

Parameterschätzung bei klassierten Daten

Ein Vergleich konkurrierender Schätzmethoden

S. Niermann
M. D. Jöhnk*
Diskussionspapier 218
ISSN 0949-9962

*Adresse der Autoren:
Institut für Quantitative Wirtschaftsforschung
Königsworther Platz 1
30167 Hannover
Fax-Nr. 0049 511 762-3923

1 Einleitung

Bei der empirischen Arbeit ergibt sich häufig das Problem, daß die Daten lediglich in der Form einer Häufigkeitstabelle vorliegen. Wenn man an der Schätzung von Parametern interessiert ist, können dann modifizierte Maximum-Likelihood-Verfahren oder der Minimum-Quantil-Abstands-Schätzer angewandt werden. Diese beiden Methoden werden ausführlich bei MCDONALD und RANSOM (1979) beschrieben. Eine weitere Möglichkeit der Parameterschätzung bieten Quantilschätzer, die als Spezialfall des Minimum-Quantils-Abstands-Schätzers aufgefaßt werden können (JOHNSON, KOTZ und BALAKRISHNAN (1994)).

Schon für den Fall, daß nur zwei Parameter zu schätzen sind, sind diese Verfahren häufig technisch sehr aufwendig. Als Konsequenz daraus werden sie auch in der Praxis nur selten angewandt.

In der vorliegenden Arbeit wird ein neues Schätzverfahren vorgestellt, welches auf die meisten Verteilungen mit zwei Parametern angewandt werden kann. Dieses Verfahren ist leicht anzuwenden und hat gute theoretische Eigenschaften. Um die Eigenschaften dieses Verfahrens mit denen der bekannten Verfahren bei endlichen Stichproben zu vergleichen, werden Simulationen durchgeführt.

2 Schätzmethoden

Bei gegebenen Klassengrenzen $(\tilde{x}_0, \tilde{x}_1, \dots, \tilde{x}_k)$ handelt es sich bei den Besetzungszahlen n_i und den davon abgeleiteten Größen um Zufallsgrößen. Die relativen Häufigkeiten werden mit \hat{p}_i und die kumulierten relativen Häufigkeiten mit \hat{F}_i notiert.

Die Klassengrenzen können als empirische Quantile interpretiert werden. Dabei muß aber beachtet werden, daß nicht die Quantile $\hat{Q}(w)$, sondern die jeweils zugehörigen w bei dieser Interpretation stichprobenabhängig sind.

Es werden verschiedene Schätzmethoden untersucht und verglichen:

- a) Maximum-Likelihood-Methode

Ansatz:

Wähle Parametervektor $\boldsymbol{\theta}$ so, daß

$$L(\boldsymbol{\theta}) = \prod p_i(\boldsymbol{\theta})^{n_i} \quad (1)$$

maximiert wird.

- b) Chi-Quadrat-Minimum-Methode

Ansatz:

Wähle Parametervektor $\boldsymbol{\theta}$ so, daß

$$X^2 = \sum \frac{[n_i - np_i(\boldsymbol{\theta})]^2}{np_i(\boldsymbol{\theta})} \quad (2)$$

minimiert wird.

c) Optimale Anpassung der Verteilungsfunktion

Wähle Parametervektor θ so, daß

$$S_1^2 = \sum_i \sum_j [\hat{F}_i \Leftrightarrow F_i(\boldsymbol{\theta})][\hat{F}_j \Leftrightarrow F_j(\boldsymbol{\theta})]w_{ij} = [\hat{\mathbf{F}} \Leftrightarrow \mathbf{F}(\boldsymbol{\theta})]' \mathbf{W}_1 [\hat{\mathbf{F}} \Leftrightarrow \mathbf{F}(\boldsymbol{\theta})] \quad (3)$$

mit noch näher zu bestimmender Gewichtungsmatrix \mathbf{W}_1 minimiert wird.

d) Linearisierung der Verteilungsfunktion

Die Methode c) entspricht der Anpassung einer nichtlinearen Regressionsfunktion an die empirische Verteilungsfunktion. Diese Regressionsfunktion kann häufig linearisiert werden. Damit können dann die Ergebnisse der linearen Regressionsanalyse auf diesen Fall übertragen werden.

Die Linearisierung wird unter Verwendung geeigneter Transformationsfunktionen $g(\cdot)$ und $h(\cdot)$ durchgeführt. Diese ermöglichen eine Darstellung der Verteilungsfunktion in der Form

$$h(F(x)) = \beta_0 + \beta_1 \cdot g(x).$$

Die Idee dieses Verfahrens besteht nun darin, ein Streudiagramm aus den $g(\tilde{x}_i)$ und den $h(\hat{F}_i)$ zu erstellen. Die Punkte in diesem Streudiagramm sollten um eine Gerade streuen, sofern das angenommene Modell den Daten angemessen ist. Die Schätzwerte der Regressionsparameter werden nun verwendet, um die Modellparameter zu schätzen. Zusätzlich kann mit diesem Verfahren natürlich die Modellvalidität überprüft werden.

Folglich werden die Parameter β_0 und β_1 so gewählt, daß

$$\begin{aligned} S_2^2 &= \sum_i \sum_j [h(\hat{F}_i) \Leftrightarrow \beta_0 \Leftrightarrow \beta_1 \cdot g(x_i)]w_{ij}^2 [h(\hat{F}_j) \Leftrightarrow \beta_0 \Leftrightarrow \beta_1 \cdot g(x_j)] \\ &= [h(\hat{\mathbf{F}}) \Leftrightarrow h(\mathbf{F}(\beta_0, \beta_1))]'\mathbf{W}_2 [h(\hat{\mathbf{F}}) \Leftrightarrow h(\mathbf{F}(\beta_0, \beta_1))] \end{aligned} \quad (4)$$

minimiert wird, wobei \mathbf{W}_2 eine Gewichtungsmatrix ist.

Es folgen einige Beispiele für die Linearisierung der Verteilungsfunktion.

(i) Die Normalverteilung

Bei der Normalverteilung gilt

$$F(x) = \Phi\left(\frac{x \Leftrightarrow \mu}{\sigma}\right)$$

und

$$\Phi^{-1}(F(x)) = \Leftrightarrow \frac{\mu}{\sigma} + \frac{1}{\sigma}x$$

Hier ist also $h(y)$ die Umkehrfunktion der Verteilungsfunktion der Standardnormalverteilung und $g(x) = x$. Zu beachten ist, daß die Parameter μ und σ der Normalverteilung sich nicht unmittelbar als Regressionskoeffizienten ergeben. Vielmehr ist $\beta_0 = \Leftrightarrow \mu/\sigma$ und $\beta_1 = 1/\sigma$. Daraus folgt $\hat{\sigma} = 1/\hat{\beta}_1$ und $\hat{\mu} = \Leftrightarrow \hat{\beta}_0/\hat{\beta}_1$. Auf dieser Linearisierung beruht auch das Normalverteilungspapier.

- (ii) Ähnlich wie bei der Normalverteilung ist die Situation bei vielen anderen Verteilungen, die lediglich linear parametrisiert sind. In allen Fällen gilt $g(x) = x$. Genannt seien:

- (iia) Die Rechteckverteilung:

$$F(x) = \frac{x \Leftrightarrow a}{b \Leftrightarrow a} = \Leftrightarrow \frac{a}{b \Leftrightarrow a} + \frac{1}{b \Leftrightarrow a} x$$

Hier ist die Verteilungsfunktion bereits linear in den Koeffizienten. Es gilt daher $h(y) = y$.

- (iib) Die zweiparametrische Exponentialverteilung:

$$F(x) = 1 \Leftrightarrow \exp\{\Leftrightarrow \lambda(x \Leftrightarrow \xi)\}$$

$$\ln(1 \Leftrightarrow F(x)) = \lambda \xi \Leftrightarrow \lambda x$$

Es ist also $h(y) = \ln(1 \Leftrightarrow y)$.

- (iic) Die logistische Verteilung:

$$F(x) = \frac{\exp\left(\frac{x-m}{d}\right)}{1 + \exp\left(\frac{x-m}{d}\right)}$$

$$\ln\left(\frac{F(x)}{1 \Leftrightarrow F(x)}\right) = \Leftrightarrow \frac{m}{d} + \frac{1}{d} x$$

Hier ist also $h(y) = \ln \frac{y}{1-y}$.

- (iid) Die Gumbel-Verteilung (Extremwertverteilung):

$$F(x) = 1 \Leftrightarrow \exp(\Leftrightarrow \exp\left(\frac{x \Leftrightarrow m}{d}\right))$$

$$\ln(\Leftrightarrow \ln(1 \Leftrightarrow F(x))) = \Leftrightarrow \frac{m}{d} + \frac{1}{d} x$$

Es ist also $h(y) = \ln(\Leftrightarrow \ln(1 \Leftrightarrow y))$.

- (iie) Die Perk-Verteilung (Hyperbelsekans - Verteilung):

$$F(x) = \frac{2}{\pi} \arctan\left(\exp\left(\frac{x \Leftrightarrow m}{d}\right)\right)$$

$$\ln\left(\tan\left(\frac{\pi}{2} F(x)\right)\right) = \Leftrightarrow \frac{m}{d} + \frac{1}{d} x$$

Es ist also $h(y) = \ln \tan\left(\frac{\pi}{2} y\right)$.

- (iif) Die Cauchy - Verteilung:

$$F(x) = \frac{1}{\pi} \arctan\left(\frac{x \Leftrightarrow m}{d}\right) + \frac{1}{2}$$

$$\tan\left(\pi\left(F(x) \Leftrightarrow \frac{1}{2}\right)\right) = \Leftrightarrow \frac{m}{d} + \frac{1}{d} x$$

Hier ist $h(y) = \tan\left(\pi\left(y \Leftrightarrow \frac{1}{2}\right)\right)$.

(iii) Viele Verteilungen haben die Eigenschaft, daß der Logarithmus der jeweiligen Zufallsvariablen eine linear parametrische Verteilung aufweist. Ein bekanntes Beispiel hierfür ist die Gibrat - Verteilung. Für Verteilungen dieses Typs gilt $g(x) = \ln(x)$.

(iiia) Die Gibrat - Verteilung (Lognormalverteilung):

$$F(x) = \Phi\left(\frac{\ln x \Leftrightarrow \mu}{\sigma}\right)$$

$$\Phi^{-1}(F(x)) = \Leftrightarrow \frac{\mu}{\sigma} + \frac{1}{\sigma} \ln x$$

(iiib) Die Pareto - Verteilung (Logexponentialverteilung):

$$F(x) = 1 \Leftrightarrow \left(\frac{b}{x}\right)^p$$

$$\ln(1 \Leftrightarrow F(x)) = p \cdot \ln b \Leftrightarrow p \cdot \ln x$$

Diese Linearisierung wird im Pareto - Diagramm ausgenutzt.

(iiic) Die Fisk - Verteilung (Loglogistische Verteilung):

$$F(x) = \frac{x^p}{c^p + x^p}$$

$$\ln\left(\frac{F(x)}{1 \Leftrightarrow F(x)}\right) = \Leftrightarrow p \cdot \ln c + p \cdot \ln x$$

(iiid) Die Weibull - Verteilung (Loggumbelverteilung):

$$F(x) = 1 \Leftrightarrow \exp(\Leftrightarrow a x^b)$$

$$\ln(\Leftrightarrow \ln(1 \Leftrightarrow F(x))) = \ln a + b \cdot \ln x$$

(iv) Weitere Möglichkeiten der Linearisierung eröffnen die Johnson - Transformationen (Vgl. JOHNSON 1949).

e) Minimum-Quantil-Abstands-Methode

Wähle Parametervektor θ so, daß

$$S_3^2 = [\hat{\mathbf{Q}} \Leftrightarrow \mathbf{Q}(\theta)]' \mathbf{W}_3 [\hat{\mathbf{Q}} \Leftrightarrow \mathbf{Q}(\theta)] \quad (5)$$

minimiert wird.

Auch hier ist \mathbf{W}_3 noch näher zu bestimmen.

Weitere Schätzmethode werden nicht betrachtet, da sie nicht eingeordnet werden können. Bemerkenswert sind die Ansätze von LAWLESS (1982) und CHENG und CHEN (1988). Diese Ansätze basieren auf der Linearisierung der Überlebensfunktion bei Überlebensdaten, die als Häufigkeitstabelle vorliegen. Sie beruhen auf der Annahme, daß die Überlebensfunktion innerhalb der vorgegebenen Klassengrenzen konstant ist. Diese

Annahme ist jedoch in der Regel verletzt und wird auch nicht benötigt, wenn Methode d) angewandt wird.

Alle betrachteten Schätzverfahren beruhen darauf, daß sie eine Zielfunktion maximieren bzw. minimieren. Es versteht sich von selbst, daß zwei Schätzverfahren, die dieselbe Zielfunktion benutzen, auch zu denselben Schätzern führen werden. Ist die Zielfunktion eines Verfahrens eine monotone Funktion der Zielfunktion eines anderen Verfahrens, dann müssen ebenfalls beide Verfahren zu den selben Schätzern führen. Ist mit wachsendem n eine Zielfunktion der Grenzwert einer anderen, oder ist eine Zielfunktion eine monotone Funktion des Grenzwertes der anderen Zielfunktion, dann werden die Schätzer asymptotisch übereinstimmen (asymptotische Äquivalenz).

3 Vergleich der Schätzverfahren

3.1 ML - Schätzer und χ^2 - Minimum - Schätzer

Der ML - Schätzer beruht auf der Zielfunktion:

$$L = \prod p_i^{n_i} \quad \text{mit } p_i = p_i(\boldsymbol{\theta})$$

Diese hat die äquivalente (monoton transformierte) Zielfunktion

$$\ln \bar{L} = \sum n_i \cdot \ln p_i \Leftrightarrow \sum n_i \cdot \ln \hat{p}_i$$

Da der zweite Term nicht von θ abhängt, hat \bar{L} ihr Minimum an derselben Stelle wie L . Es gilt:

$$\ln \bar{L} = \sum n_i \ln \frac{p_i^{n_i}}{n_i}$$

Für die Funktion $\ln x$ gilt die Approximation $\ln x \approx (x^2 \Leftrightarrow 1)/2x$ mit besonders guten Näherungswerten, wenn $x \approx 1$. So ist auch

$$\begin{aligned} \ln \bar{L} &\approx \sum n_i \left(\frac{p_i^2 n_i^2}{n_i^2} \Leftrightarrow 1 \right) \frac{n_i}{2n p_i} \\ &= \frac{n}{2} \Leftrightarrow \frac{1}{2n} \sum \frac{n_i^2}{p_i} = \Leftrightarrow \frac{1}{2} \chi^2 \end{aligned}$$

Zunächst gilt also, daß die Zielfunktionen L und χ^2 näherungsweise äquivalent sind, wobei die Näherung umso besser ist, je näher $p_i(\boldsymbol{\theta})$ an $\frac{n_i}{n}$ liegt. Bei richtiger Modellspezifikation wird man Parameterwerte $\hat{\boldsymbol{\theta}}$ finden können, so daß $plim \frac{n_i}{n} = p_i(\hat{\boldsymbol{\theta}})$ und daher beide Verfahren zu asymptotisch gleichen Schätzern führen. Bei falscher Spezifikation gilt zwar nach wie vor $plim \frac{n_i}{n} = p_i$, dann wird es aber nicht gelingen, $p_i = p_i(\boldsymbol{\theta})$ darzustellen.

3.2 χ^2 - Minimum - Schätzer und Anpassung der Verteilungsfunktion

Es kann gezeigt werden, daß beide Schätzverfahren dieselbe Zielfunktion benutzen und daher identische Schätzer hervorrufen.

Der Vergleich wird zunächst dadurch erschwert, daß der Chi-Quadrat-Minimum-Schätzer k relative Häufigkeiten \hat{p}_i mit k Modellwahrscheinlichkeiten $p_i(\boldsymbol{\theta})$ vergleicht, während bei Anpassung der Verteilungsfunktionen nur $k \Leftrightarrow 1$ kumulierte Paare $\hat{F}_i, F_i(\boldsymbol{\theta})$ verwendet werden können. In der Tat haben die Vektoren $\hat{\mathbf{p}} = (\hat{p}_1, \dots, \hat{p}_k)$ und $\mathbf{p}(\boldsymbol{\theta}) = (p_1(\boldsymbol{\theta}), \dots, p_k(\boldsymbol{\theta}))$ je ein überflüssiges Element, da das jeweils letzte aus den anderen $k \Leftrightarrow 1$ bestimmt werden kann. Es gilt:

$$\hat{p}_k = 1 \Leftrightarrow \sum_1^{k-1} p_i$$

$$p_k(\boldsymbol{\theta}) = 1 \Leftrightarrow \sum_{i=1}^{k-1} p_i(\boldsymbol{\theta})$$

Gleichwohl verwendet χ^2 alle Elemente der Vektoren $\hat{\mathbf{p}}$ und $\mathbf{p}(\boldsymbol{\theta})$. Es kommt also darauf an, das χ^2 -Kriterium geeignet umzuformen.

$$\begin{aligned} X^2 &= \sum \frac{(n_i \Leftrightarrow np_i)^2}{np_i} = n \sum \frac{(\hat{p}_i \Leftrightarrow p_i)^2}{p_i} \\ &= n \sum_{i=1}^{k-1} \frac{(\hat{p}_i \Leftrightarrow p_i)^2}{p_i} + n \frac{[(1 \Leftrightarrow \sum_{i=1}^{k-1} \hat{p}_i) \Leftrightarrow (1 \Leftrightarrow \sum_{i=1}^{k-1} p_i)]^2}{1 \Leftrightarrow \sum_{i=1}^{k-1} p_i} \\ &= n \left\{ \sum_{i=1}^{k-1} \frac{(\hat{p}_i \Leftrightarrow p_i)^2}{p_i} + \frac{(\sum_{i=1}^{k-1} p_i \Leftrightarrow \sum_{i=1}^{k-1} \hat{p}_i)^2}{1 \Leftrightarrow \sum_{i=1}^{k-1} p_i} \right\} \\ &= n \left\{ \sum_{i=1}^{k-1} \frac{(\hat{p}_i \Leftrightarrow p_i)^2}{p_i} + \frac{[\sum_{i=1}^{k-1} (\hat{p}_i \Leftrightarrow p_i)]^2}{1 \Leftrightarrow \sum_{i=1}^{k-1} p_i} \right\} \end{aligned}$$

In Matrizendarstellung lautet dies:

Sei $\mathbf{1}$ ein Einsenvektor und $\mathbf{1}\mathbf{1}'$ demzufolge eine quadratische Einsenmatrix. Sei weiterhin $\check{\mathbf{x}}$ eine aus den Elementen des Vektors x aufgebaute Diagonalmatrix, dann gilt

$$\frac{1}{n} X^2 = (\hat{\mathbf{p}} \Leftrightarrow \mathbf{p})' \left[(\check{\mathbf{p}})^{-1} + \frac{\mathbf{1}\mathbf{1}'}{p_k} \right] (\hat{\mathbf{p}} \Leftrightarrow \mathbf{p})$$

Bei dieser Darstellung ist zu beachten, daß die Vektoren \mathbf{p} und $\hat{\mathbf{p}}$ aus jeweils $k \Leftrightarrow 1$ Elementen bestehen und daß $p_k = 1 \Leftrightarrow \sum_{i=1}^k p_i = 1 \Leftrightarrow \mathbf{1}'\mathbf{p}$. Die in [...] gesetzte Matrix ist übrigens die Inverse der Varianz-Kovarianz-Matrix von $\hat{\mathbf{p}}$:

$$E(\hat{\mathbf{p}} \Leftrightarrow \mathbf{p})(\hat{\mathbf{p}} \Leftrightarrow \mathbf{p})' = \check{\mathbf{p}} \Leftrightarrow \mathbf{p}\mathbf{p}'.$$

Dies gilt, weil

$$\begin{aligned} &(\check{\mathbf{p}} \Leftrightarrow \mathbf{p}\mathbf{p}')((\check{\mathbf{p}})^{-1} \Leftrightarrow \frac{\mathbf{1}\mathbf{1}'}{p_k}) \\ &= \mathbf{I} \Leftrightarrow \mathbf{p}\mathbf{1}' + \frac{\mathbf{p}\mathbf{1}'}{p_k} \Leftrightarrow \frac{\mathbf{p}(1 \Leftrightarrow p_k)\mathbf{1}'}{p_k} = \mathbf{I} \quad q.e.d. \end{aligned}$$

Daraus folgt die Darstellung:

$$\frac{X^2}{n} = (\hat{\mathbf{p}} \Leftrightarrow \mathbf{p})' \mathbf{V}_{\hat{\mathbf{p}}}^{-1} (\hat{\mathbf{p}} \Leftrightarrow \mathbf{p})$$

Da aber mit (Kumulationsmatrix)

$$\mathbf{K} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 1 & 1 & \cdots & 1 \end{pmatrix}$$

$$(\hat{\mathbf{F}} \Leftrightarrow \mathbf{F}) = \mathbf{K}(\hat{\mathbf{p}} \Leftrightarrow \mathbf{p})$$

und

$$\mathbf{V}_{\hat{\mathbf{F}}} = \begin{pmatrix} F_1(1 \Leftrightarrow F_1) & F_1(1 \Leftrightarrow F_2) & \cdots & F_1(1 \Leftrightarrow F_k) \\ F_1(1 \Leftrightarrow F_2) & F_2(1 \Leftrightarrow F_2) & \cdots & F_2(1 \Leftrightarrow F_k) \\ \vdots & \vdots & \ddots & \vdots \\ F_1(1 \Leftrightarrow F_k) & F_2(1 \Leftrightarrow F_k) & \cdots & F_k(1 \Leftrightarrow F_k) \end{pmatrix} \quad (6)$$

$$= E(\hat{\mathbf{F}} \Leftrightarrow \mathbf{F})(\hat{\mathbf{F}} \Leftrightarrow \mathbf{F})' = E(\mathbf{K}(\hat{\mathbf{p}} \Leftrightarrow \mathbf{p})(\hat{\mathbf{p}} \Leftrightarrow \mathbf{p})'\mathbf{K}') = \mathbf{K}\mathbf{V}_{\hat{\mathbf{p}}}\mathbf{K}'$$

gilt

$$\mathbf{V}_{\hat{\mathbf{F}}}^{-1} = \mathbf{K}'^{-1}\mathbf{V}_{\hat{\mathbf{p}}}^{-1}\mathbf{K}^{-1},$$

wobei

$$\mathbf{K}^{-1} = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ \Leftrightarrow 1 & 1 & \ddots & & \vdots \\ 0 & \Leftrightarrow 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \Leftrightarrow 1 & 1 \end{pmatrix}$$

(Dekumulationsmatrix). Daher ist auch,

$$\begin{aligned} \frac{X^2}{n} &= (\hat{\mathbf{p}} \Leftrightarrow \mathbf{p})'\mathbf{V}_{\hat{\mathbf{p}}}^{-1}(\hat{\mathbf{p}} \Leftrightarrow \mathbf{p}) = (\hat{\mathbf{p}} \Leftrightarrow \mathbf{p})'\mathbf{K}'\mathbf{K}'^{-1}\mathbf{V}_{\hat{\mathbf{p}}}^{-1}\mathbf{K}^{-1}\mathbf{K}(\hat{\mathbf{p}} \Leftrightarrow \mathbf{p}) \\ &= (\hat{\mathbf{F}} \Leftrightarrow \mathbf{F})'\mathbf{V}_{\hat{\mathbf{F}}}^{-1}(\hat{\mathbf{F}} \Leftrightarrow \mathbf{F})' \end{aligned}$$

Die Schätzverfahren stimmen also überein und $\mathbf{V}_{\hat{\mathbf{F}}}^{-1}$ ist eine angemessene Gewichtungsmatrix \mathbf{W}_1 .

3.3 Anpassung der Verteilungsfunktion mit und ohne Linearisierung

Eine Approximation des Störterms v_i der linearen Regressionsgleichung kann durch eine Taylor-Entwicklung des Störterms der nichtlinearen Regressionsgleichung

$$u_i = \hat{F}_i \Leftrightarrow F(x_i)$$

gewonnen werden:

$$h(F(x_i)) = h(\hat{F}_i \Leftrightarrow u_i).$$

Dies liefert

$$\begin{aligned}
h(\hat{F}_i) &= h(F(x_i)) + (\hat{F}_i \Leftrightarrow F(x_i)) \cdot h'(F(x_i)) + \dots \\
&= h(F(x_i)) + u_i \cdot h'(F(x_i)) + \dots \\
&= h(F(x_i)) + u_i \cdot h'(\hat{F}_i) + \dots \\
&= a + b \cdot g(x_i) + u_i \cdot h'(\hat{F}_i) + \dots
\end{aligned}$$

Für $n \rightarrow \infty$ konvergiert der Approximationsfehler hierbei gegen Null.

Sei \mathbf{U} die Kovarianzmatrix des Störvektors \mathbf{u} . Dann ergibt sich die Kovarianzmatrix der Störgrößen \mathbf{v} approximativ durch

$$\mathbf{V} = \text{diag}(h'(\hat{F}_i)) \mathbf{U} \text{diag}(h'(\hat{F}_i))$$

und

$$\mathbf{V}^{-1} = \text{diag}\left(\frac{1}{h'(\hat{F}_i)}\right) \mathbf{U}^{-1} \text{diag}\left(\frac{1}{h'(\hat{F}_i)}\right). \quad (7)$$

Die Matrix \mathbf{V}^{-1} ist eine geeignete Gewichtungsmatrix \mathbf{W}_2 für die Zielfunktion in Gleichung 4.

Die Zielfunktion S_2^2 unterscheidet sich nur insofern von der Zielfunktion S_1^2 , als die Störterme bei der Linearisierung lediglich approximativ übertragen werden. Mit einer steigenden Anzahl von Beobachtungen konvergiert dieser Approximationsfehler gegen Null. Die in c) und d) beschriebenen Methoden sind damit asymptotisch äquivalent.

3.4 Optimale Anpassung der Verteilungsfunktion und Chi-Quadrat-Minimum-Methode

Sofern die Quantilfunktion linear in den Koeffizienten ist, und das ist immer der Fall, wenn die Verteilungsfunktion linearisierbar ist, ist die Schätzung der Minimum-Quantil-Abstands-Methode nichts anderes als die Umkehrregression zur Schätzung nach Linearisierung der Verteilungsfunktion. Da die Umkehrung der Regressionsfunktion zur Veränderung der Schätzergebnisse führt, werden die Parameterschätzungen unterschiedlich ausfallen. Da aber bei richtiger Spezifizierung des Modells mit wachsendem Stichprobenumfang schließlich exakt auf einer Geraden liegen müssen, kann asymptotisch kein Unterschied mehr bestehen.

Ist die Quantilfunktion nicht linearisierbar, ändert sich im Prinzip nichts. Nach wie vor bedeutet die Minimierung der Quantilabstände die Umkehrung zur optimalen Anpassung der Verteilungsfunktion, und mit wachsendem Stichprobenumfang muß bei richtiger Spezifizierung des Modells der Unterschied schließlich verschwinden.

4 Ein Anwendungsbeispiel: die zweiparametrische Weibull-Verteilung

Betrachtet wird die zweiparametrische Weibull-Verteilung mit dem Formparameter b und dem Skalenparameter a . Diese Verteilung ist interpretierbar als Verallgemeinerung der Exponentialverteilung, die sich als Spezialfall für $b = 1$ ergibt. Während sich bei der Exponentialverteilung konstante Ausfallraten ergeben, lassen sich mit der Weibull-Verteilung auch zunehmende ($b > 1$) oder abnehmende ($b < 1$) Ausfallraten modellieren.

Die Dichtefunktion dieser Verteilung ist gegeben durch

$$f(x) = a \cdot b \cdot x^{b-1} \cdot \exp(-a \cdot x^b) \quad a, b, x > 0.$$

Die Verteilungsfunktion lautet entsprechend:

$$F(x) = 1 - \exp(-a \cdot x^b).$$

Das Schätzproblem besteht darin, die Parameter a und b bei Vorliegen klassierter Daten zu schätzen. Bei gegebenen Klassengrenzen $(\tilde{x}_0, \tilde{x}_1, \dots, \tilde{x}_k)$ handelt es sich bei den Besetzungszahlen n_i und davon abgeleiteten Größen um Zufallsgrößen. Für den hier betrachteten Fall gilt dabei: $\tilde{x}_0 = 0$ und $\tilde{x}_k = \infty$.

Die verschiedenen Verfahren zur Schätzung der Parameter a und b werden dargestellt und in einer Simulationsstudie verglichen.

4.1 Maximum-Likelihood-Schätzung

Die Likelihoodfunktion für den hier betrachteten Fall lautet:

$$\begin{aligned} L(n_1, n_2, \dots, n_k; a, b) &= \prod_i \{P(X \text{ realisiert sich in Klasse } i)(a, b)\}^{n_i} & (8) \\ &= \prod_i \{F(\tilde{x}_i; a, b) - F(\tilde{x}_{i-1}; a, b)\}^{n_i} \\ &= \prod_i \{\exp(-a \cdot \tilde{x}_{i-1}^b) - \exp(-a \cdot \tilde{x}_i^b)\}^{n_i} \end{aligned}$$

Diese Funktion – bzw. die logarithmierte Funktion – wird dann bezüglich der Parameter a und b maximiert. Die Schätzwerte lassen sich nicht explizit angeben. Die Maximierung erfolgt mittels numerischer Verfahren.

4.2 Der Minimum-Quantil-Abstands-Schätzer

Das Verfahren der Minimum-Quantil-Abstand(MQA)-Schätzung stellt eine Verallgemeinerung des Verfahrens der Quantilschätzung dar, bei dem alle aus der

Häufigkeitstabelle bekannten empirischen Quantile zu der Schätzung genutzt werden. Hierbei wird die quadratische Form in Gleichung 5 bezüglich der Parameter a und b minimiert. Für den Fall, daß die Gewichtungsmatrix \mathbf{W} einer Einheitsmatrix entspricht, enthält Gleichung 5 die Summe der quadrierten Abstände der empirischen von den theoretischen Quantilen. Diese Schätzung wird als ungewichtete MQA-Schätzung bezeichnet.

Eine geeignete Gewichtungsmatrix wurde von CARMODY, EUBANK und LARICCIA (1984) vorgeschlagen. Die so ermittelten Schätzer sind ebenfalls nur numerisch bestimmbar. Außerdem sind sie nach CARMODY, EUBANK und LARICCIA eindeutig, konsistent und asymptotisch normalverteilt. Allerdings ist zu beachten, daß bei vorgegebenen Klassengrenzen die Zellhäufigkeiten und die daraus abgeleiteten Häufigkeiten als Zufallsgrößen aufzufassen sind.

4.3 Optimale Anpassung der Verteilungsfunktion nach Linearisierung

Bei der nicht linearisierten Variante wird der Ausdruck in Gleichung 3 minimiert, wobei die Inverse der Kovarianzmatrix des Vektors der kumulierten relativen Häufigkeiten, $\mathbf{V}_{\hat{F}}$, eine geeignete Gewichtungsmatrix darstellt.

Die Verteilungsfunktion ist aber in dem hier betrachteten Fall linearisierbar:

$$\hat{F}(\tilde{x}_i) = 1 \Leftrightarrow \exp(\Leftrightarrow a \tilde{x}_i^b) \Leftrightarrow \ln(\Leftrightarrow \ln(1 \Leftrightarrow \hat{F}(\tilde{x}_i))) = \ln(a) + b \cdot \ln(\tilde{x}_i) \quad (9)$$

Das Problem der Parameterschätzung einer zweiparametrischen Weibull-Verteilung wurde damit zurückgeführt auf das Problem der Schätzung der Parameter in dem verallgemeinerten linearen Modell

$$\ln(\Leftrightarrow \ln(1 \Leftrightarrow F_i)) = \beta_0 + \beta_1 \cdot \ln(\tilde{x}_i) + v_i \quad \text{für } i = 1, \dots, k \Leftrightarrow 1 \quad (10)$$

Zur Bestimmung der Kovarianzmatrix des Vektors \mathbf{v} wird Gleichung 7 verwendet. Mit $h(\hat{F}) = \ln(\Leftrightarrow \ln(1 \Leftrightarrow \hat{F}))$ folgt:

$$n \cdot \text{Var}(\mathbf{v}) = \begin{pmatrix} \frac{1-F_1}{F_1 \cdot \ln^2(1-F_1)} & \frac{1-F_2}{F_2 \cdot \ln(1-F_1) \cdot \ln(1-F_2)} & \cdots & \frac{1-F_{k-1}}{F_{k-1} \cdot \ln(1-F_1) \cdot \ln(1-F_{k-1})} \\ \frac{1-F_2}{F_2 \cdot \ln(1-F_1) \cdot \ln(1-F_2)} & \frac{1-F_2}{F_2 \cdot \ln^2(1-F_2)} & \cdots & \frac{1-F_{k-1}}{F_{k-1} \cdot \ln(1-F_2) \cdot \ln(1-F_{k-1})} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1-F_{k-1}}{F_{k-1} \cdot \ln(1-F_1) \cdot \ln(1-F_{k-1})} & \frac{1-F_{k-1}}{F_{k-1} \cdot \ln(1-F_2) \cdot \ln(1-F_{k-1})} & \cdots & \frac{1-F_{k-1}}{F_{k-1} \cdot \ln^2(1-F_{k-1})} \end{pmatrix} \quad (11)$$

Auf der Grundlage der Gleichung 10 wird eine gewichtete Kleinst-Quadrate-Schätzung durchgeführt, wobei die Inverse von $\text{Var}(\mathbf{v})$ als Gewichtungsmatrix verwendet werden kann.

Für die Parameter der Weibull-Verteilung gilt dann: $\hat{b} = \hat{\beta}_1$ und $\hat{a} = \exp(\hat{\beta}_0)$.

Außerdem kann die gesamte Asymptotik des verallgemeinerten linearen Modells unmittelbar auf den Formparameter der Weibull-Verteilung übertragen werden.

4.4 Beschreibung der Simulationen

Um die Eignung der verschiedenen Schätzverfahren zu überprüfen, wurden Simulationen durchgeführt. Zunächst wurde eine einfache Zufallsstichprobe aus einer Weibull-Verteilung mit den Parametern $a = 1$ und $b = 0.5, 1, 2$ gezogen. Es werden 5 Klassen gebildet, wobei die Klassengrenzen dem 20-, 40-, 60-, bzw. 80%-Punkt der theoretischen Verteilung entsprechen. Die Schätzung der Parameter erfolgt dann auf der Grundlage der resultierenden Häufigkeitsverteilungen. Insgesamt werden 8 Schätzmethoden verglichen*.

Verfahren	Akürzung
Maximum-Likelihood-Methode	ML
Linearisierung der Verteilungsfunktion (ungewichtet)	AVLU
(gewichtet)	AVL
Anpassung der Verteilungsfunktion (ungewichtet)	AVU
(gewichtet mit F)	AVF
(gewichtet mit \hat{F})	AVFDach
Anpassung der Quantilfunktion (ungewichtet)	MQAU
(gewichtet)	MQA

Die Simulationsergebnisse – Mittelwert, empirische Varianz und mittlerer quadratischer Abstand vom wahren Parameterwert – sind im Anhang dokumentiert. Hierbei ist zu beachten, daß zur Erhöhung der Übersichtlichkeit die Varianzen und die mittleren quadratischen Fehler mit 100 multipliziert wurden.

Zunächst ist auffällig, daß nahezu alle Verfahren bezüglich der beiden zu schätzenden Parameter Verzerrungen aufweisen, die jedoch mit wachsendem Stichprobenumfang verschwinden. Verzerrte, aber asymptotisch unverzerrte Schätzer werden in Simulationsstudien häufig das folgende, charakteristische Muster aufweisen: Verzerrungen weisen bei unterschiedlichen Stichprobenumfängen das gleiche Vorzeichen auf, nehmen mit wachsendem Stichprobenumfang aber ab. Beispielsweise ergibt sich bei dem Maximum-Likelihood-Schätzer für den Parameter a :

Parameter- konstellation	Mittel der Schätzwerte für a		
	$n=50$	$n=100$	$n=200$
$a=1, b=0.5$	1.0217	1.0040	1.0033
$a=1, b=1$	1.0158	1.0052	1.0032
$a=1, b=2$	1.0200	1.0064	1.0035

Dieses Muster deutet eindeutig darauf hin, daß der ML-Schätzer für den Parameter a – ebenso wie übrigens der ML-Schätzer für den Parameter b – den wahren Parameterwert systematisch überschätzt. Das auf der gewichteten Schätzung einer Regressionsgerade beruhende Verfahren AVL ist bezüglich des Parameters a ebenfalls nach oben verzerrt, der Parameter b wird bei diesem Verfahren unterschätzt. Die Größenordnung der Verzerrungen der Verfahren ML und AVL sind vergleichbar.

*Alle Simulationen wurden mit S-Plus mit 1000 Wiederholungen durchgeführt. Schätzer, die numerisch zu bestimmen sind, wurden unter Verwendung der Funktion *nlmin* ermittelt. Außerdem muß angemerkt werden, daß in die von CARMODY, EUBANK und LARICCIA (1984) vorgeschlagene Gewichtungsmatrix der gesuchte Parameter b eingeht, so daß dieser normalerweise vorausgeschätzt werden muß. Im Rahmen der Simulationen wurde der wahre Parameterwert verwendet. Deshalb schneidet das Verfahren MQA wohl etwas zu gut ab.

Daneben ist zu beobachten, daß die Verzerrung eine untergeordnete Rolle für den Mittleren Quadratischen Fehler spielt. Der MQF stimmt in allen Fällen fast mit der empirischen Varianz überein.

Um die Güte der Verfahren zu vergleichen, werden in Anlehnung an das Konzept der relativen Effizienz die Mittleren Quadratischen Fehler der alternativen Schätzverfahren verglichen. Zu diesem Zweck werden die Mittleren Quadratischen Fehler aller Verfahren mit dem MQF des jeweils besten Verfahrens verglichen, indem Quotienten gebildet werden. Dieses Vorgehen dient der Erhöhung der Anschaulichkeit der Ergebnisse. Zusätzlich werden den Schätzverfahren die entsprechenden Rangwerte zugeordnet.

a=1, b=0.5												
	Schätzwerte für a						Schätzwerte für b					
	n=50		n=100		n=200		n=50	n=100	n=200			
AVLU	1.052	(4)	1.033	(4)	1.033	(5)	1.241	(6)	1.139	(6)	1.251	(7)
AVL	1	(1)	1	(1)	1.011	(2)	1.154	(4)	1.030	(2)	1.155	(6)
ML	1.030	(3)	1.013	(3)	1.019	(3)	1.067	(2)	1	(1)	1.098	(3)
MQUAU	1.182	(8)	1.141	(5)	1.123	(8)	2.243	(8)	2.321	(8)	2.390	(8)
MQA	1.023	(2)	1.011	(2)	1.019	(4)	1.221	(5)	1.116	(4)	1.129	(4)
AVU	1.057	(5)	1.302	(7)	1.080	(7)	1	(1)	1.123	(5)	1.143	(5)
AVF	1.171	(7)	1.178	(6)	1	(1)	1.088	(3)	1.107	(3)	1.058	(2)
AVFDach	1.137	(6)	1.211	(8)	1.043	(6)	1.289	(7)	1.174	(7)	1	(1)

a=1, b=1												
	Schätzwerte für a						Schätzwerte für b					
	n=50		n=100		n=200		n=50	n=100	n=200			
AVLU	1,043	(4)	1,066	(6)	1,114	(4)	1,169	(6)	1,080	(5)	1,109	(7)
AVL	1	(1)	1,032	(2)	1,108	(3)	1,136	(4)	1,048	(3)	1,033	(3)
ML	1,030	(3)	1,046	(4)	1,117	(7)	1	(1)	1	(1)	1	(1)
MQUAU	1,053	(5)	1,067	(7)	1,115	(6)	1,301	(8)	1,322	(8)	1,286	(8)
MQA	1,009	(2)	1,040	(3)	1,114	(5)	1,066	(2)	1,008	(2)	1,019	(2)
AVU	1,119	(6)	1,057	(5)	1,105	(2)	1,147	(5)	1,073	(4)	1,056	(4)
AVF	1,148	(7)	1	(1)	1	(1)	1,097	(3)	1,117	(6)	1,085	(6)
AVFDach	1,204	(8)	1,166	(8)	1,155	(8)	1,260	(7)	1,142	(7)	1,064	(5)

a=1, b=2												
	Schätzwerte für a						Schätzwerte für b					
	n=50		n=100		n=200		n=50	n=100	n=200			
AVLU	1,072	(6)	1,031	(6)	1,028	(6)	1,185	(8)	1,163	(8)	1,145	(8)
AVL	1,023	(2)	1	(1)	1,006	(2)	1,072	(3)	1,029	(5)	1,082	(6)
ML	1,055	(5)	1,014	(3)	1,012	(4)	1	(1)	1,006	(3)	1,054	(4)
MQUAU	1,073	(7)	1,021	(5)	1,019	(5)	1,088	(6)	1,060	(6)	1,086	(7)
MQA	1,026	(3)	1,003	(2)	1,006	(3)	1,022	(2)	1,012	(4)	1,066	(5)
AVU	1	(1)	1,042	(8)	1,050	(7)	1,073	(4)	1,006	(2)	1,019	(3)
AVF	1,257	(8)	1,017	(4)	1	(1)	1,076	(5)	1	(1)	1,016	(2)
AVFDach	1,028	(4)	1,036	(7)	1,108	(8)	1,173	(7)	1,079	(7)	1	(1)

Bezüglich des Parameters a dominiert das Verfahren AVL das ML-Verfahren in jedem der betrachteten Fälle. Bezüglich des Parameters b schneiden die ML-Schätzwerte besser ab. Angesichts der Tatsache, daß lediglich die auf der Linearisierung der Verteilungsfunktion

beruhenden verfahren AVL und AVLU analytisch darstellbare Lösungen besitzen, alle anderen Verfahren lediglich unter Verwendung numerischer Verfahren lösbar sind, kann zusammengefaßt werden, daß das Verfahren AVL sowohl hinsichtlich der Güte der Eigenschaften der Parameterschätzer als auch hinsichtlich des benötigte Rechenaufwandes in der Anwendung zu empfehlen ist.

A Die Simulationsergebnisse

		a=1, b=0.5					
		Schätzwerte für a			Schätzwerte für b		
		n=50	n=100	n=200	n=50	n=100	n=200
AVLU	Mittel	1.0219	1.0049	1.0034	0.5162	0.5072	0.4995
	Varianz	3.128	1.173	0.6425	0.8343	0.3452	0.1808
	MQF	3.172	1.174	0.643	0.8598	0.3464	0.1807
AVL	Mittel	1.0268	1.0066	1.0047	0.4868	0.4927	0.4927
	Varianz	2.947	1.134	0.6276	0.7827	0.3082	0.1618
	MQF	3.016	1.137	0.6292	0.7992	0.3132	0.1669
ML	Mittel	1.0217	1.004	1.0033	0.5131	0.505	0.4988
	Varianz	3.062	1.151	0.6339	0.7227	0.302	0.1586
	MQF	3.106	1.152	0.6343	0.7391	0.3042	0.1586
MQAU	Mittel	1.0357	1.0111	1.0062	0.4996	0.4965	0.4959
	Varianz	3.442	1.286	0.6959	1.556	0.7055	0.344
	MQF	3.566	1.297	0.699	1.554	0.706	0.3454
MQA	Mittel	1.0152	1.0026	1.0021	0.5196	0.5073	0.5
	Varianz	3.064	1.15	0.6366	0.847	0.3345	0.1631
	MQF	3.084	1.149	0.6346	0.8461	0.3394	0.1631
AVU	Mittel	1.0262	1.0123	0.9989	0.511	0.5054	0.50384
	Varianz	3.1225	1.4664	0.6728	0.6814	0.3391	0.1639
	MQF	3.188	1.48	0.6722	0.6927	0.3417	0.1652
AVF	Mittel	1.0243	1.0101	1.0074	0.5179	0.5037	0.5031
	Varianz	3.475	1.3303	0.6175	0.7426	0.3357	0.1521
	MQF	3.5305	1.339	0.6226	0.7535	0.3368	0.1529
AVFDach	Mittel	1.0030	0.9995	0.9999	0.5118	0.4993	0.4984
	Varianz	3.4319	1.379	0.6499	0.8936	0.3575	0.1444
	MQF	3.4294	1.3774	0.6492	0.8928	0.3572	0.1445

		a=1, b=1					
		Schätzwerte für a			Schätzwerte für b		
		n=50	n=100	n=200	n=50	n=100	n=200
AVLU	Mittel	1.0164	1.0052	1.0034	1.0315	1.0127	1.0048
	Varianz	2.687	1.387	0.6726	2.929	1.381	0.6883
	MQF	2.711	1.388	0.6731	3.026	1.396	0.6899
AVL	Mittel	1.0214	1.008	1.0045	0.9692	0.9852	0.9909
	Varianz	2.557	1.339	0.6678	2.848	1.333	0.6351
	MQF	2.6	1.344	0.6692	2.94	1.354	0.6427
ML	Mittel	1.0158	1.0052	1.0032	1.2039	1.0108	1.0028
	Varianz	2.655	1.361	0.6746	2.533	1.282	0.6219
	MQF	2.677	1.362	0.6749	2.588	1.292	0.622
MQAU	Mittel	1.0222	1.0081	1.0048	1.0179	1.009	1.0008
	Varianz	2.692	1.385	0.6719	3.337	1.702	0.8009
	MQF	2.739	1.39	0.6735	3.366	1.708	0.8002
MQA	Mittel	1.0147	1.0051	1.0032	1.0055	1.0016	0.9984
	Varianz	2.603	1.354	0.6726	2.758	1.303	0.6341
	MQF	2.623	1.355	0.673	2.758	1.302	0.6337
AVU	Mittel	1.0211	1.0072	1.0037	1.0260	1.01566	1.00301
	Varianz	2.867	1.3726	0.6667	2.904	1.3633	0.6565
	MQF	2.909	1.3765	0.66745	2.969	1.3865	0.6568
AVF	Mittel	1.016	1.0067	0.9984	1.0279	1.01521	1.0104
	Varianz	2.962	1.299	0.6044	2.7647	1.4213	0.665
	MQF	2.984	1.3024	0.6041	2.8398	1.443	0.6751
AVFDach	Mittel	0.9951	0.9999	1.0022	0.9992	1.0063	0.9974
	Varianz	3.1304	1.5196	0.69815	3.265	1.4731	0.6618
	MQF	3.1296	1.5181	0.69796	3.262	1.4756	0.6618

		a=1, b=2					
		Schätzwerte für a			Schätzwerte für b		
		n=50	n=100	n=200	n=50	n=100	n=200
AVLU	Mittel	1.0201	1.0063	1.003	2.2569	2.0262	2.0123
	Varianz	2.996	1.31	0.6401	12.53	5.829	2.871
	MQF	3.034	1.313	0.6403	12.84	5.892	2.883
AVL	Mittel	1.0252	1.009	1.005	1.9412	1.9727	1.9859
	Varianz	2.834	1.267	0.6246	11.27	5.145	2.705
	MQF	2.895	1.274	0.6264	11.61	5.214	2.722
ML	Mittel	1.02	1.0064	1.0035	2.0439	2.0208	2.0124
	Varianz	2.95	1.29	0.6298	10.65	5.056	2.639
	MQF	2.987	1.292	0.6304	10.83	5.094	2.652
MQAU	Mittel	1.0242	1.0082	1.0045	2.075	2.0345	2.0206
	Varianz	2.98	1.295	0.6334	11.24	5.245	2.694
	MQF	3.036	1.301	0.6348	11.79	5.37	2.734
MQA	Mittel	1.0222	1.0077	1.0043	1.9746	1.988	1.9939
	Varianz	2.857	1.273	0.626	11.01	5.116	2.681
	MQF	2.903	1.278	0.6267	11.07	5.125	2.682
AVU	Mittel	1.01088	1.0085	1.0039	2.0461	2.0312	2.0106
	Varianz	2.8211	1.322	0.6531	11.42	5.006	2.556
	MQF	2.8301	1.328	0.654	11.62	5.098	2.565
AVF	Mittel	1.0136	1.0071	1.0064	2.0493	2.0179	2.0126
	Varianz	3.543	1.292	0.6194	11.42	5.038	2.543
	MQF	3.558	1.296	0.6229	11.65	5.065	2.556
AVFDach	Mittel	0.9995	0.9964	0.9974	2.0021	1.999	2.002
	Varianz	2.914	1.3197	0.6904	12.73	5.471	2.519
	MQF	2.911	1.3196	0.6904	12.71	5.466	2.516

B Literatur

Carmody, T.J., Eubank, R.L. and V.N. LaRiccia (1984). "A family of minimum quantile distance estimators for the three-parameter Weibull distribution", *Statistische Hefte*, Vol. 25, 69-82.

Cheng, K.F. and C.H. Chen (1988). "Estimation of the Weibull parameters with grouped data", *Commun. Statist. - Theory Meth.*, Vol. 17(2), 325-341.

Gehan, E.A. and M.M. Siddiqui (1973). "Simple regression methods for survival time studies", *Journal of the American Statistical Association*, Vol. 68(344), 848-856.

Johnson, N.L. (1949). "Systems of frequency curves generated by methods of translation", *Biometrika*, Vol. 36, 149-176.

Johnson, N.L, Kotz, S. and N. Balakrishnan (1994). "Continuous univariate distributions", Vol. 1, 2nd Edition, New York: Wiley.

Lawless, J.F. (1982). "Statistical models and methods for lifetime data", New York: Wiley.

McDonald, J.B. and M.R. Ransom (1979). "Alternative parameter estimates based upon grouped data", *Commun. Statist. - Theory Meth.*, Vol. A8(9), 899-917.