FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
FACHBEREICH WIRTSCHAFTS- WISSENSCHAFTEN

Lehrstuhl für Statistik und Ökonometrie

Diskussionspapier
89 / 2012

Quasi-arithmetische Mittelwerte und Normalverteilung

Ingo Klein

Lange Gasse 20 · D-90403 Nürnberg

Quasi-arithmetische Mittelwerte und Normalverteilung

Ingo Klein

Lehrstuhl für Statistik und Ökonometrie
Universität Erlangen-Nürnberg
ingo.klein@wiso.uni-erlangen.de

Abstract

J.M. Keynes (1911) shows how distributions look like for which the arithmetic, the geometric and the harmonic mean are "most probable values". We propose a general class of distributions for which the quasi-arithmetic means are ML-estimators such that these distributions can be transformed into an normal or a truncated normal distribution. As special cases we get for example the generalized logarithmic distributions introduced by Chen (1995).

Keywords: ML-estimator, quasi-arithmetic mean, exponential family, generalized logarithmic distribution, inverse transformed normal distribution

1. Einleitung

Klein & Grottke (2008) greifen Arbeiten von J.M. Keynes aus dem Jahre (1907), (1908) und (1911) auf, in denen er, in heutige Terminologie übersetzt, untersucht, welche Verteilungsgesetze zu den wesentlichen Mittelwerten (=arithmetisches, harmonisches und geometrisches Mittel und Median) gehören, wenn man diese als ML-Schätzer betrachtet. Wenn man eine Reihe von Präzisierungen der Ergebnisse von Keynes vornimmt, die vor allem den Träger der Verteilungen und die Normalisierungskonstante betreffen, gelangt man zu Verteilungsgesetzen, die zwar zur Exponentialfamilie gehören, aber zunächst noch unspezifizierte Komponenten (=Funktionen) aufweisen. Spezifiziert man diese, so gelangt man in der Tat zu konkreten und wohlbekannten Verteilungen. Wir wollen der Frage nachgehen, ob sich nicht gewisse Setzungen dieser unspezifizierten Funktionen auszeichnen lassen, wenn man nicht nur die von Keynes diskutierten hauptsächlichen Mittelwerte sondern allgemein die Klasse der quasi-arithmetischen Mittelwerte betrachtet, zu denen diese hauptsächlichen Mittelwerte mit Ausnahme des Medians gehören.

Die grundlegende Idee ist, spezielle Verteilungsfamilien in den Mittelpunkt zu rücken, die das arithmetische Mittel als ML-Schätzer für den interessierenden Parameter besitzen. Dann werden via Variablentransformation neue Verteilungsgesetze konstruiert, für die quasi-arithmetische Mittelwerte ML-Schätzer für diesen Parameter sind. Letztlich wird das Vorgehen verallgemeinert, das bei der Konstruktion der logarithmischen Normalverteilung angewandt wird.

Einige der von Keynes vorgeschlagenen Dichten weisen enge Bezüge zu Verteilungen von Zufallsvariablen auf, die sich in eine Normalverteilung bzw. gestutzte Normalverteilung transformieren lassen. Es wird deshalb eine allgemeine Klasse von Verteilungen angegeben, für die quasi-arithmetische Mittelwerte (oder einfache Funktionen davon) ML-Schätzer sind und zusätzlich enge Bezüge zur (gestutzten) Normalverteilung bestehen. Diese Klasse verallgemeinert die logarithmische Verteilung und besitzt als Spezialfall die verallgemeinerte logarithmische Verteilung, wie sie z.B. von Chen (1995) vorgeschlagen wurde.

Die Arbeit gliedert sich wie folgt. Wir spezifizieren in Anlehnung an Klein & Grottke (2008) diese Kriteriumsfunktion derart, dass sich quasi-arithmetische Mittel als ML-Schätzer ergeben. Die zugehörige Dichtefamilie enthält als Spezialfälle die von Keynes angegebenen Formeln für die Dichten, die zu arithmetischem, geometrischem und harmonischem Mittelwert gehören. In dieser allgemeinen Darstellung kommen unspezifizierte Parameter vor. Dies betrifft insbesondere die Normalisierungskonstante. Nachdem die Verteilungsfamilien, für die sich quasi-arithmetische Mittel als ML-Schätzer ergeben recht unspezifiziert sind, betrachten wir konkret solche Verteilungen, die sich per Variablentransformation in eine u.U. gestutzte Normalverteilung transformieren lassen. Diese Klasse ist immer noch hinreichend flexibel, um als Modell für vielfältige Anwendungen zu dienen.

2. Kriteriumsfunktionen und zugehörige Verteilungen

Allgemein wird eine Kriteriumsfunktion $\psi(x;\theta)$ betrachtet, die die Beobachtung x mit dem unbekannten Parameter θ verknüpft. Geht man von stochastisch unabhängigen und mit der Dichte f identisch verteilten Zufallsvariablen $X_1, X_2, \ldots, X_n$ aus, so kann ein Schätzer für θ durch Auflösung der Gleichung

$$\sum_{i=1}^{n} \psi(X_i;\theta) = 0 \tag{1}$$

nach θ bestimmt werden. Bekanntlich führt die Wahl von $\psi(x;\theta) = x - \theta$ zum arithmetischen Mittel und von $\psi(x;\theta) = \mathrm{sign(x-\theta)}$ zum Median von $X_1, X_2, \ldots, X_n$.
Wenn $\partial\psi(x;\theta)/\partial\theta$ unabhängig von x lässt sich nachweisen (siehe Klein & Grottke (2008)), dass die Setzung

$$C'(\theta) = \left(-\frac{\partial\psi(x;\theta)}{\partial\theta}\right)\phi'(\theta)$$

zu

$$f(x;\theta) = K\exp\left(\psi(x;\theta)\frac{C'(\theta)}{-\partial\psi(x;\theta)/\partial\theta} + C(\theta) + b(x)\right) \tag{2}$$

als zu der Kriteriumsfunktion (1) gehörenden Dichte führt. Die Funktion $b(x)$ ist geeignet zu wählen. Von der Normierungskonstanten K ist in Abhängigkeit von ψ, C' und b nachzuweisen, dass sie unabhängig von dem Parameter ist.

3. Quasi-arithmetische Mittelwerte

Sei u eine auf $D \subseteq \mathbb{R}$ f.ü. streng monotone Funktion und legt u die Kriteriumsfunktion entsprechend

$$\psi(x;\theta) = u(x) - u(\theta), \quad x,\theta \in D \tag{3}$$

fest, dann ist $\partial\psi(x;\theta)/\partial\theta = -u'(\theta)$ unabhängig von x. Die zugehörige Dichte lautet

$$f(x;\theta) = K\exp\left((u(x)-u(\theta))\frac{C'(\theta)}{u'(\theta)} + C(\theta) + b(x)\right), \quad x,\theta \in D \tag{4}$$

mit K derart, dass $\int_D f(x;\theta) = 1$ ist.
Diese Dichte ist bezüglich der Transformation u nicht eindeutig festgelegt. Betrachtet man eine affin-lineare Transformation von (x) (z.B. $v(x) = a + bu(x)$), dann gilt die äquivalente Darstellung für die Dichte (4):

$$f(x;\theta) = K\exp\left((v(x)-v(\theta))\frac{C'(\theta)}{v'(\theta)} + C(\theta) + b(x)\right), \quad x,\theta \in D.$$

Hängt K nicht von θ ab, so erhält man mittels der Invarianzeigenschaft von ML-Schätzern (Zehna (1966)) für den ML-Schätzer von θ:

$$\hat{\theta} = u^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} u(X_i)\right), \tag{5}$$

mit u^{-1} als Umkehrfunktion von u auf dem Bildbereich von D. Dieser Schätzer wird quasi-arithmetisches Mittel genannt (siehe z.B. Aczél (1966, S. 276), Bullen et al. (1988, S. 215)). u heißt Generator.

Jede affin-lineare Transformation des Generators u führt zu demselben quasi-arithmetischen Mittel, wie sich leicht zeigen lässt.

Wenn die Konstante K von dem Parameter θ abhängt, erhält man im Allg. nicht mehr das quasi-arithmetische Mittel als ML-Schätzer für θ, wie Klein & Grottke (2008) gezeigt haben.

4. Transformierte Verteilungen

Wenn man weiß, dass das arithmetische Mittel in einer Verteilungsfamilie der ML-Schätzer ist, so lässt sich aus dieser Verteilungsfamilie eine neue Verteilungsfamilie ableiten, für die quasi-arithmetische Mittelwerte ML-Schätzer sind, wie die folgende Überlegung zeigt.

Dichten, für die das arithmetische Mittel ML-Schätzer sind, stellen sich entsprechend

$$f^{**}(x;\theta) = K \exp\left((x-\theta)C'(\theta) + C(\theta) + b(x)\right). \tag{6}$$

dar (siehe Klein & Grottke (2008)). Konstruiert man nun für eine f.ü. differenzierbare, streng monotone Transformation u aus (6) eine neue Dichte

$$f^{*}(x;\theta) = f^{**}(u(x);\theta)|u'(x)|, \tag{7}$$

dann ist $u(X)$ entsprechend (6) verteilt, wenn X mit der Dichte f^* verteilt ist. Wählt man schließlich $u(\theta)$ als neuen Parameter in (7), so gelangt man zu der Dichte

$$f(x;\theta) = f^{*}(x;u(\theta)) = K|u'(x)| \exp\left((u(x)-u(\theta))C'(u(\theta)) + C(u(\theta)) + b(x)\right). \tag{8}$$

Mit $\partial C(u(\theta))/\partial\theta = C'(u(\theta))u'(\theta)$ ist

$$f(x;\theta) = K \exp\left((u(x)-u(\theta))\frac{\partial C(u(\theta))/\partial\theta}{u'(\theta)} + C(u(\theta)) + b(x) + \ln|u'(x)|\right), \tag{9}$$

womit die Dichte der allgemeinen Darstellung (4) genügt, für die das quasi-arithmetische Mittel $u^{-1}(1/n\sum_{i=1}^{n} u(X_i))$ der ML-Schätzer ist.

Es stellt sich die Frage, welche Verteilungen der Form (6) als Ausgangspunkt für die Konstruktion von Verteilungen der Form (9) gewählt werden können. Nahe liegend ist es, die Normalverteilung mit Lageparameter und die Exponentialverteilung mit Skalenparameter zu wählen, da Teicher (1961) gezeigt hat, dass beide Verteilungsfamilien eine zentrale Rolle spielen, wenn es um die Charakterisierung des Verteilungstyps durch die Form des ML-Schätzers geht. Wir konzentrieren uns auf die Normalverteilung als Ausgangsverteilung.

5. Invers u-transformierte Normalverteilung

5.1 Definition und Eigenschaften

Einige der von Keynes vorgeschlagenen Lösungen berücksichtigen die Differenz $u(x) - u(\theta)$ quadratisch im Exponenten. Wählt man dieses Konstruktionsprinzip, so lassen sich allgemein Dichten betrachten, die per Variablentransformation in eine Normalverteilung bzw. eine gestutzte Normalverteilung überführt werden können.

Sei $u : D \subseteq \mathbb{R} \to \mathbb{R}$ eine f.ü. differenzierbare, streng monotone Abbildung. Setzt man in (4)

$$C(\theta) = k^2 u(\theta)^2 \quad \text{und} \quad b(x) = -k^2 u(x)^2 + \ln |u'(x)|,$$

womit $C'(\theta)/u'(\theta) = 2k^2 u(\theta)$ ist, folgt die Dichte

$$\begin{aligned} f(x;\theta) &= K|u'(x)| \exp\left(2k^2(u(x) - u(\theta))u(\theta) + k^2 u(\theta)^2 - k^2 u(x)^2\right) \\ &= K|u'(x)| \exp\left(-k^2(u(x) - u(\theta))^2\right). \end{aligned}$$

Dann bestimmt sich die Konstante durch

$$K \int_D |u'(x)| \exp\left(-k^2(u(x) - u(\theta))^2\right) dx = 1.$$

Setzt man $y = u(x)$, dann ist die Konstante durch

$$K \int_{u(D)} \exp\left(-k^2(y - u(\theta))^2\right) dy = 1$$

gegeben. Wenn $u(D) = \mathbb{R}$ ist, folgt sofort $K = 1/\sqrt{2\pi\sigma^2}$ mit $\sigma^2 = 1/(2k^2)$. Damit ist

$$f(x;\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} |u'(x)| \exp\left(-\frac{1}{2\sigma^2}(u(x) - u(\theta))^2\right), \quad x, \theta \in \mathbb{R}. \tag{10}$$

In diesem Falle kann von einer invers u-transformierten Normalverteilung gesprochen werden. Diese Bezeichung folgt Freeman & Modarres (2006), die von der ,,power normal distribution" als einer Verteilung sprechen, die über eine inverse Box-Cox-Transformation erzeugt wird. Diese Verteilung wird im nächsten Abschnitt noch aufgegriffen, wenn es um die invers u-transformierte gestutzte Normalverteilung geht.

Definition 1 *Sei u eine Transformation. Dann heißt die Zufallsvariable X invers u-transformiert normalverteilt, wenn $u(X)$ normalverteilt ist.*

Seien $X_1, \ldots, X_n$ stochastisch unabhängig, identisch invers u-transformiert normalverteilt. Die invers u-transformierten Normalverteilungen gehören offensichtlich zur Exponentialfamilie, $\sum_{i=1}^n d(X_i)$ ist vollständig suffizient für θ und $1/n \sum_{i=1}^n u(X_i)$ UMVUE für $u(\theta)$. Der ML-Schätzer für θ ist das quasi-arithmetische Stichprobenmittel $u^{-1}(1/n \sum_{i=1}^n u(X_i))$ und für den ML-Schätzer des durch Reparametrisierung in (10) eingeführten Parameters σ^2 gilt

$$\hat{\sigma}^2_{ML} = \frac{1}{n} \sum_{i=1}^n \left(u(X_i) - \frac{1}{n} \sum_{i=1}^n u(X_i) \right)^2 . \tag{11}$$

Weiterhin gelten die Aussagen des Hauptsatzes der Stichprobentheorie, d.h.

1. $1/n \sum_{i=1}^n u(X_i)$ ist normalverteilt mit Mittelwert $u(\theta)$ und Varianz σ^2,
2. $n\hat{\sigma}^2_{ML}/\sigma^2$ ist χ^2-verteilt mit $n-1$ Freiheitsgraden und
3. $1/n \sum_{i=1}^n u(X_i)$ und $\hat{\sigma}^2_{ML}$ sind stochastisch unabhängig.

Wenn $u(\theta)$ streng monoton ist, lassen sich UMP-Tests für θ betreffende Hypothesen mittels $\sum_{i=1}^n u(X_i)$ bekanntem σ^2 durchführen.
Auch die Bayes-Inferenz ist für invers u-transformierte Normalverteilungen analog zur Bayes-Inferenz bei Normalverteilung durchführbar. Bekanntlich ist für den Lokalisationsparameter die zur Normalverteilung konjugierte a priori-Vetrteilung wieder normalverteilt. Ein äquivalentes Ergebnis für invers u-transformierte Normalverteilungen wird im folgenden Satz nachgewiesen.

Satz 1 *Seien die a priori-Verteilung für θ eine u-transformierte Normalverteilung mit den Parametern $u(\xi)$ und τ^2 und die bedingte Verteilung von X gegeben θ ebenfalls eine invers u-transformierte Normalverteilung mit den Parametern $u(\theta)$ und σ^2, dann ist auch die a posteriori-Verteilung von θ gegeben die einfache Zufallsstichprobe $X_1 = x_1, \ldots, X_n = x_n$ als Kopien von X invers u-transformiert normalverteilt mit den Parametern*

$$\frac{\tau^2 \sum_{i=1}^n u(x_i) + \sigma^2 u(\xi)}{n\tau^2 + \sigma^2} \quad \textit{und} \quad \frac{\tau^2\sigma^2}{n\tau^2 + \sigma^2}.$$

D.h die invers u-transformierte Normalverteilung ist konjugiert zu sich selbst.

Den Bayes-Schätzer bei quadratischer Verlustfunktion bzw. den a posteriori-Bayes-Schätzer erhält man als Mittelwert der a posteriori-Verteilung, d.h. als

$$E[\theta|X_1, \ldots, X_n] = \int \theta \frac{1}{\sqrt{2\pi \frac{\tau^2\sigma^2}{n\tau^2+\sigma^2}}} |u'(\theta)| \exp\left(-\frac{1}{2\frac{\tau^2\sigma^2}{n\tau^2+\sigma^2}} \left(u(\theta) - \frac{\tau^2 \sum_{i=1}^n u(x_i) + \sigma^2 u(\xi)}{n\tau^2 + \sigma^2} \right)^2 \right) d\theta.$$

Nach Substitution von $y = u(x)$ ist

$$E[\theta|X_1, \ldots, X_n] = E_\Phi[u^{-1}(\theta)|X_1, \ldots, X_n],$$

wobei Φ eine Normalverteilung mit Mittelwert

$$\frac{\tau^2 \sum_{i=1}^n u(x_i) + \sigma^2 u(\xi)}{n\tau^2 + \sigma^2}$$

und Varianz

$$\frac{\tau^2 \sigma^2}{n\tau^2 + \sigma^2}$$

ist.
Der a posteriori-Bayes-Schätzer ist im Regelfall mit numerischer oder Monte Carlo-Integration zu berechnen.

5.2 Exemplarische Transformationen

Spezielle invers u-transformierte Normalverteilungen sind die Normalverteilung selbst ($u(x) = x$) und die logarithmische Normalverteilung ($u(x) = \ln x$). Für eine solche invers u-transformierte Normalverteilung ist das quasi-arithmetische Mittel der ML-Schätzer für θ.
Weitere Transformationen werden z.B. von Tukey (1957), Box & Cox (1964), Hoyle (1973), Manley (1976), Goto & Inoue (1980), John & Draper (1980), Bickel & Doksum (1981), Sakia (1992), Yeo & Johnson (2000), Tan et al. (2004), Albada & Robinbson (2007) und Hossain (2011) behandelt. Goto & Hamasaki (2002) erweitern speziell die Box-Cox-Transformation auf bivariate Verteilungen. Dieses Vorgehen ist generell auf invers u-transformierte Normalverteilungen übertragbar.
Für invers u-transformierte Verteilungen werden surjektive Transformationen mit Bildbereich $\mathbb{R}$ benötigt.

5.2.1 Zweiseitige Potenz-Transformation

Als Beispiel[1] betrachten wir

$$u(x) = \mathrm{sign(x)|x|}^\lambda, \quad \lambda \neq 0.$$

Für $\lambda > 0$ ist diese Funktion auf ganz $\mathbb{R}$ streng monoton zunehmend und dort invertierbar. Dieser Fall wird von Tan et al. (2004), S. 611 betrachtet.

[1] Zu beachten ist, dass für $\lambda < 0$ u im strengen Sinne nicht Generator eines quasi-arithmetischen Mittelwertes sein kann, da u an der Stelle 0 nicht stetig ist, von einem Generator aber gefordert wird, dass er überall stetig sein soll. Dies wirkt sich aber auf die Möglichkeit der Konstruktion eines ML-Schätzers nebst zugehöriger Verteilungsklasse nicht aus.

Für $\lambda \geq 1$ kommt man sofort zu einer bimodalen invers u-transformierten Normalverteilung (mit $f(0;\theta) = 0$). Für $0 < \lambda < 1$ ergibt sich eine spitzgipflige Verteilung, da in diesem Fall $u'(x) = \lambda|x|^{\lambda-1} \to \infty$ mit $x \to 0$ ist.

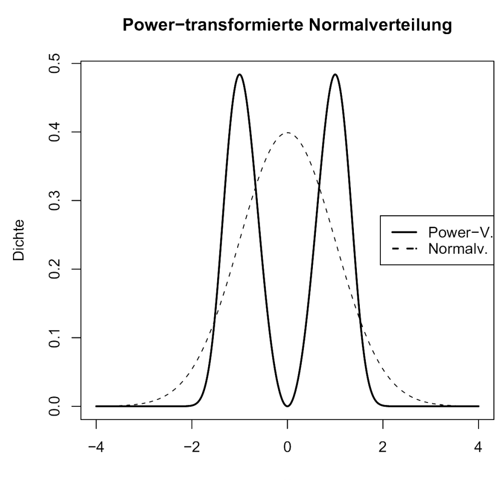
Der ML-Schätzer für θ ist durch

$$\mathrm{sign}\left(\frac{1}{\mathrm{n}}\sum_{\mathrm{i=1}}^{\mathrm{n}}\mathrm{sign}(\mathrm{X_i})|\mathrm{X_i}|^{\lambda}\right)\left(\mathrm{sign}\left(\frac{1}{\mathrm{n}}\sum_{\mathrm{i=1}}^{\mathrm{n}}\mathrm{sign}(\mathrm{X_i})|\mathrm{X_i}|^{\lambda}\right)\frac{1}{\mathrm{n}}\sum_{\mathrm{i=1}}^{\mathrm{n}}\mathrm{sign}(\mathrm{X_i})|\mathrm{X_i}|^{\lambda}\right)^{1/\lambda}.$$

gegeben. Der ML-Schätzer für σ^2 lautet entsprechend (11)

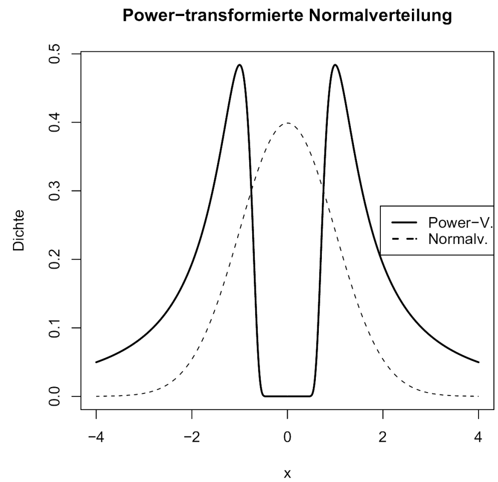
$$\hat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}\left(\mathrm{sign}(\mathrm{X_i})|\mathrm{X_i}|^{\lambda} - \frac{1}{\mathrm{n}}\sum_{\mathrm{i=1}}^{\mathrm{n}}\mathrm{sign}(\mathrm{X_i})|\mathrm{X_i}|^{\lambda}\right)^2.$$

Den typischen Verlauf der zugehörigen Dichte zeigt die folgende Grafik für die Parameter $\sigma^2 = 1$, $\theta = 0$ und $\lambda = 2$:



Die Ränder besitzen deutlich weniger Wahrscheinlichkeitsmasse als die Standardnormalverteilung.

Für $\lambda < 0$ ist die Dichte im Punkte $x = 0$ zwar nicht definiert und kann dort durch die Setzung $f(0;\theta) = 0$ unstetig ergänzt werden. Die Dichte ist wiederum bimodal, wie die folgende Grafik für die Parameter $\sigma^2 = 1$, $\theta = 0$ und $\lambda = -2$ zeigt:



Für $\lambda < 0$ besitzt die invers u-transformierte Normalverteilung in den Rändern deutlich mehr Wahrscheinlichkeitsmasse als die Standardnormalverteilung.
Für den ML-Schätzer für θ gilt

$$\operatorname{sign}\left(\frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\operatorname{sign}(X_i)|X_i|^{|\lambda|}}}\right)\left(\operatorname{sign}\left(\frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\operatorname{sign}(X_i)|X_i|^{|\lambda|}}}\right)\frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\operatorname{sign}(X_i)|X_i|^{|\lambda|}}}\right)^{1/|\lambda|},$$

sofern $X_i \neq 0$, $i = 1, 2, \ldots, n$ und $\frac{1}{n}\sum_{i=1}^{n}\operatorname{sign}(X_i)|X_i|^{\lambda} \neq 0$. Die ML-Schätzung von σ^2 erfolgt analog zum Fall $\lambda > 0$.
Robert (1991) diskutiert eine ebenfalls bimodale verallgemeinerte inverse Normalverteilung, die sich als einfache Erweiterung des Konzeptes der invers u-transformierten Normalverteilungen darstellen lässt. Er definiert

$$f(x;\theta) = K\frac{1}{|y|^{\alpha}}\exp\left(-\frac{1}{2\sigma^2}\left(\frac{1}{y}-\frac{1}{\theta}\right)^2\right)$$

für $\alpha > 1$, $\sigma > 0$. Für $\alpha = 2$ ergibt sich der hier betrachtete Spezialfall der inversen Normalverteilung mit $u(x) = 1/x$, für die das verallgemeinerte harmonische Mittel der Form

$$\frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\operatorname{sign}(X_i)|X_i|}}$$

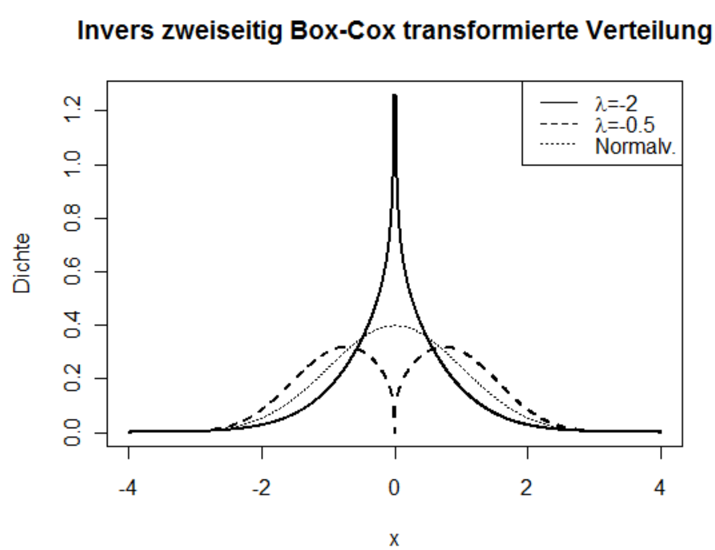
ML-Schätzer für θ ist. Für $\alpha \neq 2$ ist die Konstante K eine komplizierte Funktion von θ, α, σ, die wiederum bei der ML-Schätzung zu berücksichtigen ist, wie sich Robert (1991) entnehmen lässt, der die Konstante explizit berechnet.

5.2.2 Zweiseitige Box-Cox-Transformation

Die zweiseitige Box-Cox-Transformation lautet

$$u(x) = \frac{\mathrm{sign(x)}|\mathrm{x}|^{\lambda} - 1}{\lambda}, \quad x \in \mathbb{R}.$$

Sie ist eine affin-lineare Transformation der bereits diskutierten zweiseitigen Potenz-Transformation. Damit liefern sowohl die zweiseitige Potenz-Transformation als auch die zweiseitige Box-Cox-Transformation dasselbe quasi-arithemtische Mittel als ML-Schätzer für θ.
Die Fälle der zweiseitig invers Box-Cox-transformierten und der zweiseitig Potenz-transformierten Normalverteilung unterscheiden sich nur in dem Auftreten von λ als zusätzlichem Skalenparameter.
Die folgende Grafik illustriert typische Verläufe der zweiseitig Box-Cox transformierten Normalverteilung für $\theta = 0$, $\sigma^2 = 1$ und $\lambda = 2$ bzw. $\lambda = 0.5$.



5.2.3 Modulus Transformation für $\lambda > 0$

John & Draper (1980) diskutieren in Verallgemeinerung der bekannten Box-Cox-Transformation für $\lambda \neq 0$ mit

$$u(x) = \mathrm{sign(x)}\frac{(|\mathrm{x}| + 1)^{\lambda} - 1}{\lambda}, \quad \mathrm{x} \in \mathbb{R}$$

die sog. Modulus Transformation und

$$u'(x) = (|x| + 1)^{\lambda}, \quad x \in \mathbb{R}.$$

Mit $\lambda \to 0$ ergibt sich

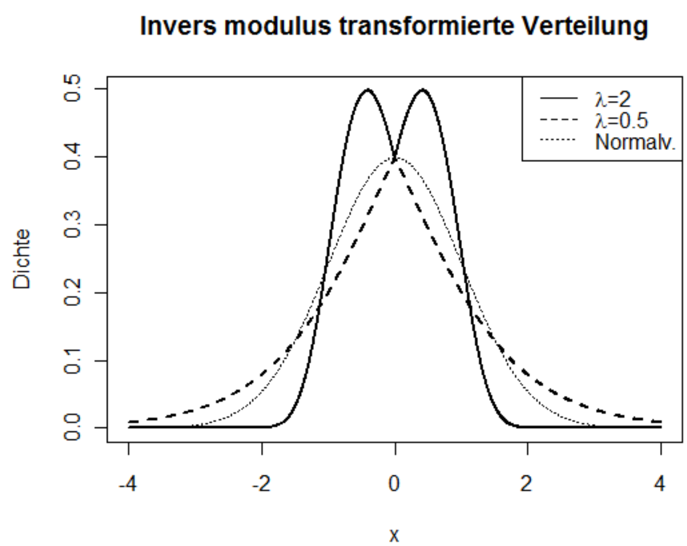
$$u(x) = \mathrm{sign(x)} \log(|\mathrm{x}| + 1), \quad \mathrm{x} \in \mathbb{R}.$$

und

$$u'(x) = \frac{1}{|x| + 1} \quad x \in \mathbb{R}.$$

Diese Familie von Transformation besitzt nur für $\lambda > 0$ ganz $\mathbb{R}$ als Bildbereich. Für $\lambda > 1$ erhält man eine bimodale Verteilung, für $\lambda = 1$ offenbar die Normalverteilung und für $\lambda \in (0, 1)$ eine spitzgipflige Verteilung, die der Laplace-Verteilung ähnelt. Diese Transformation ist nicht geeignet, um Schiefe in einer Verteilung zu korrigieren. Die zugehörige invers u-transformierte Verteilung ist stets wieder symmetrisch.

Die folgende Grafik vermittelt einen Eindruck der invers u-transformierten Dichtefunktionen für die Parametersetzung $\theta = 0$, $\sigma^2 = 1$ und $\lambda = 2$ bzw. $\lambda = 0.5$.



5.2.4 Transformation nach Yeo & Johnson (2000)

Yeo & Johnson betrachten ebenfalls eine auf $\mathbb{R}$ bijektive Transformation

$$u(x) = \begin{cases} \frac{(x+1)^\lambda - 1}{\lambda} & \text{für } x \geq 0 \text{ und } \lambda \neq 0 \\ \frac{(1-x)^{2-\lambda} - 1}{\lambda - 2} & \text{für } x < 0 \text{ und } \lambda \neq 2 \end{cases}$$

und

$$u'(x) = \begin{cases} (x+1)^{\lambda - 1} & \text{für } x \geq 0 \text{ und } \lambda \neq 0 \\ (1-x)^{1-\lambda} & \text{für } x < 0 \text{ und } \lambda \neq 2 \end{cases}$$

Mit $\lambda \to 0$ ergibt sich

$$u(x) = \ln(1+x), \quad x \geq 0$$

bzw. für $\lambda \to 2$

$$u(x) = -\ln(1-x) \;\; x < 0,$$

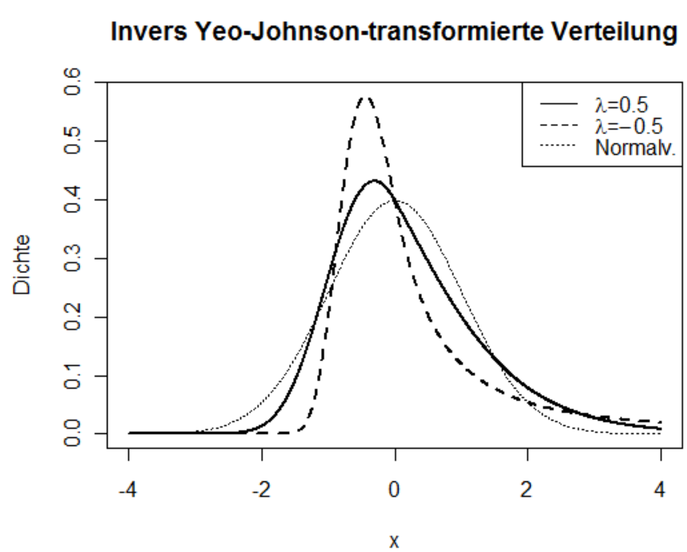
so dass

$$u'(x) = 1/(1+x), \quad x \geq 0 \;\text{ und }\; u'(x) = 1/(1-x)\;; x < 0.$$

Die Transformation verfügt über zahlreiche wünschenswerte Eigenschaften, wie z.B. die Stetigkeit an der Stelle 0. Sie ist so konstruiert, dass sie auch Schiefe in der Ausgangsverteilung korrigieren kann, d.h. die zugehörige invers u-transformierte Verteilung ist schief.

$$f(x;\theta,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \begin{cases} (1+x)^{\lambda-1} \exp\left(-\frac{1}{2\sigma^2}\left(\frac{(x+1)^\lambda-1}{\lambda} - \frac{(\theta+1)^\lambda-1}{\lambda}\right)^2\right) & \text{für } x \geq 0 \\ (1-x)^{1-\lambda} \exp\left(-\frac{1}{2\sigma^2}\left(\frac{(1-x)^{2-\lambda}-1}{\lambda-2} - \frac{(1-\theta)^{2-\lambda}-1}{2-\lambda}\right)^2\right) & \text{für } x < 0 \end{cases}$$

Für $\theta = 0$, $\sigma^2 = 1$ und $\lambda = 0.5$ bzw. $\lambda = -0.5$ zeigt die folgende Grafik die Verläufe der invers u-transformierten Dichte.



5.2.5 sinh-Transformation

Manley (1976) betrachtet ebenfalls zwei auf ganz $\mathbb{R}$ surjektive Transformationen

$$u(x) = \sinh(\lambda x) \;\; x \in \mathbb{R}, \; \lambda \neq 0$$

bzw.

$$u(x) = \operatorname{asinh}(\lambda x) = \log(\lambda x + \sqrt{1 + (\lambda x)^2}) \;\; x \in \mathbb{R},\; \lambda \neq 0.$$

asinh bezeichnet die inverse Sinushyperbolicusfunktion. Mit

$$d\sinh(x)/dx = \cosh(x) \quad \text{und} \quad d\operatorname{asinh}(x)/dx = 1/\sqrt{1+x^2}$$

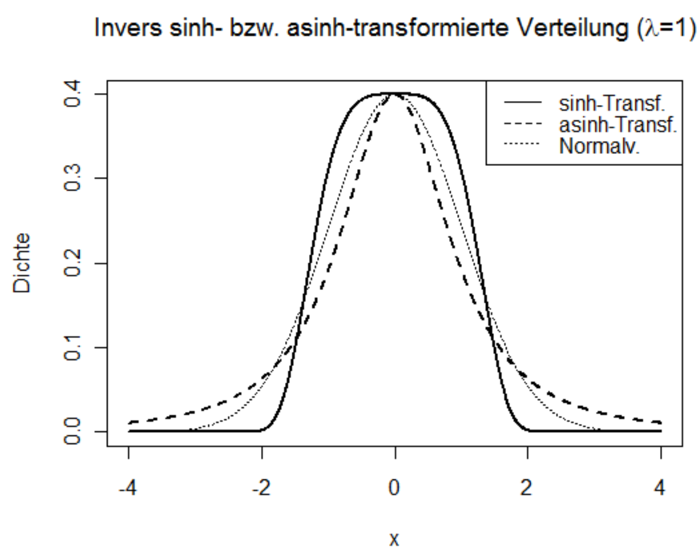
sind die zugehörigen invers u-transformierten Normalverteilungen für $\lambda = 1$:

$$f(x;\theta,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \cosh(x) e^{-1/2(\sinh(x)-\sinh(\theta))^2/\sigma^2}$$

bzw.

$$f(x;\theta,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{\sqrt{1+x^2}} e^{-1/2(\operatorname{asinh}(x)-\operatorname{asinh}(\theta))^2/\sigma^2}.$$

Ähnlich wie bei der Modulus-Transformation können auch diese Transformationen keinerlei Schiefe korrigieren. Die zugehörigen invers u-transformierten Normalverteilungen sind entweder platykurtisch (sinh-Transformation) oder leptokurtisch (asinh-Transformation). Typische Verläufe zeigt die folgende Grafik für die Parameter $\theta = 0$, $\sigma^2 = 1$, $\lambda = 1$:



6. Invers u-transformierte gestutzte Normalverteilung

6.1 Definition und Eigenschaften

Ist der Bildbereich $u(D) = (a, b)$ ein Intervall der reellen Zahlen, so dass $a \neq -\infty$ oder $b \neq \infty$ sind, so hängt die Konstante K von θ ab. Wir betrachten den Fall eines beidseitig endlichen

Intervalls (a, b). Ist u streng monoton zunehmend, so ist

$$K \int_a^b \exp\left(-k^2(y - u(\theta))^2\right) dy = 1,$$

womit

$$\frac{1}{K} = \left(\Phi\left(\frac{b - u(\theta)}{\sigma}\right) - \Phi\left(\frac{a - u(\theta)}{\sigma}\right)\right) \frac{1}{\sqrt{2\pi\sigma^2}}$$

gilt. Es wird wiederum $\sigma^2 = 1/(2k^2)$ gesetzt. Somit erhält man für die Dichte von $y = u(x)$ eine beidseitig gestutzte Normalverteilung. In diesem Fall kann in Anlehnung an die u-transformierte Normalverteilung X u-gestutzt normalverteilt genannt werden.

Definition 2 *Sei u eine Transformation. Dann heißt die Zufallsvariable X u-transformiert gestutzt normalverteilt, wenn $u(X)$ gestutzt normalverteilt ist.*

Die zugehörige Dichte der invers u-transformierten gestutzten Normalverteilung ist

$$f(x;\theta) = \frac{1}{\Phi\left(\frac{b-u(\theta)}{\sigma}\right) - \Phi\left(\frac{a-u(\theta)}{\sigma}\right)} \frac{1}{\sqrt{2\pi\sigma^2}} |u'(x)| \exp\left(-\frac{1}{2\sigma^2}\left(u(x) - u(\theta)\right)^2\right) \tag{12}$$

für $a < u(x) < b$ und $x \in D \subseteq \mathbb{R}$. Φ bezeichnet die Verteilungsfunktion der Standardnormalverteilung.

Der ML-Schätzer für θ löst die Gleichung

$$\frac{1}{n}\sum_{i=1}^{n} u(X_i) = u(\theta) + \sigma \frac{\phi\left((b - u(\theta))/\sigma\right) - \phi\left((a - u(\theta))/\sigma\right)}{\Phi\left((b - u(\theta))/\sigma\right) - \Phi\left((a - u(\theta))/\sigma\right)} \equiv \tau(\theta).$$

Dasselbe Ergebnis stellt sich ein, wenn u streng monoton abnehmend mit Bildbereich $D(a, b)$ ist. In diesem Fall ist $dy = -|u'(x)|dx$ und die Bedingung für die Konstante K lautet ebenfalls

$$K\left(-\int_b^a \exp\left(-k^2(y - u(\theta))^2\right)\right) dy = K \int_a^b \exp\left(-k^2(y - u(\theta))^2\right) dy = 1.$$

D.h. K hängt sowohl von θ, k^2 als auch von weiteren Parametern ab, die z.B. die Transformation u spezifizieren. Reparametrisiert man $k^2 = 1/(2\sigma^2)$, so ist der ML-Schätzer von σ^2 im Gegensatz zum nicht-gestutzten Fall nicht mehr explizit angebbar; er muss iterativ numerisch bestimmt werden.

Wenn a hinreichend klein bzw. b hinreichend groß sind, so dass $\phi((a - u(\theta))/\sigma) \approx \phi((b - u(\theta))/\sigma) \approx 0$ und $\Phi((a - u(\theta))/\sigma) \approx 0$ bzw. $\Phi((b - u(\theta))/\sigma) \approx 1$ sind, ergibt sich zumindest näherungsweise wiederum das quasi-arithmetische Mittel als ML-Schätzer für θ.

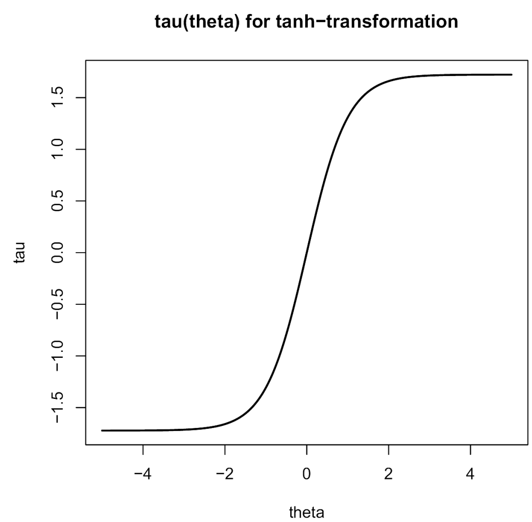
6.2 Spezielle Transformationen

6.2.1 tanh-Transformation

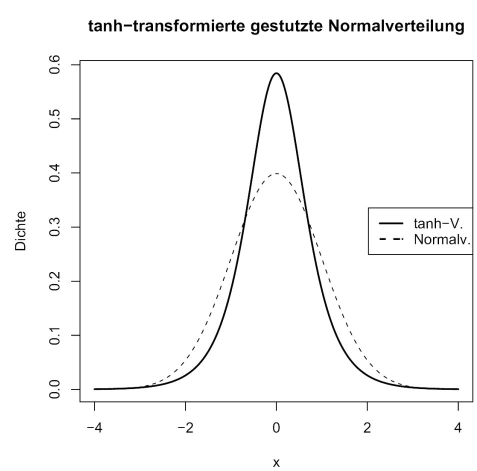
Zu einer zweiseitig gestutzten u-Normalverteilung gelangt man mit der Transformation

$$u(x) = \lambda \tanh(x), \quad x \in \mathbb{R}$$

mit dem Bildbereich $(-\lambda, \lambda)$. Für diese Funktion stellt sich $\tau(\theta)$ grafisch wie folgt dar:



τ ist streng monoton und numerisch invertierbar. Die zugehörige Dichte ist deutlich spitzgipfliger als die Standardnormalverteilung, wie die folgende Grafik für die Parametersetzung $\theta = 0$, $\sigma^2 = 1$ und $\lambda = 1.75$ zeigt:



Die Beschränktheit der Transformation bewirkt in den Rändern eine deutlich geringere Wahrscheinlichkeitsmasse als die der Standardnormalverteilung.
Der ML-Schätzer für θ ist durch Lösung der Gleichung

$$\frac{1}{n}\sum_{i=1}^{n}\tanh(X_i) = \tanh(\theta) + \frac{\sigma}{\lambda}\frac{\phi\left((1-\tanh(\theta))/(\sigma/\lambda)\right) - \phi\left((-1-\tanh(\theta))/(\sigma/\lambda)\right)}{\Phi\left((1-\tanh(\theta))/(\sigma/\lambda)\right) - \Phi\left((-1-\tanh(\theta))/(\sigma/\lambda)\right)} \equiv \tau(\theta)$$

bestimmt. Nur wenn σ klein bzw. λ groß sind, ist zu erwarten, dass sich näherungsweise ein quasi-arithmetisches Mittel als ML-Schätzer ergibt.

6.2.2 Box-Cox-Transformation

Ein Beispiel für eine einseitig gestutzte u- (gestutzte) Normalverteilung liefert die verallgemeinerte logarithmische Normalverteilung, die von der Box-Cox-Transformation (siehe Box & Cox (1964), (1982)) ausgeht und von Chen (1995) eingeführt wurde. D.h. es wird für $x > 0$ und $\lambda \in \mathbb{R}$

$$u(x) = \frac{x^\lambda - 1}{\lambda}$$

betrachtet, wobei $u(x) = \ln x$ für $\lambda \to 0$ ist. Bis auf diesen Grenzfall ist der Wertebereich der Box-Cox-Transformation einseitig beschränkt. Für $\lambda > 0$ ist $u(x) > -1/\lambda$ und für $\lambda < 0$ ist $u(x) < -1/\lambda$. Die Box-Cox-Transformation ist für alle λ streng monoton zunehmend mit $u'(x) = x^{\lambda-1}$ für $x > 0$. Die verallgemeinerte logarithmische Normalverteilung von Chen lautet dann

$$f(x;\theta) = \begin{cases} \frac{1}{1-\Phi\left(-\frac{\theta^\lambda/\lambda}{\sigma}\right)}\frac{1}{\sqrt{2\pi\sigma^2}}x^{\lambda-1}\exp\left(-\frac{1}{2\sigma^2}\left(\frac{x^\lambda}{\lambda}-\frac{\theta^\lambda}{\lambda}\right)^2\right) & \text{für } \lambda > 0 \\ \frac{1}{\sqrt{2\pi\sigma^2}x}\exp\left(-\frac{1}{2\sigma^2}\left(\ln x - \ln\theta\right)^2\right) & \text{für } \lambda = 0 \\ \frac{1}{\Phi\left(\frac{-\theta^\lambda/\lambda}{\sigma}\right)}\frac{1}{\sqrt{2\pi\sigma^2}}x^{\lambda-1}\exp\left(-\frac{1}{2\sigma^2}\left(\frac{x^\lambda}{\lambda}-\frac{\theta^\lambda}{\lambda}\right)^2\right) & \text{für } \lambda < 0 \end{cases}$$

Die verallgemeinerte logarithmische Normalverteilung ist also für $\lambda \neq 0$ bezüglich der Box-Cox-Transformation gestutzt normalverteilt. Der ML-Schätzer für θ erfüllt die notwendige Bedingung

$$\frac{1}{n}\sum_{i=1}^{n}\frac{X_i^\lambda - 1}{\lambda} = \begin{cases} \frac{\theta^\lambda-1}{\lambda} + \sigma\frac{-\phi\left(-\theta^\lambda/(\lambda\sigma)\right)}{1-\Phi\left(\theta^\lambda/(\lambda\sigma)\right)} & \text{für } \lambda > 0 \\ \ln\theta & \text{für } \lambda = 0 \\ \frac{\theta^\lambda-1}{\lambda} + \sigma\frac{\phi\left(-\theta^\lambda/(\lambda\sigma)\right)}{\Phi\left(\theta^\lambda/(\lambda\sigma)\right)} & \text{für } \lambda < 0 \end{cases}$$

Chen (1995) geht approximativ von $\Phi\left(-\theta^\lambda/(\lambda\sigma)\right) = 0$ für $\lambda > 0$ und $\Phi\left(-\theta^\lambda/(\lambda\sigma)\right) = 1$ für $\lambda < 0$ aus, womit sich der ML-Schätzer von θ zu

$$\left(\frac{1}{n}\sum_{i=1}^{n} X_i^\lambda\right)^{1/\lambda}$$

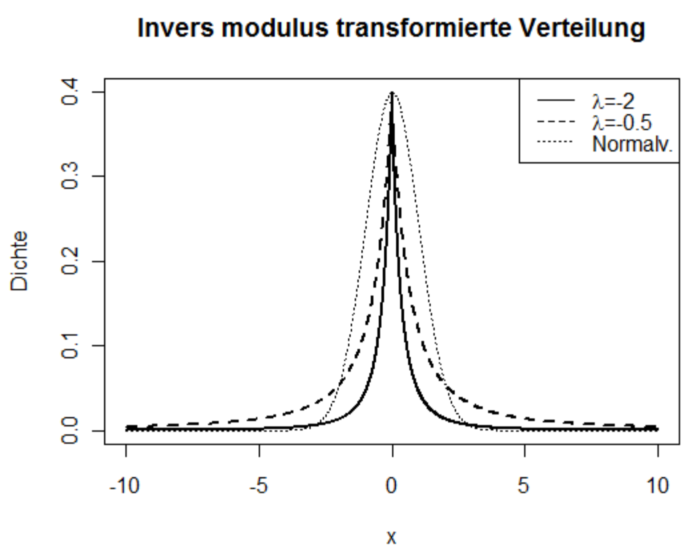
vereinfacht. D.h. er ist zumindest approximativ durch den Potenzmittelwert gegeben.

6.2.3 Modulus Transformation für $\lambda < 0$

Für $\lambda < 0$ besitzt die Modulus Transformation

$$u(x) = \mathrm{sign(x)}\frac{(|\mathrm{x}|+1)^\lambda - 1}{\lambda}, \quad \mathrm{x} \in \mathbb{R}$$

als Bildbereich das Intervall $[-1/\lambda, 1/\lambda]$. D.h. die zugehörige invers u-transformierte Verteilung lässt sich in eine symmetrisch zweiseitig gestutzte Normalverteilung transformieren. Die invers u-transformierte Normalverteilung ist in diesem Falle extrem spitzgipflig und verfügt über schwere Tails. Einen Eindruck verschafft die folgende Grafik für die Parameter $\theta = 0$, $\sigma^2 = 1$ und $\lambda = -2$ bzw. $\lambda = -0.5$.



Der ML-Schätzer für θ bestimmt sich wiederum nur näherungsweise als quasi-arithmetisches Mittel.

7. Zusammenfassung

Es wird gezeigt, dass sich das Keynessche Vorgehen der Lösung einer einfachen Differentialgleichung zur Bestimmung der Dichte für vorgebenene ML-Schätzer auf die Klasse der quasiarithmetischen Mittel anwenden lässt. Ausgehend von einer Verteilung, für die das arithmetische Mittel der ML-Schätzer (z.B. die Normalverteilung oder die Exponentialverteilung) ist, lassen sich per Variablentransformation leicht Verteilungen konstruieren, für die beliebige quasiarithmetische Mittel ML-Schätzer sind. Zu denen gehören beispielsweise transformierte (gestutzte) Normalverteilungen, mit der verallgemeinerten logarithmischen Normalverteilung für die Box-Cox-Transformation als Spezialfall.

8. Anhang: Beweis von Satz 1

Beweis: Es ist

$$\begin{aligned} L(\theta; x_1, \ldots, x_n) \cdot f_\theta(\theta) = {} & \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} |u'(x)| \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(u(x_i) - u(\theta))^2\right) \\ & \cdot \frac{1}{\sqrt{2\pi\tau^2}} |u'(\theta)| \exp\left(-\frac{1}{2\tau^2}(u(\theta) - u(\xi))^2\right). \end{aligned}$$

Mit

$$\begin{aligned} & -\frac{1}{2\sigma^2}\sum_{i=1}^{n}(u(x_i) - u(\theta))^2 - \frac{1}{2\tau^2}(u(\theta) - u(\xi))^2 \\ & = -\frac{1}{2}\left(u(\theta)^2 \frac{n\tau^2 + \sigma^2}{\tau^2\sigma^2} - 2u(\theta)\frac{\tau^2 \sum_{i=1}^{n} x_i + \sigma^2 u(\xi)}{\tau^2\sigma^2}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n} u(x_i)^2 \\ = & -\frac{1}{2\frac{\tau^2\sigma^2}{n\tau^2+\sigma^2}}\left(u(\theta) - \frac{\tau^2 \sum_{i=1}^{n} u(x_i) + \sigma^2 u(\xi)}{n\tau^2 + \sigma^2}\right)^2 \\ & + \frac{1}{2\frac{\tau^2\sigma^2}{n\tau^2+\sigma^2}}\left(\frac{\tau^2 \sum_{i=1}^{n} u(x_i) + \sigma^2 u(\xi)}{n\tau^2 + \sigma^2}\right)^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n} u(x_i)^2, \end{aligned}$$

wobei im letzten Schritt quadratisch ergänzt wurde, erhält man

$$L(\theta; x_1, \ldots, x_n) \cdot f_\theta(\theta) = |u'(\theta)| \exp\left(-\frac{1}{2\frac{\tau^2\sigma^2}{n\tau^2+\sigma^2}}\left(u(\theta) - \frac{\tau^2 \sum_{i=1}^{n} u(x_i) + \sigma^2 u(\xi)}{n\tau^2 + \sigma^2}\right)^2\right) + \text{Rest}$$

mit einem Rest

$$\text{Rest} = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} |u'(x)| \exp\left(\frac{1}{\sqrt{2\pi\tau^2}}\frac{1}{2\frac{\tau^2\sigma^2}{n\tau^2+\sigma^2}}\left(\frac{\tau^2 \sum_{i=1}^{n} u(x_i) + \sigma^2 u(\xi)}{n\tau^2 + \sigma^2}\right)^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n} u(x_i)^2\right),$$

der nicht von θ abhängt und sich bei der Berechnung der a posteriori-Dichte herauskürzt. D.h. für die a posteriori-Dichte von θ gilt:

$$\begin{aligned} f_{\theta|X_1,\dots,X_n}(\theta|x_1,\dots,x_n) &= \frac{L(\theta;x_1,\dots,x_n)\cdot f_\theta(\theta)}{\int L(\theta;x_1,\dots,x_n)\cdot f_\theta(\theta)d\theta} \\ &= \frac{|u'(\theta)|\exp\left(-\frac{1}{2\frac{\tau^2\sigma^2}{n\tau^2+\sigma^2}}\left(u(\theta)-\frac{\tau^2\sum_{i=1}^n u(x_i)+\sigma^2 u(\xi)}{n\tau^2+\sigma^2}\right)^2\right)}{\int |u'(\theta)|\exp\left(-\frac{1}{2\frac{\tau^2\sigma^2}{n\tau^2+\sigma^2}}\left(u(\theta)-\frac{\tau^2\sum_{i=1}^n u(x_i)+\sigma^2 u(\xi)}{n\tau^2+\sigma^2}\right)^2\right)d\theta} \end{aligned}$$

Da für den Nenner offensichtlich

$$\sqrt{2\pi\frac{\tau^2\sigma^2}{n\tau^2+\sigma^2}}$$

gilt, ergibt sich

$$f_{\theta|X_1,\dots,X_n}(\theta|x_1,\dots,x_n) = \frac{1}{\sqrt{2\pi\frac{\tau^2\sigma^2}{n\tau^2+\sigma^2}}}|u'(\theta)|\exp\left(-\frac{1}{2\frac{\tau^2\sigma^2}{n\tau^2+\sigma^2}}\left(u(\theta)-\frac{\tau^2\sum_{i=1}^n u(x_i)+\sigma^2 u(\xi)}{n\tau^2+\sigma^2}\right)^2\right).$$

Dies ist die Dichte einer u-transformierten Normalverteilung mit Mittelwert

$$\frac{\tau^2\sum_{i=1}^n u(x_i)+\sigma^2 u(\xi)}{n\tau^2+\sigma^2}$$

und Varianz

$$\frac{\tau^2\sigma^2}{n\tau^2+\sigma^2} \quad \square.$$

Literatur

1. Aczél, J. (1966). *Lectures on functional equations and their applications.* Academic Press, New York.

2. Albada, S.J. & Robinson, P.A. (2007). Transformation of arbitrary distributions to the normal distribution with application to EEG test–retest reliability. *Journal of Neuroscience Methods* **161**, 207-211.

3. Bickel, P.J. & Doksum, K.A. (1981). An analysis of transformations revisited. *Journal of the American Statistical Association* **76**, 296-311.

4. Box, G.E.P. & Cox, D.R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, Series B* **26**, 211-252.

5. Box, G.E.P. & Cox, D.R. (1982). An analysis of transformations revisited, rebutted. *Journal of the American Statistical Association* **77**, 209-210.

6. Bullen, P.S., Mitrinović, D.S., Vasić, P.M. (1988). *Means and their inequalities.* D. Reidel Publ. Co., Dordrecht.

7. Chen, G. (1995). Generalized log-normal distributions with reliability application. *Computational Statistics & Data Analysis* **19**, 309-319.

8. Freeman, J. & Modarres, R. (2006). Inverse Box-Cox: The power-normal distribution. *Statistics and Probability Letters* **76**, 764-772.

9. Goto, M. & Hamasaki, T. (2002). The bivariate power-normal distribution. *Bulletin of Informatics and Cybernetics* **34**, 29-49.

10. Goto, M. & Inoue, T. (1980). Some properties of the power normal Distribution. *Bulletin Biometric Society of Japan* **1**, 28-54

11. Hossain, M.Z. (2011). The use of Box-Cox transformation technique in economic and statistical analysis. *Journal of Emerging Trends in Economic and Management Science* **2**, 32-39.

12. Hoyle, M.H. (1973). Transformations: An introduction and a bibliography. *International Statistical Review* **41**, 203-223.

13. John , J.A. & Draper, N.R. (1980). An alternative family of transformations. *Applied Statistics* bf 29, 190-197.

14. Keynes, J.M. (1907). *The principles of probability.* Submitted as a fellowship dissertation to King's College Cambridge. Keynes Papers, King's College TP/A/1-2.

15. Keynes, J.M. (1908). *The principles of probability.* Submitted as a fellowship dissertation to King's College Cambridge. Keynes Papers, King's College TP/A/1-2.

16. Keynes, J.M. (1911). The principal averages and the laws of error which lead to them. *Journal of the Royal Statistical Society.* **74**, 322-331.

17. Klein, I., Grottke, M. (2008). On J.M. Keynes' "The Principal Averages and the Laws of Error which Lead to Them" - Refinement and Generalisation. IWQW-Discussion Papers, University of Erlangen-Nuremberg 7/2008.

18. Manley, B.F. (1976). Exponential data tranformation. *The Statistician* **25**, 37-42.

19. Robert, C. (1991). Generalized inverse normal distributions. *Statistics & Probability Letters* **11**, 37-41.

20. Sakia, R.M. (1992). The Box-Cox tranformaiuton technique: a review. *The Statistician* **41**, 169-178.

21. Tan, W.D., Gan, F.F. & Chang, T.C. (2004). Using normal quantile plot to select an appropriate transformation to achieve normality. *Computational Statistics & Data Analysis* **45**, 609-619.

22. Teicher, H. (1961). Maximum likelihood characterization of distributions. *Annals of Mathematical Statistics* **32**, 1214-1222.

23. Tukey, J.W. (1957). The comparative anatomy of transformations. *Annals of Mathematical Statistics* **28**, 602-632.

24. Yeo, I.-K. & Johnson, R. (2000). A new family of power transformations to improve normality or symmetry. *Biometrika* **87**, 954-959.

25. Zehna, P.W. (1966). Invariance of maximum likelihood estimation. *Annals of Mathematical Statistics* **37**, 755.