierend auf Erwartungswert und Median, hat hier gezeigt, dass die Schätzfunktion $\hat{\lambda}$ bei der Ersetzungsmethode bei Stichprobengrößen bis 5000 einen kleineren mittleren quadratischen Fehler hat. Beim Vergleich hinsichtlich des Schätzverhaltens für α erweist sich die erweiterte Maximum-Likelihood-Methode, als die mit der kleineren mittleren quadratischen Fehler für alle Stichprobengrößen. Für große Stichprobengrößen liefern die drei Methoden allerdings hinsichtlich ihrer Schätzqualität gleichwertige Schätzfunktionen. Die erweiterte Maximum-Likelihood-Methode wurde aufgrund ihrer vergleichbaren Schätzqualität mit den ungebinnednen Versionen und der Tatsache, dass sie die vorherrschende Methode in der Gammasstronomie ist, in den weiteren Untersuchungen dieser Arbeit zur Analyse verwendet.

Als nächstes wurde die Schätzfunktion $\hat{\lambda}$, angewandt auf ein reines Gammasignal, schrittweise dahingegend verbessert, dass sie einen deutlich geringeren Bias aufweist. Dabei wurde berücksichtigt, dass der Wert der Schätzfunktion für α ebenso einen Bias besitzt und der wahre Wert unbekannt ist. Ausgehend von der Idee, dass man anhand von den Werten der simulierten Verteilung von $\hat{\alpha}$, den zugehörigen Mittelwert und die Werte benutzt, die genau eine Standardabweichung davon entfernt sind, wurden diese drei Werte als neue Startparameter für Simulationen für außerdem verschiedene Werte von λ benutzt. So wurde die Abhängigkeit der mittleren Werte von $\hat{\lambda}$ in Abhängigkeit von verschiedenen vorgegebenen Werten von λ untersucht. Hier zeigt sich, dass man die Wertepaare, also vorgegebenes λ_0 und geschätzter zugehöriger Mittelwert von $\hat{\lambda}$, durch eine Gerade mit einer Steigung größer als eins approximieren kann. Mit einem linearen Fit wurde mittels der Geradenparameter die Korrektur der Schätzgröße $\hat{\lambda}$ durchgeführt und kleinere Werte des mittleren quadratischen Fehlers erzielt. Hier zeigt sich für unkorrigierte Version für $n = 300$ ein relativer Fehler für die drei vorgegebenen Werte von λ mindestens doppelt so groß sind wie der relative Fehler der unkorrigierten Version.

Ferner wurden die Ergebnisse der Schätzmethode bei Berücksichtigung der effektiven Fläche des H.E.S.S-Instruments ohne die Korrektur aus dem vorherigen Kapitel untersucht. Um den statistischen Einfluss der effektiven Fläche zu beschreiben, lässt sich diese mit der Funktion $(1 - \exp(-s \cdot E))$ mit einem geeigneten positiven Parameter s modellieren. Damit lässt sich allerdings nur die hier untersuchte Schätzmethode entwickeln, eine vollständige Parametrisierung der effektiven Fläche hingegen nicht. Untersucht man hier die Ergebnisse der Schätzgrößen

bemerkt man, dass bei einer vorgegebenen Stichprobe die geschätzten Werte für α und λ deutlich näher am wahren Wert liegen. Der Grund hierfür ist, dass durch die Modellierung mit der effektiven Fläche, bei einer bestimmten Stichprobengröße anteilmäßig mehr Ereignisse mit hohen Energien detektiert werden, als ohne effektive Fläche. Die effektive Fläche führt also dazu, dass sich die Anzahl der Ereignisse, die dann detektiert wird, stark verringert. Dies entspricht keiner Umverteilung der Ereignisse, sondern der Tatsache dass Ereignisse mit hohen Energien eher detektiert werden als Ereignisse mit niedrigen Energien, sofern die effektive Fläche berücksichtigt wird. Die effektive Fläche lässt sich auch als energieabhängige Effizienz beschreiben, die die Wahrscheinlichkeit angibt, mit der ein Ereignis detektiert wird. Misst man also Daten mit einer gewissen Stichprobenanzahl muss also berücksichtigt werden, dass das eigentliche Energiespektrum durch die effektive Fläche modifiziert wurde. Dieses Energiespektrum lässt sich nun nicht mehr mit einem Potenzgesetz beschreiben. Vergleicht man aber die Ergebnisse der Schätzung der Parameter mit und ohne effektive Fläche unter Berücksichtigung bei vergleichbaren Stichprobengrößen, zeigen sich hier keine statistisch signifikanten Unterschiede.

Des Weiteren wurde die erweiterte Likelihood-Methode auf eine Stichprobe angewandt, die neben dem reinen Gammastrahlungssignal noch Untergrundsignal enthält. Dieses Untergrundsignal ergibt sich aus der kosmischen Strahlung mit einem bekannten spektralen Index von 2.7. Um die Zahl der Untergrundereignisse im Fit anzupassen, wurden mehrere Off-Regionen gewählt. Die Ergebnisse für die Schätzfunktion $\hat{\lambda}$ haben hier gezeigt, dass für eine konstante Zahl von Ereignissen der Gammastrahlung, der Bias für eine ansteigende Anzahl der Untergrundereignisse linear zunimmt. Ab einer bestimmten Zahl der Untergrundereignisse zeigt sich hier ein Verhalten, bei dem sich die Mittelwerte der Schätzfunktion für λ nicht mehr mit der ansteigenden Gerade beschreiben lassen, was hier eine Korrektur der Schätzfunktion erschwert. Beurteilt man die Schätzmethode $\hat{\lambda}$ hinsichtlich der Anzahl der Off-Regionen und der Gesamtanzahl der Untergrundereignisse zeigt sich, dass die Berücksichtigung mehrerer Off-Regionen dazu führt, dass die Werte von $\hat{\lambda}$ weniger vom wahren Wert abweichen und dabei weniger streuen.

Zum Abschluss der Arbeit wurden zwei Likelihood-Verhältnis-Tests vorgestellt. Mit dem einen lässt sich statistisch entscheiden, ob sich eine gemessene Datenreihe hinreichend mit einem Potenzgesetz beschreiben lässt oder die Erweiterung durch einen exponentiellen Cutoff notwendig ist. Damit lässt sich zum Beispiel statistisch testen, ob ein physikalischer Mechanismus der Gammastrahlungsquelle vorliegt, der zu einem Cutoff im Energiespektrum führt. Zur Untersuchung wurden Zufallszahlen, verteilt nach einem reinen Potenzgesetz mit spektralem Index von $\alpha_0 = 2.0$, für bestimmte Stichprobengrößen simuliert. Nach dem Wilk Theorem gilt hier, dass die Teststatistik des Likelihood-Verhältnis-Tests asymptotisch der χ^2 -Verteilung folgt. Da die Testentscheidung bei einem vorgegebenen Signifikanzniveau von 0.05 allein auf dem 95%-Quantil beruht, gilt es zu untersuchen, ob das empirische Quantil der simulierten Verteilung der Teststatistik, mit dem der χ^2 -Verteilung übereinstimmt. Hier zeigt sich schon für Stichprobengrößen von nur $n=300$ eine geringe Abweichung vom wahren Quantil-Wert der χ^2 -Verteilung mit relativem Fehler von 2%. Die simulierte Verteilung lässt sich dabei allerdings erst für $n = 10000$ mit der χ^2 -Verteilung approximieren. Für Stichprobengrößen unter $n = 10000$ empfiehlt es sich die Verteilung zu simulieren und deren empirisches Quantil zu benutzen.

Testet man das vorgegebene reine Potenzgesetz gegen eine allgemeine Poisson-Verteilung, wobei für jedes Bin ein eigener Poisson-Parameter gilt, sagt hier das Theorem nach Wilk, dass die asymptotische Verteilung wiederum die χ^2 -Verteilung ist, allerdings mit einem Freiheitsgrad, der der Differenz aus der Anzahl der Bins des befüllten Histogramms und der freien Parameter des Potenzgesetzes entspricht. Hier zeigte sich, dass erst für hohe Stichprobenzahlen von $n = 5000000$ die simulierte Verteilung sich durch diese χ^2 -Verteilung beschreiben lässt. Für die Testentscheidung gilt es hier also für kleine Stichprobenzahlen die empirischen Quantile der

simulierten Verteilung der Teststatistik zu benutztten.

Alles in allem hat sich innerhalb der Arbeit gezeigt, dass alle statistischen Schätzmethoden, Schätzfunktionen für die Parameter liefern, die die Eigenschaft der Konsistenz besitzen. Dies ist eine Minimalanforderung an Schätzfunktionen. Dies gilt auch bei Berücksichtigung der effektiven Fläche und Untergrundsignalen. Wichtig zu erwähnen ist, dass die untersuchten Methoden der erweiterten Maximum-Likelihood-Methode für gebinnte Daten und ungebinnte Daten und die Ersetzungsmethode nur anhand eines vorgegebenen Parametersets untersucht wurden. Um hier wirklich aussagekräftige Beurteilungen vorzunehmen, gilt es die Schätzfunktionen für mehrere mögliche Werte für die Parameter zu untersuchen. Auch das Verhalten, das in Kapitel 7 erkannt wurde, dass die Schätzfunktion für λ für alle vorgegebenen Werte immer oberhalb des wahren Wertes liegen, gilt es noch für verschiedene Werte von α zu testen. Hier gilt auch zu untersuchen wie sich eine Korrektur auf die Werte von $\hat{\lambda}$ auswirkt, wenn man Werte von α zur Bestimmung der Fitgeraden benutzt, wesentlich weiter vom mittleren α entfernt sind. Hier bietet sich das 95%-Interval der α -Werte an. Auch das Verhalten der Schätzfunktion α sollte hier für mehrere Werte untersucht werden. Da in dieser Arbeit die Beurteilung und Verbesserung der Schätzfunktion für λ im Vordergrund stand, kann man in weiteren Untersuchungen auch eine Korrektur von α vornehmen oder das Verhalten der Schätzfunktion bei Untergrund untersuchen. Als Aufgabe für zukünftige statistische Untersuchungen von Potenzgesetzverteilungen mit exponentiellem Cutoff innerhalb der Gammaastronomie stellt sich die gleichzeitige Berücksichtigung der effektiven Fläche, Untergrund und möglicher Korrekturen aller Schätzfunktionen.

Literaturverzeichnis

, <https://www.mpi-hd.mpg.de/hfm/HESS/pages/about/>; aufgerufen am 23.Oktober 2014.

Statistische Datenanalyse, 2012. http://www.math.uni-potsdam.de/prof/i_statistik/v_vorlesungen/StatDA1112-HPI/material/uebersicht-est; abgerufen am 02.Oktober 2014.

Aharonian, F., Akhperjanian, A. G., Bazer-Bachi, A. R., et al. Observations of the Crab nebula with HESS. *Astronomy & Astrophysics*, 457:899–915, Oct. 2006. doi: [10.1051/0004-6361:20065351](https://doi.org/10.1051/0004-6361:20065351).

Aharonian, F., Akhperjanian, A. G., Anton, G., et al. Spectrum and variability of the Galactic center VHE γ -ray source HESS J1745-290. *Astronomy & Astrophysics*, 503:817–825, Sept. 2009. doi: [10.1051/0004-6361/200811569](https://doi.org/10.1051/0004-6361/200811569).

Barlow, R. J. Extended maximum likelihood. *Nucl.Instrum.Meth.*, A297:496–506, 1990. doi: [10.1016/0168-9002\(90\)91334-8](https://doi.org/10.1016/0168-9002(90)91334-8).

Berge, D., Funk, S., & Hinton, J. Background modelling in very-high-energy γ -ray astronomy. *Astronomy & Astrophysics*, 466:1219–1229, May 2007. doi: [10.1051/0004-6361:20066674](https://doi.org/10.1051/0004-6361:20066674).

Clauset, A., Rohilla Shalizi, C., & Newman, M. E. J. Power-law distributions in empirical data. *ArXiv e-prints*, June 2007.

Cowan, G. Goodness of fit and wilk's theorem, 2013. , http://wwwusers.ts.infn.it/~milotti/Didattica/StatisticaAvanzata/Cowan_2013.pdf; aufgerufen am 10.August 2014.

Cronin, J. W., Gaisser, T. K., & Swordy, S. P. Cosmic Rays at the Energy Frontier. *Scientific American*, 276:44–49, Jan. 1997. doi: [10.1038/scientificamerican0197-44](https://doi.org/10.1038/scientificamerican0197-44).

Georgii, H. *Stochastik: Einführung in die Wahrscheinlichkeitstheorie und Statistik*. De Gruyter Lehrbuch. De Gruyter, 2009. ISBN 9783110215274. URL <http://books.google.de/books?id=iKXPRcYMR7MC>.

Grupen, C. *Astroteilchenphysik: Das Universum im Licht der kosmischen Strahlung*. vieweg, 2000. ISBN 3528031581.

Hansmann-Menzemer, S. Modern methods of data analysis, 2007. http://www.phys.uni-heidelberg.de/~menzemer/Stat0708/statistik_vorlesung_7.pdf; aufgerufen am 10.Oktober 2014.

H.E.S.S. Collaboration, Abramowski, A., Acero, F., et al. Measurement of the extragalactic background light imprint on the spectra of the brightest blazars observed with H.E.S.S. *Astronomy & Astrophysics*, 550:A4, Feb. 2013. doi: [10.1051/0004-6361/201220355](https://doi.org/10.1051/0004-6361/201220355).

Hujer, R. Goodness of fit and wilk's theorem, 2005. , www.wiwi.uni-frankfurt.de/professoren/hujer/Lehre/0ek_2/mle.pdf; aufgerufen am 26.Oktober 2014.

Kolanoski, P. D. H. Die maximum-likelihood-methode, 2008. http://www-zeuthen.desy.de/~kolanosk/smd_ss08/skripte/ml.pdf; aufgerufen am 30.September 2014.

Ligustri, R., 2014. <http://apod.nasa.gov/apod/ap141017.html>; abgerufen am 01.September 2014.

Longair, M. S. *High Energy Astrophysics*. Feb. 2011.

Schmidt, P. D. V. Statistik i, vorlesungsskript, 2004. <http://www.mathematik.uni-ulm.de/stochastik/lehre/ss04/statistik1/skript/node36.html>; abgerufen am 16.Oktobe 2014.

Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die zum Gelingen dieser Arbeit beigetragen und mich währenddessen unterstützt haben. Besonderer Dank gilt:

- **Prof. Dr. Christopher van Eldik** für die Vergabe der Arbeit und die gute Betreuung währenddessen.
- **Prof. Dr. Gisela Anton** für die Übernahme des Koreferats.
- **Der gesamten H.E.S.S.-Gruppe** für die nette Atmosphäre und die hilfreichen fachlichen Diskussionen, insbesondere:
- **Susanne Raab** für die Hilfe bei zahlreichen LaTeX-Problemen, vor allem bei Bibtex.
- **Susanne Raab, Johannes Veh und Stephanie Häffner** für das teilweise Korrekturlesen dieser Arbeit.
- **Stefan Pickel** für die große Hilfe beim Verfassen des Abstracts.
- **Philipp Willmann** für die Hilfestellung bei Computerproblemen und vor allem für das Korrekturlesen dieser Arbeit. Vielen Dank dafür!

Zum Schluss ein besonderer Gruß und Dank an meine Bürokollegen Eschbach und Ziegler. Vielen Dank für die schöne Zeit unter Freunden und Gleichgesinnten. Ich werde euch vermissen!

Erklärung

Hiermit bestätige ich, dass ich diese Arbeit selbstständig und nur unter Verwendung der angegebenen Hilfsmittel angefertigt habe.

Erlangen, den 28.10.2014

Andreas Horlbeck