

Durchführung einer Regression über das Resttermverfahren am Beispiel einer bilogarithmischen Funktion

Dipl.- Ing. Björnsterne Zindler, M.Sc.

Erstellt: 21. November 2014 – Letzte Revision: 6. Dezember 2014

Inhaltsverzeichnis

1 Herleitungen	2
1.1 Einleitung zum Thema	2
1.2 Darstellung der Regressionsfunktion	3
1.2.1 Der Sonderfall $p = q = 1$	3
1.2.2 Der allgemeine Fall $p \neq 1$ oder $q \neq 1$	3
1.3 Durchführung der Regression über das Resttermverfahren	4
1.4 Der Restterm als Fehlerwertkontrolle	5
1.5 Der Pearson- Korrelationskoeffizient ersten Grades - $\rho^{(1)}$	6
1.6 Definition der Koeffizienten	7
1.6.1 Definition von q	7
1.6.2 Definition von p	7
1.6.3 Der Definitionsbereich der Arbeitsgleichung	7
1.6.4 Der Wertebereich der Arbeitsgleichung	7
2 Beispiele für eine Regression	8
2.1 Beispiel 1 – Große x_i - und y_i - Werte	8
2.2 Beispiel 2 – Kleine x_i - und y_i - Werte	11
2.3 Beispiel 3 – Falsche Voraussetzungen für x_i - und y_i - Werte	14
2.4 Beispiel 4 – Durch Modifikation der Arbeitsgleichung das Beispiel 3 korrigiert	17
2.5 Beispiel – Zusammenfassung der statistischen Werte der Beispiele 1 - 4	21
3 Zusammenfassung der Berechnungsgrundlagen	22
3.1 Bilogarithmisch monoton steigend	22
3.2 Bilogarithmisch monoton fallend	23

Literatur

[001] Keine für vorliegenden Text.

1 Herleitungen

1.1 Einleitung zum Thema

[001]

Gegeben ist eine spezielle Funktion, welche die Koeffizienten bzw. Freiheitsgrade p und q besitzt. Diese sollen regressiert werden für eine Anzahl von Datenpunkten $P_i(x_i, y_i)$.

Einleitung

Für die vorliegende Form der Funktion wurde bisher keine Regressionsvorschrift gefunden. Im Projekt SAW- 2012 war jedoch eine solche Datenbearbeitung notwendig. Daher befasst sich das vorliegende Skript mit der Herleitung eines Algorithmus, welche p und q ermittelt. Im Verlaufe dieser Herleitung muss ein Startwert für die Koeffizienten aufgestellt werden. Daher handelt es sich streng genommen nicht um eine Regression, sondern um einen Kurvenfit mit einer bekannten Wahrscheinlich des Ergebnisses, die exakteste Lösung unter den möglichen Ergebnissen zu sein. Voraussetzung für eine erfolgreiche Ermittlung der Koeffizienten ist, dass vorliegende Datenpaare im Vornherein in der Nähe der noch zu ermittelnden Arbeitsgleichung liegen.

Eine Funktion $f(x)$ sei bekannt aus den Berechnungsmöglichkeiten der Euler- Mascheroni- Konstante γ .

$$\gamma = \int_0^{\infty} f(x) \cdot dx = - \int_0^{\infty} \ln(-\ln x) \cdot dx$$

Grundgedanke $f(x) = -\ln(-\ln x)$ regressieren zu können, ist die der Resttermregression.¹

Eine durch einfache Mittel linear regressierte Funktion $Y = b + a \cdot X$ mit den Koeffizienten a und b lässt die Werte von p und q bekannt werden. Das Vorgehen dahin, wird im weiteren Verlauf erläutert.

¹Siehe vorher Dipl.- Ing. Björnstjerne Zindler, M.Sc. „Das Resttermverfahren“ <http://www.ZenithPoint.de>

1.2 Darstellung der Regressionsfunktion

Gegeben ist die parametrische Beschreibung einer speziellen, uns interessierenden Funktion.

Regressionsfkt.

1.2.1 Der Sonderfall $p = q = 1$

$$x = e^{-t} \quad y = -\ln t$$

Wobei durch Koordinatentransformation $t \rightarrow y, x \rightarrow t$ bzw. $t \rightarrow x, y \rightarrow t$ beide Darstellungen ineinander überführt werden können.

$$t = e^{-y} \quad t = -\ln x$$

\Rightarrow

$$-\ln t = y \quad e^{-t} = x$$

Die Parameterdarstellung insgesamt kann in das kartesische Koordinatensystem überführt werden.

$$t = -\ln x \quad y = \ln t$$

\Rightarrow

$$y = -\ln(-\ln x)$$

1.2.2 Der allgemeine Fall $p \neq 1$ oder $q \neq 1$

Gegeben ist die parametrische Beschreibung einer speziellen, uns interessierenden Funktion.

$$x = p \cdot e^{-q \cdot t} \quad y = -\frac{1}{q} \cdot \ln \frac{t}{p}$$

Wobei durch Koordinatentransformation $t \rightarrow y, x \rightarrow t$ bzw. $t \rightarrow x, y \rightarrow t$ beide Darstellungen ineinander überführt werden können.

$$t = p \cdot e^{-q \cdot y} \quad t = -\frac{1}{q} \cdot \ln \frac{x}{p}$$

\Rightarrow

$$-\frac{1}{q} \cdot \ln \frac{t}{p} = y \quad p \cdot e^{-q \cdot t} = x$$

Die Parameterdarstellung explizit in das kartesische Koordinatensystem überführt werden.

$$t = -\frac{1}{q} \cdot \ln \frac{x}{p} \quad y = -\frac{1}{q} \cdot \ln \frac{t}{p}$$

\Rightarrow

$$y = -\frac{1}{q} \cdot \ln \left(-\frac{1}{p \cdot q} \cdot \ln \frac{x}{p} \right)$$

Letzterer, funktioneller Ausdruck ist die Arbeitsgleichung der vorliegenden Regression.

1.3 Durchführung der Regression über das Resttermverfahren

Durchführung

Gegeben sei die lineare Funktion:

$$Z = b + a \cdot X$$

Damit werden die Transformatoren mittels Resttermverschub aus der zu regressierenden Arbeitsgleichung gebildet.

$$y = -\frac{1}{q} \cdot \ln \left(-\frac{1}{p \cdot q} \cdot \ln \frac{x}{p} \right)$$

 \Rightarrow

$$e^{-q \cdot y} = \frac{\ln p}{p \cdot q} + \frac{1}{p \cdot q} \cdot \ln \frac{1}{x}$$

Zwischen Term links und Term rechts wird die lineare Funktion als Resttermfunktion eingefügt.

$$e^{-q \cdot y} = Z = b + a \cdot X = \frac{\ln p}{p \cdot q} + \frac{1}{p \cdot q} \cdot \ln \frac{1}{x}$$

Zwischen Z und b wird die Äquivalenzkette aufgebrochen und die sich ergebenden Gleichungen sind die Transformatoren der Urlistwerte.

$$Z = e^{-q^{(a)} \cdot y} \quad b + a \cdot X = \frac{\ln p^{(a)}}{p^{(a)} \cdot q^{(a)}} + \frac{1}{p^{(a)} \cdot q^{(a)}} \cdot \ln \frac{1}{x}$$

Die Koeffizienten \tilde{b} und \tilde{a} können² definiert werden über eine lineare Regression zwischen den Werten $b + a \cdot X_i$ und Z_i . Durch Umstellen sind damit die Endwerte von p und q bekannt.

$$b = \frac{\ln p^{(a)}}{p^{(a)} \cdot q^{(a)}} \quad a = \frac{1}{p^{(a)} \cdot q^{(a)}} > 0 \quad X = -\ln x$$

 \Rightarrow

$$p = e^{\frac{\tilde{b}}{\tilde{a}}} \quad q = \frac{1}{p \cdot \tilde{a}} = \frac{1}{\tilde{a}} \cdot e^{-\frac{\tilde{b}}{\tilde{a}}}$$

Um beschriebene Regression überhaupt „starten“ zu können, sind Abschätzung der Startwerte (a) von p und q nötig - setzen des Grafs mittig in das Intervall $0 < x < 1$:

$$p^{(a)} = MAX(x_i) + MIN(x_i)$$

Für $q^{(a)}$ über die geschätzte Nullstelle x_0

$$0 = y_0 = -\frac{1}{q^{(a)}} \cdot \ln \left(-\frac{1}{p^{(a)} \cdot q^{(a)}} \cdot \ln \frac{x_0}{p^{(a)}} \right)$$

 \Rightarrow

$$x_0 = p^{(a)} \cdot e^{-p^{(a)} \cdot q^{(a)}}$$

 \Rightarrow

$$q^{(a)} = -\frac{1}{p^{(a)}} \cdot \ln \frac{x_0}{p^{(a)}}$$

Damit ist die Regression durchführbar und abgeschlossen.

²Funktionsgleichung der linearen Regression: $Z_i = \tilde{b} + \tilde{a} \cdot (b + a \cdot X_i)$

1.4 Der Restterm als Fehlerwertkontrolle

Vorliegende Regression ist fehlerbehaftet. Allein schon die Abschätzung der Nullstelle der Urlistwerte ist unsicher. Daher ist es vorteilhaft eine Ermittlung des zu erwartenden Fehlers durch zu führen. Methoden und Möglichkeiten sind in der Statistik an betreffender Stelle einsehbar. Oftmals sind die dort beschriebenen Vorgehensweisen komplex und zeitaufwändig. Abhilfe und Vereinfachung kann ein Ermitteln des (Gesamt)Restwertes sein. Dieser ist vorliegend allgemein definiert als:

$$R_i = \frac{b + a \cdot X_i}{Z_i}$$

Die Restwerte werden in das erste und zweite Zentrale Moment überführt, um sie statistisch auswerten zu können. Diese sind definiert durch:

$$ZM^{(k)} = E\left((X - \mu)^k\right) \quad ZM^{|k|} = E\left(|X - \mu|^k\right)$$

Wobei $E(\cdot)$ der Erwartungswert ist, X die Zufallsvariable darstellt und μ der wahre Mittelwert. Von Interesse ist hier nur das erste und zweite Zentrale Moment. Das dritte (Exzess) und das vierte (Wölbung) Moment ist hier nicht von Belang.

- **Das erste Zentrale Moment:**

Das erste Zentrale Moment ist bewiesenermaßen immer Null. Ist dies nicht der Fall, ist entweder μ falsch oder die Zufallsvariablen sind nicht vollständig bekannt. Der Mittelwert im vorliegenden Fall ist definiert mit $\mu = 1 = R$, was aus der Definition des Restterms erfolgt. Damit ist das erste Zentrale Moment ein Gradmesser der Fehlerhaftigkeit der (Restterm)Regression.

$$ZM^{(1)} = \frac{1}{n} \cdot \sum_{i=1}^n (R_i - 1) \rightarrow 0$$

- **Das erste Zentrale absolute Moment:**

$$ZM^{|1|} = \frac{1}{n} \cdot \sum_{i=1}^n |R_i - 1| \rightarrow 0$$

Hierbei handelt es sich um die mittlere absolute Abweichung bezüglich des Restwertes. Sind der Durchschnitt der einzelnen Restwerte gleich dem Median und sind die einzelnen Restwerte gleichzeitig normalverteilt, muss zusätzlich gelten:

$$ZM^{|1|} = \sqrt{\frac{2}{\pi}} \cdot ZM^{(2)}$$

⇒

$$\sqrt{\frac{2}{\pi} \cdot \frac{ZM^{(2)}}{(ZM^{|1|})^2}} = 1$$

- **Das zweite Zentrale (absolute) Moment:**

Ist laut Definition in der Statistik die Varianz der Zufallsvariablen und damit das Quadrat der Standardabweichung. Wobei das zweite Zentrale und das zweite Zentrale absolute Moment gleiche Ergebnisse liefert.

$$ZM^{(2)} \equiv \text{Var}(R_i) \equiv \sigma_i^2 = \frac{1}{n} \cdot \sum_{i=1}^n (R_i - 1)^2 = \frac{1}{n} \cdot \sum_{i=1}^n |R_i - 1|^2 \rightarrow 0$$

1.5 Der Pearson- Korrelationskoeffizient ersten Grades - $\rho^{(1)}$

Korrelationskoeff. Ist ermittelbar über den Anstieg a des Restterms.

$$a = \rho^{(1)} \cdot \frac{\sigma_Y}{\sigma_X}$$

Mit:

$$a > 0$$

\Rightarrow

$$\rho^{(1)} > 0$$

Bedeutet im Konkreten, dass nur positiv linear korrelierende Datenpaare nach vorliegender Vorschrift regressiert werden können. Gegebenenfalls ist dies vorher zu überprüfen.

Daher:

$$0 < \rho^{(1)} \leq +1$$

\Rightarrow

$$0 < a \cdot \sigma_X \leq \sigma_Y$$

Über eine andere allgemeine Berechnungsgrundlage des Korrelationskoeffizienten folgt:

$$\rho^{(1)} = \frac{\text{Cov}(X, Y)}{\sigma_X \cdot \sigma_Y}$$

\Rightarrow

$$0 < \text{Cov}(X, Y) \leq \sigma_X \cdot \sigma_Y$$

Da $\sigma_X > 0$ und $\sigma_Y > 0$ gilt, muss sein:

$$\text{Cov}(X, Y) > 0$$

1.6 Definition der Koeffizienten

1.6.1 Definition von q

Aus $+\infty > q > \frac{1}{t} \cdot \ln p$ und $+\infty > t > 0$ folgt, dass gilt:

Koeffizienten

$$q > \frac{1}{t} \cdot \ln p > 0$$

\Rightarrow

$$q > 0$$

1.6.2 Definition von p

Aus der Arbeitsgleichung sind die Anforderungen für p direkt ablesbar.

$$y = -\frac{1}{q} \cdot \ln \left(-\frac{1}{p \cdot q} \cdot \ln \frac{x}{p} \right)$$

\Rightarrow

$$0 < -\frac{1}{p \cdot q} \cdot \ln \frac{x}{p} < +\infty$$

\Rightarrow

$$p > x > 0$$

\Rightarrow

$$p > 0$$

1.6.3 Der Definitionsbereich der Arbeitsgleichung

$$0 < x < +1$$

\Rightarrow

$$0 < p \cdot e^{-q \cdot t} < +1$$

\Rightarrow

$$+\infty > q > \frac{1}{t} \cdot \ln p$$

\Rightarrow

$$+\infty > t > 0$$

1.6.4 Der Wertebereich der Arbeitsgleichung

$$-\infty < y < +\infty$$

\Rightarrow

$$-\infty < -\frac{1}{q} \cdot \ln \frac{t}{p} < +\infty$$

\Rightarrow

$$+\infty > t > 0$$

2 Beispiele für eine Regression

2.1 Beispiel 1 – Große x_i - und y_i - Werte

Beispiel 1

Urlistwerte, transformierte Werte und Koeffizienten der linearen Regression als Ergebnis der Abzissenwerte $b + a \cdot X_i$ und der Ordinatenwerte Z_i .³

$$b + a \cdot X_i = \frac{\ln p^{(a)}}{p^{(a)} \cdot q^{(a)}} + \frac{1}{p^{(a)} \cdot q^{(a)}} \cdot \ln \frac{1}{x_i} \quad Z_i = e^{-q^{(a)} \cdot y_i}$$

⇒

x_i	$b + a \cdot X_i$	y_i	Z_i	\tilde{b}	\tilde{a}
100	2,371	-1000	2,508	0,284	0,926
500	0,780	+100	0,912		
1000	0,094	+900	0,437		

Vorher gewählte Schätzparameter:

$$\begin{aligned} p^{(a)} &= 1000 + 100 = 1100 \\ x_0^{(a)} &= 400 \\ q^{(a)} &= 0,92 \cdot 10^{-3} \end{aligned}$$

Die endgültigen Koeffizienten aus \tilde{a} und \tilde{b} berechnet.

$$\begin{aligned} p &= 1,359 \\ x_0 &= 0,461 \\ q &= 0,795 \end{aligned}$$

Die vorläufige Regressionsgleichung.

$$\tilde{y} = -1,258 \cdot \ln(-0,926 \cdot \ln(0,736 \cdot \tilde{x}))$$

Einfügen der Schätzparameter und endgültige Regressionsgleichung.

$$\tilde{y} \cdot p^{(a)} = -1,258 \cdot \ln\left(-0,926 \cdot \ln \frac{0,736 \cdot \tilde{x}}{p^{(a)}}\right)$$

⇒

$$y^{(r)} = -1383,648 \cdot \ln(-0,926 \cdot \ln(0,000668 \cdot x))$$

Mit dem Definitionsbereich:

$$-\infty < y \cdot p^{(a)} = -\frac{1}{q} \cdot \ln\left(-\frac{1}{p \cdot q} \cdot \ln \frac{x}{p^{(a)} \cdot p}\right) < +\infty$$

⇒

$$0 < x < p^{(a)} \cdot p$$

⇒

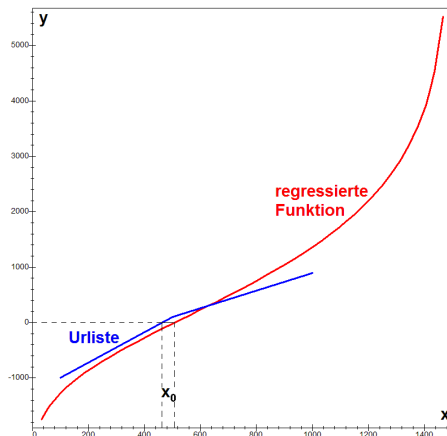
$$0 < x < 1494,9$$

Numerische und grafische Darstellung.

x_i	y_i	$y_i^{(r)}$
100	-1000	-1271
500	+100	-21
1000	+900	+1041

³Funktionsgleichung der linearen Regression: $Z_i = \tilde{b} + \tilde{a} \cdot (b + a \cdot X_i)$

⇒



Der lineare Korrelationskoeffizient nach Pearson mit dem Kurzbezeichner $X_i \leftarrow b + a \cdot \tilde{X}_i$.

i	X_i	Z_i	$X_i - X_M$	$Z_i - Z_M$	$(X_i - X_M) \cdot (X_i - X_M)$	$(Z_i - Z_M) \cdot (Z_i - Z_M)$	$(X_i - X_M) \cdot (Z_i - Z_M)$
1	2,371	2,508	+1,289	+1,222	1,662	1,493	1,575
2	0,780	0,912	-0,302	-0,374	0,091	0,140	0,113
3	0,094	0,437	-0,988	-0,849	0,976	0,721	0,839
$n = 3$	Σ = 3,245	Σ = 3,857	Σ = -0,001	Σ = -0,001	Σ = 2,729	Σ = 2,354	Σ = 2,527

⇒

$$X_M = \frac{3,245}{3} = 1,082 \quad Y_M = \frac{3,857}{3} = 1,286$$

⇒

$$\sigma_X = \sqrt{\frac{2,729}{3}} = 0,954 \quad \sigma_Y = \sqrt{\frac{2,354}{3}} = 0,886$$

⇒

$$\text{Cov}(X, Y) = \frac{2,527}{3} = 0,842 > 0$$

⇒

$$\rho^{(1)} = \frac{0,842}{0,954 \cdot 0,886} = 0,996 > 0$$

Und:

$$0 < a \cdot \sigma_X = 0,926 \cdot 0,954 = 0,883 < 0,886 = \sigma_Y$$

Sowie:

$$0 < \text{Cov}(X, Y) \leq \sigma_X \cdot \sigma_Y$$

⇒

$$0 < 0,842 < 0,954 \cdot 0,886 = 0,845$$

Die Summe des Restterms:

i	X_i	Z_i	$\frac{X_i}{Z_i}$	$\frac{X_i}{Z_i} - 1$	$\left \frac{X_i}{Z_i} - 1 \right $	$\left(\frac{X_i}{Z_i} - 1 \right)^2$
1	2,371	2,508	0,945	-0,055	0,055	0,003
2	0,780	0,912	0,855	-0,145	0,145	0,021
3	0,094	0,437	0,215	-0,785	0,785	0,616
$n = 3$	Σ = 3,245	Σ = 3,857	Σ = 2,015	Σ = -0,985	Σ = 0,985	Σ = 0,640

⇒

$$ZM^{|1|} \equiv R_{MAX} = \frac{0,985}{3} = 0,328$$

Sowie:

$$ZM^{(1)} = -\frac{0,985}{3} = -0,328$$

Und:

$$ZM^{(2)} = \frac{0,640}{3} = 0,213$$

⇒

$$\sigma_i = \sqrt{\frac{0,640}{3}} = 0,462$$

Schlussendlich:

$$\sqrt{\frac{2}{\pi} \cdot \frac{ZM^{(2)}}{(ZM^{|1|})^2}} = 1$$

⇒

$$\sqrt{\frac{2}{\pi} \cdot 3 \cdot \frac{0,640}{0,985^2}} = 1,122 \approx 1$$

2.2 Beispiel 2 – Kleine x_i - und y_i - Werte

Urlistwerte, transformierte Werte und Koeffizienten der linearen Regression als Ergebnis der Abzissenwerte $b + a \cdot X_i$ und der Ordinatenwerte Z_i .⁴ Beispiel 2

$$b + a \cdot X_i = \frac{\ln p^{(a)}}{p^{(a)} \cdot q^{(a)}} + \frac{1}{p^{(a)} \cdot q^{(a)}} \cdot \ln \frac{1}{x_i} \quad Z_i = e^{-q^{(a)} \cdot y_i}$$

⇒

x_i	$b + a \cdot X_i$	y_i	Z_i	\tilde{b}	\tilde{a}
0,11	2,409	-0,85	2,179	0,400	0,868
0,19	1,812	-0,46	1,524		
0,30	1,314	-0,18	1,179		
0,41	0,973	+0,09	0,921		
0,49	0,779	+0,37	0,712		
0,61	0,539	+0,66	0,546		
0,70	0,389	+1,04	0,386		
0,79	0,257	+1,50	0,253		
0,89	0,127	+2,26	0,126		

Vorher gewählte Schätzparameter.

$$\begin{aligned} p^{(a)} &= 0,11 + 0,89 = 1 \\ x_0^{(a)} &= 0,4 \\ q^{(a)} &= 0,916 \end{aligned}$$

Die endgültigen Koeffizienten aus \tilde{a} und \tilde{b} berechnet.

$$\begin{aligned} p &= 1,047 \\ x_0 &= 0,331 \\ q &= 1,100 \end{aligned}$$

Die vorläufige Regressionsgleichung.

$$\tilde{y} = -0,909 \cdot \ln(-0,868 \cdot \ln(0,955 \cdot \tilde{x}))$$

Einfügen der Schätzparameter und endgültige Regressionsgleichung.

$$\tilde{y} \cdot p^{(a)} = -0,909 \cdot \ln\left(-0,868 \cdot \ln \frac{0,955 \cdot \tilde{x}}{p^{(a)}}\right)$$

⇒

$$y^{(r)} = -0,909 \cdot \ln(-0,868 \cdot \ln(0,955 \cdot x))$$

Mit dem Definitionsbereich:

$$-\infty < y \cdot p^{(a)} = -\frac{1}{q} \cdot \ln\left(-\frac{1}{p \cdot q} \cdot \ln \frac{x}{p^{(a)} \cdot p}\right) < +\infty$$

⇒

$$0 < x < p^{(a)} \cdot p$$

⇒

$$0 < x < 1,047$$

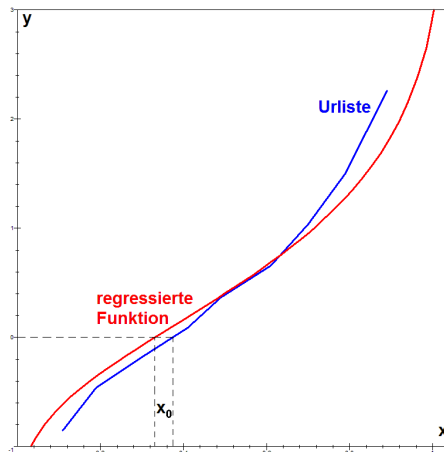
⁴Funktionsgleichung der linearen Regression: $Z_i = \tilde{b} + \tilde{a} \cdot (b + a \cdot X_i)$

2 Beispiele für eine Regression

Numerische und grafische Darstellung.

x_i	y_i	$y_i^{(r)}$
0,11	-0,85	-0,610
0,19	-0,46	-0,357
0,30	-0,18	-0,074
0,41	+0,09	+0,187
0,49	+0,37	+0,379
0,61	+0,66	+0,688
0,70	+1,04	+0,955
0,79	+1,50	+1,280
0,89	+2,26	+1,780

⇒



Der lineare Korrelationskoeffizient nach Pearson mit dem Kurzbezeichner $X_i \leftarrow b + a \cdot \tilde{X}_i$.

i	X_i	Z_i	$X_i - X_M$	$Z_i - Z_M$	$(X_i - X_M) \cdot (X_i - X_M)$	$(Z_i - Z_M) \cdot (Z_i - Z_M)$	$(X_i - X_M) \cdot (Z_i - Z_M)$
1	2,409	2,179	+1,454	+1,309	2,114	1,713	1,903
2	1,812	1,524	+0,857	+0,654	0,734	0,428	0,560
3	1,314	1,179	+0,359	+0,309	0,129	0,095	0,111
4	0,973	0,921	+0,018	+0,051	0,000	0,003	0,001
5	0,779	0,712	-0,176	-0,158	0,031	0,025	0,028
6	0,539	0,546	-0,416	-0,324	0,173	0,105	0,135
7	0,389	0,386	-0,566	-0,484	0,320	0,234	0,274
8	0,257	0,253	-0,698	-0,617	0,487	0,381	0,431
9	0,127	0,126	-0,828	-0,744	0,686	0,554	0,616
$n = 9$	Σ = 8,599	Σ = 7,826	Σ = +0,004	Σ = -0,004	Σ = 4,674	Σ = 3,565	Σ = 4,059

⇒

$$X_M = \frac{8,599}{9} = 0,955 \quad Y_M = \frac{7,826}{9} = 0,870$$

⇒

$$\sigma_X = \sqrt{\frac{4,674}{9}} = 0,721 \quad \sigma_Y = \sqrt{\frac{3,565}{9}} = 0,629$$

⇒

$$\text{Cov}(X, Y) = \frac{4,059}{9} = 0,451 > 0$$

⇒

$$\rho^{(1)} = \frac{0,451}{0,721 \cdot 0,629} = 0,994 > 0$$

Und:

$$0 < a \cdot \sigma_X = 0,868 \cdot 0,721 = 0,626 < 0,629 = \sigma_Y$$

Sowie:

$$0 < \text{Cov}(X, Y) \leq \sigma_X \cdot \sigma_Y$$

⇒

$$0 < 0,451 < 0,721 \cdot 0,629 = 0,454$$

Die Summe des Restterms:

i	X_i	Z_i	$\frac{X_i}{Z_i}$	$\frac{X_i}{Z_i} - 1$	$\left \frac{X_i}{Z_i} - 1 \right $	$\left(\frac{X_i}{Z_i} - 1 \right)^2$
1	2,409	2,179	1,106	+0,106	0,106	0,011
2	1,812	1,524	1,189	+0,189	0,189	0,036
3	1,314	1,179	1,115	+0,115	0,115	0,013
4	0,973	0,921	1,056	+0,056	0,056	0,003
5	0,779	0,712	1,094	+0,094	0,094	0,009
6	0,539	0,546	0,987	-0,013	0,013	0,000
7	0,389	0,386	1,008	+0,008	0,008	0,000
8	0,257	0,253	1,016	+0,016	0,016	0,000
9	0,127	0,126	1,008	+0,008	0,008	0,000
$n = 9$	Σ = 8,599	Σ = 7,826	Σ = 9,579	Σ = +0,579	Σ = 0,605	Σ = 0,072

⇒

$$ZM^{[1]} \equiv R_{MAX} = \frac{0,605}{9} = 0,067$$

Sowie:

$$ZM^{(1)} = \frac{0,579}{9} = 0,064$$

Und:

$$ZM^{(2)} = \frac{0,072}{9} = 0,008$$

⇒

$$\sigma_i = \sqrt{\frac{0,072}{9}} = 0,089$$

Schlussendlich:

$$\sqrt{\frac{2}{\pi} \cdot \frac{ZM^{(2)}}{(ZM^{[1]})^2}} = 1$$

⇒

$$\sqrt{\frac{2}{\pi} \cdot 9 \cdot \frac{0,072}{0,067^2}} = 9,586 \neq 1$$

2.3 Beispiel 3 – Falsche Voraussetzungen für x_i - und y_i - Werte

Beispiel 3

Urlistwerte, transformierte Werte und Koeffizienten der linearen Regression als Ergebnis der Abzissenwerte $b + a \cdot X_i$ und der Ordinatenwerte Z_i .⁵

$$b + a \cdot X_i = \frac{\ln p^{(a)}}{p^{(a)} \cdot q^{(a)}} + \frac{1}{p^{(a)} \cdot q^{(a)}} \cdot \ln \frac{1}{x_i} \quad Z_i = e^{-q^{(a)} \cdot y_i}$$

⇒

x_i	$b + a \cdot X_i$	y_i	Z_i	\tilde{b}	\tilde{a}
0,89	0,252	-0,85	1,481	1,293	-0,224
0,79	0,510	-0,46	1,237		
0,70	0,772	-0,18	1,087		
0,61	1,070	+0,09	0,959		
0,49	1,544	+0,37	0,843		
0,41	1,930	+0,66	0,737		
0,30	2,606	+1,04	0,618		
0,19	3,594	+1,50	0,500		
0,11	4,777	+2,26	0,352		

Widerspricht der Forderung $a > 0$. Es wird dennoch weiter gerechnet als Negativbeispiel.

Vorher gewählte Schätzparameter.

$$\begin{aligned} p^{(a)} &= 0,11 + 0,89 = 1 \\ x_0^{(a)} &= 0,63 \\ q^{(a)} &= 0,462 \end{aligned}$$

Die endgültigen Koeffizienten aus \tilde{a} und \tilde{b} berechnet.

$$\begin{aligned} p &= 0,00314 \\ x_0 &= 0,270 \\ q &= -1419,59 \end{aligned}$$

Spätestens hier widerspricht der ermittelte Wert von $q = -1419,59$ der Forderung $q > 0$ aus den Ausgangsdefinitionen der Koeffizienten. Es wird dennoch weiter gerechnet als Negativbeispiel.

Die vorläufige Regressionsgleichung.

$$\tilde{y} = 0,000704 \cdot \ln(0,224 \cdot \ln(318,574 \cdot \tilde{x}))$$

Einfügen der Schätzparameter und endgültige Regressionsgleichung.

$$\tilde{y} \cdot p^{(a)} = 0,000704 \cdot \ln\left(0,224 \cdot \ln \frac{318,574 \cdot \tilde{x}}{p^{(a)}}\right)$$

⇒

$$y^{(r)} = 0,000704 \cdot \ln(0,224 \cdot \ln(318,574 \cdot x))$$

Mit dem Definitionsbereich:

$$-\infty < y \cdot p^{(a)} = -\frac{1}{q} \cdot \ln\left(-\frac{1}{p \cdot q} \cdot \ln \frac{x}{p^{(a)} \cdot p}\right) < +\infty$$

⇒

$$0 < x < p^{(a)} \cdot p$$

⇒

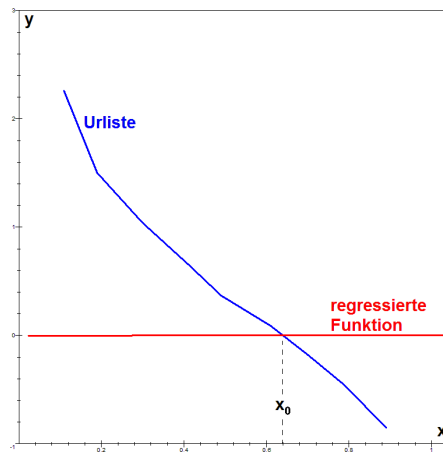
$$0 < x < 0,00314$$

⁵Funktionsgleichung der linearen Regression: $Z_i = \tilde{b} + \tilde{a} \cdot (b + a \cdot X_i)$

Numerische und grafische Darstellung.

x_i	y_i	$y_i^{(r)}$
0,89	-0,85	+0,002
0,79	-0,46	+0,002
0,70	-0,18	+0,001
0,61	+0,09	+0,001
0,49	+0,37	+0,001
0,41	+0,66	+0,001
0,30	+1,04	+0,000
0,19	+1,50	-0,000
0,11	+2,26	-0,002

⇒



Der lineare Korrelationskoeffizient nach Pearson mit dem Kurzbezeichner $X_i \leftarrow b + a \cdot \tilde{X}_i$.

i	X_i	Z_i	$X_i - X_M$	$Z_i - Z_M$	$(X_i - X_M) \cdot (X_i - X_M)$	$(Z_i - Z_M) \cdot (Z_i - Z_M)$	$(X_i - X_M) \cdot (Z_i - Z_M)$
1	0,252	1,481	-1,643	+0,613	2,699	0,376	-1,007
2	0,510	1,237	-1,385	+0,369	1,918	0,136	-0,511
3	0,772	1,087	-1,123	+0,219	1,261	0,048	-0,246
4	1,070	0,959	-0,825	+0,091	0,681	0,008	-0,075
5	1,544	0,843	-0,351	-0,025	0,123	0,001	+0,009
6	1,930	0,737	+0,035	-0,131	0,001	0,017	-0,005
7	2,606	0,618	+0,711	-0,250	0,506	0,063	-0,178
8	3,594	0,500	+1,699	-0,368	2,887	0,135	-0,625
9	4,777	0,352	+2,882	-0,516	8,306	0,266	-1,487
$n = 9$	$\Sigma = 17,055$	$\Sigma = 7,814$	$\Sigma = 0,000$	$\Sigma = 0,002$	$\Sigma = 18,382$	$\Sigma = 1,050$	$\Sigma = 4,125$

⇒

$$X_M = \frac{17,055}{9} = 1,895 \quad Y_M = \frac{7,814}{9} = 0,868$$

2 Beispiele für eine Regression

$$\Rightarrow \sigma_X = \sqrt{\frac{18,382}{9}} = 1,429 \quad \sigma_Y = \sqrt{\frac{1,050}{9}} = 0,342$$

$$\Rightarrow \text{Cov}(X, Y) = -\frac{4,125}{9} = -0,458 \stackrel{!}{<} 0$$

$$\Rightarrow \rho^{(1)} = -\frac{0,458}{1,429 \cdot 0,342} = -0,938 \stackrel{!}{<} 0$$

Und:

$$0 \stackrel{!}{>} a \cdot \sigma_X = -0,224 \cdot 1,429 = -0,320 < 0,342 = \sigma_Y$$

Sowie:

$$0 < \text{Cov}(X, Y) \leq \sigma_X \cdot \sigma_Y$$

\Rightarrow

$$0 \stackrel{!}{>} -0,458 < 1,429 \cdot 0,342 = 0,489$$

Die Summe des Restterms:

i	X_i	Z_i	$\frac{X_i}{Z_i}$	$\frac{X_i}{Z_i} - 1$	$\left \frac{X_i}{Z_i} - 1 \right $	$\left(\frac{X_i}{Z_i} - 1 \right)^2$
1	0,252	1,481	0,170	-0,830	0,830	0,689
2	0,510	1,237	0,412	-0,588	0,588	0,346
3	0,772	1,087	0,710	-0,290	0,290	0,084
4	1,070	0,959	1,116	+0,116	0,116	0,013
5	1,544	0,843	1,832	+0,832	0,832	0,692
6	1,930	0,737	2,619	+1,619	1,619	2,621
7	2,606	0,618	4,217	+3,217	3,217	10,349
8	3,594	0,500	7,188	+6,188	6,188	38,291
9	4,777	0,352	13,571	+12,571	12,571	158,030
$n = 9$	Σ = 17,055	Σ = 7,814	Σ = 13,835	Σ = +22,835	Σ = 26,251	Σ = 211,115

\Rightarrow

$$ZM^{(1)} \equiv R_{MAX} = \frac{26,251}{9} = 2,917$$

Sowie:

$$ZM^{(1)} = \frac{22,835}{9} = 2,537$$

Und:

$$ZM^{(2)} = \frac{211,115}{9} = 23,457$$

\Rightarrow

$$\sigma_i = \sqrt{\frac{211,115}{9}} = 4,843$$

Schlussendlich:

$$\sqrt{\frac{2}{\pi} \cdot \frac{ZM^{(2)}}{(ZM^{(1)})^2}} = 1$$

\Rightarrow

$$\sqrt{\frac{2}{\pi} \cdot 9 \cdot \frac{23,457}{2,917^2}} = 3,974 \neq 1$$

2.4 Beispiel 4 – Durch Modifikation der Arbeitsgleichung das Beispiel 3 korrigiert

Modifikation der Arbeitsgleichung.

Beispiel 4

$$\hat{y} = -\frac{1}{q} \cdot \ln \left(-\frac{1}{p \cdot q} \cdot \ln \frac{\hat{x}}{p} \right)$$

$$\Rightarrow y = -\frac{1}{q} \cdot \ln \left(-\frac{1}{p \cdot q} \cdot \ln \frac{p-x}{p} \right)$$

$$\Rightarrow e^{-q \cdot y} = \frac{1}{p \cdot q} \cdot \ln p + \frac{1}{p \cdot q} \cdot \ln \frac{1}{p-x}$$

Modifikation des Schätzparameters $q^{(a)}$ und Darstellung von x_0 .

$$0 = y_0 = -\frac{1}{q^{(a)}} \cdot \ln \left(-\frac{1}{p^{(a)} \cdot q^{(a)}} \cdot \ln \frac{p^{(a)} - x_0}{p^{(a)}} \right)$$

$$\Rightarrow x_0 = p^{(a)} \cdot \left(1 - e^{-p^{(a)} \cdot q^{(a)}} \right)$$

$$\Rightarrow q^{(a)} = -\frac{1}{p^{(a)}} \cdot \ln \frac{p^{(a)} - x_0}{p^{(a)}}$$

Urlistwerte, transformierte Werte und Koeffizienten der linearen Regression als Ergebnis der Abzissenwerte $b + a \cdot X_i$ und der Ordinatenwerte Z_i .⁶

$$b + a \cdot X_i = \frac{\ln p^{(a)}}{p^{(a)} \cdot q^{(a)}} + \frac{1}{p^{(a)} \cdot q^{(a)}} \cdot \ln \frac{1}{p^{(a)} - x_i} \quad Z_i = e^{-q^{(a)} \cdot y_i}$$

\Rightarrow

x_i	$b + a \cdot X_i$	y_i	Z_i	\tilde{b}	\tilde{a}
0,89	2,221	-0,85	2,328	-0,025	1,038
0,79	1,570	-0,46	1,580		
0,70	1,211	-0,18	1,196		
0,61	0,947	+0,09	0,914		
0,49	0,677	+0,37	0,692		
0,41	0,531	+0,66	0,519		
0,30	0,359	+1,04	0,356		
0,19	0,212	+1,50	0,225		
0,11	0,117	+2,26	0,106		

Vorher gewählte Schätzparameter.

$$p^{(a)} = 0,11 + 0,89 = 1$$

$$x_0^{(a)} = 0,63$$

$$q^{(a)} = 0,994$$

Die endgültigen Koeffizienten aus \tilde{a} und \tilde{b} berechnet.

$$p = 0,976$$

$$x_0 = 0,372$$

$$q = 0,987$$

Die vorläufige Regressionsgleichung.

$$\tilde{y} = -1,013 \cdot \ln(-1,038 \cdot \ln(1 - 1,025 \cdot \tilde{x}))$$

⁶Funktionsgleichung der linearen Regression: $Z_i = \tilde{b} + \tilde{a} \cdot (b + a \cdot X_i)$

Einfügen der Schätzparameter und endgültige Regressionsgleichung.

$$\tilde{y} \cdot p^{(a)} = -1,013 \cdot \ln \left(-1,038 \cdot \ln \frac{1 - 1,025 \cdot \tilde{x}}{p^{(a)}} \right)$$

⇒

$$y^{(r)} = -1,013 \cdot \ln (-1,038 \cdot \ln (1 - 1,025 \cdot x))$$

Mit dem Definitionsbereich:

$$-\infty < y \cdot p^{(a)} = -\frac{1}{q} \cdot \ln \left(-\frac{1}{p \cdot q} \cdot \ln \frac{p - x}{p^{(a)} \cdot p} \right) < +\infty$$

⇒

$$p > x > (1 - p^{(a)}) \cdot p$$

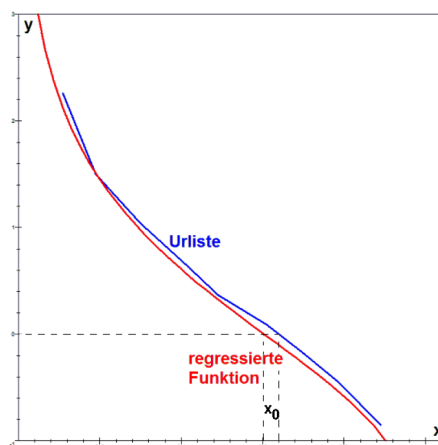
⇒

$$0 < x < 0,976$$

Numerische und grafische Darstellung.

x_i	y_i	$y_i^{(r)}$
0,89	-0,85	-0,937
0,79	-0,46	-0,550
0,70	-0,18	-0,274
0,61	+0,09	-0,018
0,49	+0,37	+0,328
0,41	+0,66	+0,577
0,30	+1,04	+0,977
0,19	+1,50	+1,512
0,11	+2,26	+2,114

⇒



Der lineare Korrelationskoeffizient nach Pearson mit dem Kurzbezeichner $X_i \leftarrow b + a \cdot \tilde{X}_i$.

i	X_i	Z_i	$X_i - X_M$	$Z_i - Z_M$	$(X_i - X_M) \cdot (X_i - X_M)$	$(Z_i - Z_M) \cdot (Z_i - Z_M)$	$(X_i - X_M) \cdot (Z_i - Z_M)$
1	2,221	2,328	+1,349	+1,448	1,820	2,097	1,953
2	1,570	1,580	+0,698	+0,700	0,487	0,490	0,489
3	1,211	1,196	+0,339	+0,316	0,115	0,100	0,107
4	0,947	0,914	+0,075	+0,034	0,006	0,001	0,003
5	0,677	0,692	-0,195	-0,188	0,038	0,035	0,037
6	0,531	0,519	-0,341	-0,361	0,116	0,130	0,123
7	0,359	0,356	-0,513	-0,524	0,263	0,275	0,269
8	0,212	0,225	-0,660	-0,655	0,436	0,430	0,432
9	0,117	0,106	-0,755	-0,774	0,570	0,599	0,584
$n = 9$	Σ = 7,845	Σ = 7,916	Σ = -0,003	Σ = -0,004	Σ = 3,851	Σ = 4,157	Σ = 3,997

⇒

$$X_M = \frac{7,845}{9} = 0,872 \quad Y_M = \frac{7,916}{9} = 0,880$$

⇒

$$\sigma_X = \sqrt{\frac{3,851}{9}} = 0,654 \quad \sigma_Y = \sqrt{\frac{4,157}{9}} = 0,680$$

⇒

$$\text{Cov}(X, Y) = \frac{3,997}{9} = 0,444 > 0$$

⇒

$$\rho^{(1)} = -\frac{0,444}{0,654 \cdot 0,680} = 0,999 > 0$$

Und:

$$0 < a \cdot \sigma_X = 1,038 \cdot 0,654 = 0,679 < 0,680 = \sigma_Y$$

Sowie:

$$0 < \text{Cov}(X, Y) \leq \sigma_X \cdot \sigma_Y$$

⇒

$$0 < 0,444 < 0,654 \cdot 0,680 = 0,445$$

Die Summe des Restterms:

i	X_i	Z_i	$\frac{X_i}{Z_i}$	$\frac{X_i}{Z_i} - 1$	$\left \frac{X_i}{Z_i} - 1 \right $	$\left(\frac{X_i}{Z_i} - 1 \right)^2$
1	2, 221	2, 328	0, 954	-0, 046	0, 046	0, 002
2	1, 570	1, 580	0, 994	-0, 006	0, 006	0, 000
3	1, 211	1, 196	1, 013	+0, 013	0, 013	0, 000
4	0, 947	0, 914	1, 036	+0, 036	0, 036	0, 001
5	0, 677	0, 692	0, 978	-0, 022	0, 022	0, 000
6	0, 531	0, 519	1, 023	+0, 023	0, 023	0, 000
7	0, 359	0, 356	1, 008	+0, 008	0, 008	0, 000
8	0, 212	0, 225	0, 942	-0, 058	0, 058	0, 003
9	0, 117	0, 106	1, 104	+0, 104	0, 104	0, 011
$n = 9$	Σ = 7, 845	Σ = 7, 916	Σ = 9, 052	Σ = +0, 052	Σ = 0, 316	Σ = 0, 017

⇒

$$ZM^{|1|} \equiv R_{MAX} = \frac{0,316}{9} = 0,035$$

Sowie:

$$ZM^{(1)} = \frac{0,052}{9} = 0,006$$

Und:

$$ZM^{(2)} = \frac{0,017}{9} = 0,002$$

⇒

$$\sigma_i = \sqrt{\frac{0,017}{9}} = 0,043$$

Schlussendlich:

$$\sqrt{\frac{2}{\pi} \cdot \frac{ZM^{(2)}}{(ZM^{|1|})^2}} = 1$$

⇒

$$\sqrt{\frac{2}{\pi} \cdot 9 \cdot \frac{0,017}{0,035^2}} = 8,917 \neq 1$$

2.5 Beispiel – Zusammenfassung der statistischen Werte der Beispiele 1 - 4

Es werden die relevanten, berechneten, statistischen Werte gesammelt und nochmals dargestellt.

Zusammenfassung

Beispiel	σ_X	σ_Y	Cov (X_i, Y_i)	$\rho^{(1)}$	$ZM^{(1)}$ \equiv R_{MAX}	$ZM^{(1)}$ \equiv \ominus	$\sqrt{ZM^{(2)}}$ \equiv σ_i
1	0,954	0,886	+0,842	+0,996	0,328	-0,328	0,462
2	0,721	0,629	+0,451	+0,994	0,067	+0,064	0,089
3	1,429	0,342	-0,458	-0,938	2,917	+2,537	4,843
4	0,654	0,680	+0,444	+0,999	0,035	+0,006	0,043

3 Zusammenfassung der Berechnungsgrundlagen

Zusammenfassung Zum Schluss sollen die notwendigen Berechnungsgrundlagen aufgelistet werden.

3.1 Bilogarithmisch monoton steigend

Parametrische Darstellung der Arbeitsgleichung

$$x(t) = p \cdot e^{-q \cdot t} \quad y(t) = -\frac{1}{q} \cdot \ln \frac{t}{p}$$

Kartesische Darstellung der Arbeitsgleichung

$$y(x) = -\frac{1}{q} \cdot \ln \left(-\frac{1}{p \cdot q} \cdot \ln \frac{x}{p} \right)$$

Resttermdarstellung der Arbeitsgleichung - die Transformatoren

$$(b + a \cdot X_1)(x) = \frac{\ln p^{(a)}}{p^{(a)} \cdot q^{(a)}} + \frac{1}{p^{(a)} \cdot q^{(a)}} \cdot \ln \frac{1}{x_i} \quad Z_i(y) = e^{-q^{(a)} \cdot y_i}$$

Definition der Startkoeffizienten (a) , x_0 wird geschätzt oder andersweitig ermittelt

$$p^{(a)} = MAX(x_i) + MIN(x_i)$$

$$x_0 = p^{(a)} \cdot e^{-p^{(a)} \cdot q^{(a)}}$$

$$q^{(a)} = -\frac{1}{p^{(a)}} \cdot \ln \frac{x_0}{p^{(a)}}$$

Durchführung der Linearen Regression, Ermittlung der Koeffizienten \tilde{a} und \tilde{b}

$$Z_i = \tilde{b} + \tilde{a} \cdot (b + a \cdot X_i)$$

Definition der endgültigen Koeffizienten p und q

$$p = e^{\frac{\tilde{b}}{\tilde{a}}}$$

$$q = \frac{1}{p \cdot \tilde{a}}$$

$$x_0 = p \cdot e^{-p \cdot q}$$

Der Definitionsbereich

$$0 < x < p^{(a)} \cdot p$$

Der Restterm als Fehlerwertkontrolle, die maximale Abweichung (Betrag aller Abweichungen)

$$ZM^{(1)} \equiv R_{Max} = \frac{1}{n} \cdot \sum_{i=1}^n \left| \frac{b + a \cdot X_i}{Z_i} - 1 \right| \rightarrow 0$$

Das erste und zweite Zentrale Moment als Fehlerwertkontrolle

$$ZM^{(1)} = \frac{1}{n} \cdot \sum_{i=1}^n \left(\frac{b + a \cdot X_i}{Z_i} - 1 \right) \rightarrow 0$$

$$ZM^{(2)} \equiv \sigma = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^n \left(\frac{b + a \cdot X_i}{Z_i} - 1 \right)^2} \rightarrow 0$$

3.2 Bilogarithmisch monoton fallend

Parametrische Darstellung der Arbeitsgleichung

Zusammenfassung

$$x(t) = p \cdot (1 - e^{-q \cdot t}) \quad y(t) = -\frac{1}{q} \cdot \ln \frac{t}{p}$$

Kartesische Darstellung der Arbeitsgleichung

$$y(x) = -\frac{1}{q} \cdot \ln \left(-\frac{1}{p \cdot q} \cdot \ln \frac{p-x}{p} \right)$$

Resttermdarstellung der Arbeitsgleichung - die Transformatoren

$$(b + a \cdot X_i)(x) = \frac{\ln p^{(a)}}{p^{(a)} \cdot q^{(a)}} + \frac{1}{p^{(a)} \cdot q^{(a)}} \cdot \ln \frac{1}{p^{(a)} - x_i} \quad Z_i(y) = e^{-q^{(a)} \cdot y_i}$$

Definition der Startkoeffizienten (a) , x_0 wird geschätzt oder andersweitig ermittelt

$$p^{(a)} = MAX(x_i) + MIN(x_i)$$

$$x_0 = p^{(a)} \cdot (1 - e^{-p^{(a)} \cdot q^{(a)}})$$

$$q^{(a)} = -\frac{1}{p^{(a)}} \cdot \ln \frac{p^{(a)} - x_0}{p^{(a)}}$$

Durchführung der Linearen Regression, Ermittlung der Koeffizienten \tilde{a} und \tilde{b}

$$Z_i = \tilde{b} + \tilde{a} \cdot (b + a \cdot X_i)$$

Definition der endgültigen Koeffizienten p und q

$$p = e^{\frac{\tilde{b}}{\tilde{a}}}$$

$$q = \frac{1}{p \cdot \tilde{a}}$$

$$x_0 = p \cdot (1 - e^{-p \cdot q})$$

Der Definitionsbereich

$$(1 - p^{(a)}) \cdot p < x < p$$

Der Restterm als Fehlerwertkontrolle, die maximale Abweichung (Betrag aller Abweichungen)

$$ZM^{(1)} \equiv R_{Max} = \frac{1}{n} \cdot \sum_{i=1}^n \left| \frac{b + a \cdot X_i}{Z_i} - 1 \right| \rightarrow 0$$

Das erste und zweite Zentrale Moment als Fehlerwertkontrolle

$$ZM^{(1)} = \frac{1}{n} \cdot \sum_{i=1}^n \left(\frac{b + a \cdot X_i}{Z_i} - 1 \right) \rightarrow 0$$

$$ZM^{(2)} \equiv \sigma = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^n \left(\frac{b + a \cdot X_i}{Z_i} - 1 \right)^2} \rightarrow 0$$

L^AT_EX

