ACTUARIAL RESEARCH CLEARING HOUSE
1997 VOL. 1

GENERATING RANDOM VARIATES WITH A GIVEN FORCE OF MORTALITY AND FINDING A SUITABLE FORCE OF MORTALITY BY THEORETICAL QUANTILE-QUANTILE PLOTS

Jeffrey S. Pai [1]

Warren Centre for Actuarial Studies & Research, University of Manitoba

Winnipeg, Manitoba R3T 5V4 Canada

Abstract: The problem of the computer generation of random variables with a given force of mortality can be done by applying the connection between the cumulative distribution function $F(t)$ and the cumulative force of mortality $M(t)$. For generating a random variate with a given cumulative mortality, it suffices to invert an exponential random variate E. This inversion method works well for most of the mortality laws. For Makeham's law, we suggest the composition method which requires only three simple steps. Based on the relationship between $F(t)$ and $M(t)$, we also suggest a graphical approach to find a suitable force of mortality for a set of survival time. Several graphical examples are presented to illustrate our methodology under different laws such as Pareto, Weibull, de Moivre, Gompertz and Makeham.

[1] Room 640, Drake Centre, 181 Freedman crescent, Winnipeg, Manitoba R3T 5V4 Canada. E-mail: jpai@cc.umanitoba.ca

1. INTRODUCTION

The analytical laws of mortality such as Gompertz and Makeham are often used by actuaries to describe the survival distributions. Numerical approaches to evaluate features such as mean and variance of these distributions may be very difficult not to mention the characteristics of their life contingent functions.

The use of sampling based methods is an alternative to the numerical approaches. Let $t_1, \cdots, t_J$ denote samples generated from a probability density function, say $f(t)$. The mean of any function of t, say $E[g(t)]$, can by approximated by the ergodic average

$$\int g(t)f(t)\mathrm{d}t \approx \frac{1}{J}\sum_{j=1}^{J} g(t_j). \tag{1}$$

we note that the right hand side converges to the left hand side of (1) in probability as $J \to \infty$. For example, let $f(t)$ be the probability density function of the random variable $T(x)$ (the future lifetime for a person of age x) and $g(t) = \exp[-\delta T(x)]$ be the present value random variable of the benefit of a whole life insurance policy. δ denotes the force of interest.

Consider the relationship between the cumulative distribution function, say $F(t)$, and the force of mortality, say $\mu(t)$ as follows:

$$F(t) \quad = \quad 1 - \exp[-M(t)], \quad \text{where} \quad M(t) = \int_0^t \mu(s)\mathrm{d}s. \tag{2}$$

This connection helps us to generate random variates through the use of exponential random variates. Section 2 discusses the generation of random variates with a given force of mortality by the inversion method. We also suggest the Newton-Raphson method, the thinning method, and the composition method in the case of the Makeham's law.

Section 3 presents a type of graphical methods for data analysis. We suggest a theoretical quantile-quantile plot (Q-Q plot) to explore the observed data and to find a suitable force of mortality for the data. We present several graphical examples to illustrate our methodology.

2. GENERATION OF RANDOM VARIATES

From (2) we note that $M(t)$ has an exponential probability function (with mean 1). Thus, for generating a random variate with a given cumulative force of mortality it suffices to invert an exponential random variate.

Inversion method (Cinlar, 1975)

(i) Generate E from the exponential distribution.

(ii) Calculate $T = M^{-1}(E)$.

Table 1 summarizes the result from the inversion method for several useful laws of force of mortality.

2.1. Newton-Raphson Iterations

In the case of the Makeham's law,

$$M(t) = at + \frac{bc^x(c^t - 1)}{\ln(c)},$$

where $M^{-1}(E)$ is not explicitly known, we can solve $M(T) = E$ for T by Newton-Raphson method as follows.

Inversion method using Newton-Raphson iterations

(i) Generate E from the exponential distribution

(ii) Set $T_0 = 0$

(iii) Calculate $\Delta_i = \frac{E - aT_i - bc^x(c^{T_i} - 1)/\ln(c)}{a + bc^{x+T_i}}$

(iv) Calculate $T_{i+1} = T_i + \Delta_i$

(v) Retain T_{i+1} if $|\Delta_i| < \epsilon$, go to (iii) otherwise

An appropriate stopping rule, say ϵ, must be added in practical applications. The computing time depends on the starting value, T_0, and the precision required, ϵ. A better starting value can be obtained by taking the second-degree Taylor polynomial of c^t. That is,

$$M(t) = at + \frac{bc^x}{\ln(c)}[1 + \ln(c)t + \frac{(\ln(c))^2}{2}t^2 - 1]. \tag{3}$$

The positive root of the quadratic equation is

$$T_0 = \frac{[(a + bc^x)^2 + 2Ebc^x \ln(c)]^{1/2} - a - bc^x}{bc^x \ln(c)}. \tag{4}$$

2.2. Thinning Method

Not like the Newton-Raphson method, the thinning method doesn't need any stopping rule. Suppose that we can find a force of mortality, say $\mu_0(t)$, such that $\mu(t) \leq \mu_0(t)$ for all t and the sampling with respect to $\mu_0(t)$ is easy and computational inexpensive, we can apply the thinning method to Makeham's law. For example, let $\mu_0(t) = (a/c^x + b)c^{x+t}$.

Thinning method (Lewis and Shedler, 1977)

(i) Set $T_0 = 0$

(ii) Generate Δ_i from the Gompertz's law

with parameters $b_0 = a/c^x + b$, $c_0 = c$ and $x_0 = x + T_i$

(iii) Generate U from a uniform distribution (0,1)

(iv) Calculate $T_{i+1} = T_i + \Delta_i$

(v) Retain T_{i+1} if $U \leq \frac{a+bc^{x+T_{i+1}}}{ac^{T_{i+1}}+bc^{x+T_{i+1}}}$, go to (ii) otherwise

2.3. Composition Method

Both the Newton-Raphson method and the thinning method are iterative method which is, sometimes, not satisfactory in terms of the computing time and the precision. A non-iterative approach is suggested by using the composition method. Let's first decompose the Makeham's law, $\mu(t)$, to two parts, $\mu_1(t)$ and $\mu_2(t)$. That is, $\mu(t) = \mu_1(t) + \mu_2(t)$, where $\mu_1(t) = bc^{x+t}$ has Gompertz's law and $\mu_2(t) = a$ is a small adjustment to $\mu_1(t)$.

Composition method (Devroye, 1986, page 263)

(i) Generate G from the Gompertz's law (see Table 1)

(ii) Generate E from the exponential distribution

(iii) Retain $T = \text{Min}[E/a, G]$

Proof:

$$\begin{aligned} Pr[T > t] &= Pr[\text{Min}[E/a, G] > t] \\ &= Pr[\{E > at\} \text{ and } \{G > t\}] \\ &= Pr[E > at]Pr[G > t] \\ &= exp[-at - \frac{bc^x(c^t - 1)}{\ln(c)}]. \end{aligned}$$

We note that only two exponential random variates are needed for generating one random variate from the Makeham's law.

As an application, consider an annuity payable continuously at the rate of 1 per year as long as at least one of three lives, say (w), (y), and (z), survives (the last-survival status). Suppose that (w=30), (y=40), and (z=50) follow the Makeham's law, the Gomperzt's law, and the Weibull's law respectively, and their future lifetimes are mutually independent. We first draw 3 samples independently from the three analytical laws of mortality with age $x = 30$, 40, and 50 respectively. We keep the largest value T and calculate the present value $g = (1 - exp[-\delta T])/\delta$. We repeat J times to have a total of J samples, say $g_j, j = 1, \cdots, J$. Then the expected value and the variance of the present value random variable can be approximated by the sample mean and variance of $g_j, j = 1, \cdots, J$.

The composition method can be used for any revised law. For example, $\mu_1(t)$ could be a Pareto's or Weibull's law while $\mu_2(t)$ is a constant adding an extra risk to $\mu_1(t)$. One can also apply this approach to a life having an additional risk to an analytical law in certain age interval. For example, let $\mu_1(t)$ be a Makeham's law (which has been found suitable for adult ages) and let $\mu_2(t)$ be a constant if $t < 13$, 0 otherwise.

3. THEORETICAL QUANTILE-QUANTILE PLOTS

Let us suppose that $t_1, \cdots, t_n$ are independent identically distributed (i.i.d.) random variables from the probability density function with the force of mortality $\mu(t)$. Let $s_1, \cdots, s_n$ be the value of data sorted from smallest to largest, so that s_i is the $p_i = (i - 0.5)/n$ empirical quantile. The relationship between the cumulative force of mortality and the cumulative probability distribution can be written as

$$M(s_i) \quad = \quad -\ln[1 - F(s_i)] \quad (i = 1, \cdots, n). \tag{5}$$

If we estimate $F(s_i)$ by p_i, we have $m_i = \hat{M}(s_i) = \ln[\frac{n}{n-i+0.5}]$. We expect to see, roughly, a 45^0 line pass the origin if we plot $M(s_i)$ versus m_i under the situation where the data are truly from the underlying distribution. Table 2 shows the Q-Q plots under different laws. The following simulation results are carried out as stated in Section 2 with sample size 10000.

3.1. Pareto's law $M(t) = a\ln(1 + x + t)$

PLOT: $\ln(1 + x + s_i)$ versus $\ln[\frac{n}{n-i+0.5}]$

EXPECT: a straight line cross the origin with slope a

RESULT: Figure 1 shows the result with $a = 0.5$, $x = 50$ and a line $Y = 0.5x$.

3.2. de Moivre's law $M(t) = a \ln(\frac{\omega - x}{\omega - x - t})$

PLOT: $\ln(\omega - x - s_i)$ versus $\ln[\frac{n}{n-i+0.5}]$

EXPECT: a line with intercept $a \ln(\omega - x)$ and slope $-a$

RESULT: Figure 2 shows the result with $a = 1.1$, $\omega = 110$, $x = 50$ and a line $Y = a\ \ln(\omega - x) - aX$.

3.3. Weibull's law $M(t) = \frac{a[(x+t)^{b+1} - x^{b+1}]}{b+1}$

The Weibull's law can be written as

$$\ln(M(t) + \frac{ax^{b+1}}{b+1}) = \ln(\frac{a}{b+1}) + (b+1)\ln(t+x).$$

If we plot $\ln(s_i + x)$ versus $\ln(m_i + \frac{ax^{b+1}}{b+1})$ we will see a regression line with intercept $\ln(\frac{a}{b+1})$ and slope $(b+1)$. We can initiate the plot by letting $\frac{ax^{b+1}}{b+1} = 0$ and estimate it by the least square estimates, $Y = \hat{A} + \hat{B}X$. That is,

$$\frac{ax^{b+1}}{b+1} \approx x^{\hat{B}} e^{\hat{A}}.$$

After a few iterations, we should get a satisfactory line.

PLOT: $\ln(s_i + x)$ versus $\ln(\ln[\frac{n}{n-i+0.5}] + x^{\hat{B}} e^{\hat{A}})$

EXPECT: a line with intercept $\ln(\frac{a}{b+1})$ and slope $(b+1)$.

RESULT: Figure 3 shows the plots and the line $Y = \ln(\frac{a}{b+1}) + (b+1)X$ for 4 iterations.

3.4. Gompertz's law $M(t) = \frac{bc^x(c^t-1)}{\ln(c)}$

Gompertz's law can be written as

$$\ln(M(t) + \frac{bc^x}{\ln(c)}) = \ln(\frac{bc^x}{\ln(c)}) + t\ \ln(c).$$

If we plot s_i versus $\ln(m_i + \frac{bc^x}{\ln(c)})$ we will see a linear relationship for large t. Several iterations might be needed to update $\frac{bc^x}{\ln(c)}$ as stated in the previous subsection.

PLOT: s_i versus $\ln(\ \ln[\frac{n}{n-i+0.5}] + \frac{bc^x}{\ln(c)})$

EXPECT: a line with intercept $\ln(\frac{bc^x}{\ln(c)})$ and slope $\ln(c)$

RESULT: Figure 4 shows the plots with $b = 0.000007$, $c = 1.12$, $x = 50$ and a line $Y = \ln(\frac{bc^x}{\ln(c)}) + \ln(c)X$.

Same plot can be used to the Makeham's law for small values of a (usually the case). Figure 5 shows the plot with $a = 0.001$, $b = 0.000007$, $c = 1.12$, $x = 50$ and a line $Y = \ln(\frac{bc^x}{\ln(c)}) + \ln(c)X$.

4. Conclusion

We have recognized several methods to draw samples form Makeham's law. While the adaptive rejection method that Scollnik (1995) suggested works for all log-concave densities, the composition method we suggest is simpler and more powerful especially when a very large sample size is required to achieve the desired accuracy. Sampling methods can be used easily to solve many complicated actuarial functions. The last-survival example in Section 2 is just one of them. In terms of estimation, London (1988) has suggested several useful methods to parametric survival models. All those methods can be used only when the certain law has been identified. Our method by Q-Q plots is a preliminary study to explore data and, hopefully, identify a suitable law visually.

References:

1. Cinlar, E. (1975): Introduction to stochastic processes. Prentice-Hall, Englewood Cliffs, N.J.

2. Devroye, L. (1986): Non-Uniform random variate generation. New York: Springer-Verlag.

3. Lewis, P.A.W. and Shedler, G.S. (1979): Simulation of nonhomogeneous Poisson processes by thinning. *Naval Research Logistice Quarterly*, 26, 403-413.

4. London, D. (1988): Survival models and thier estimation. Winsted and New Britain, Connecticut: Actex Publications.

5. Scollnik, D,P.M. (1995): Simulating random variates from Makeham's distribution and from others with exact or nearly log-concave densities. *Transactions of the Society of Actuaries*, Volume XLVII, 41-69.

Table 1: Inversion method

law	$\mu(t)$	$M(t)$	$M^{-1}(E)$
Pareto	$\frac{a}{1+x+t}$, $a > 0$	$a \ln(1+x+t)$	$\exp[\frac{E}{a}] - 1 - x$
de Moivre	$\frac{a}{\omega-x-t}$, $a > 0$	$a \ln(\omega - x) - a \ln(\omega - x - t)$	$(\omega - x)(1 - e^{-\frac{E}{a}})$
Weibull	$a(x+t)^b$, $a, b > 0$	$\frac{a(x+t)^{b+1}}{b+1} - \frac{ax^{b+1}}{b+1}$	$[\frac{E(b+1)}{a} + x^{b+1}]^{\frac{1}{b+1}} - x$
Gompertz	bc^{x+t}, $b > 0, c \geq 1$	$\frac{bc^{x+t}}{\ln(c)} - \frac{bc^x}{\ln(c)}$	$\frac{\ln(bc^x + E\ \ln(c)) - \ln(b)}{\ln(c)} - x$

Table 2: Theoretical Q-Q plots

law	X axis	Y axis
Pareto	$\ln(1 + x + s_i)$	m_i
de Moivre	$\ln[(\omega - s_i)/(\omega - x - s_i)]$	m_i
Weibull	$\ln(x + s_i)$	$\ln(m_i + \text{adjustment})$
Gompertz	s_i	$\ln(m_i + \text{adjustment})$
Makeham	s_i	$\ln(m_i + \text{adjustment})$

Figure 1: Result of the Q-Q plot for the Pareto's law ($a = 0.5$, $x = 50$) and the line $Y = aX$.

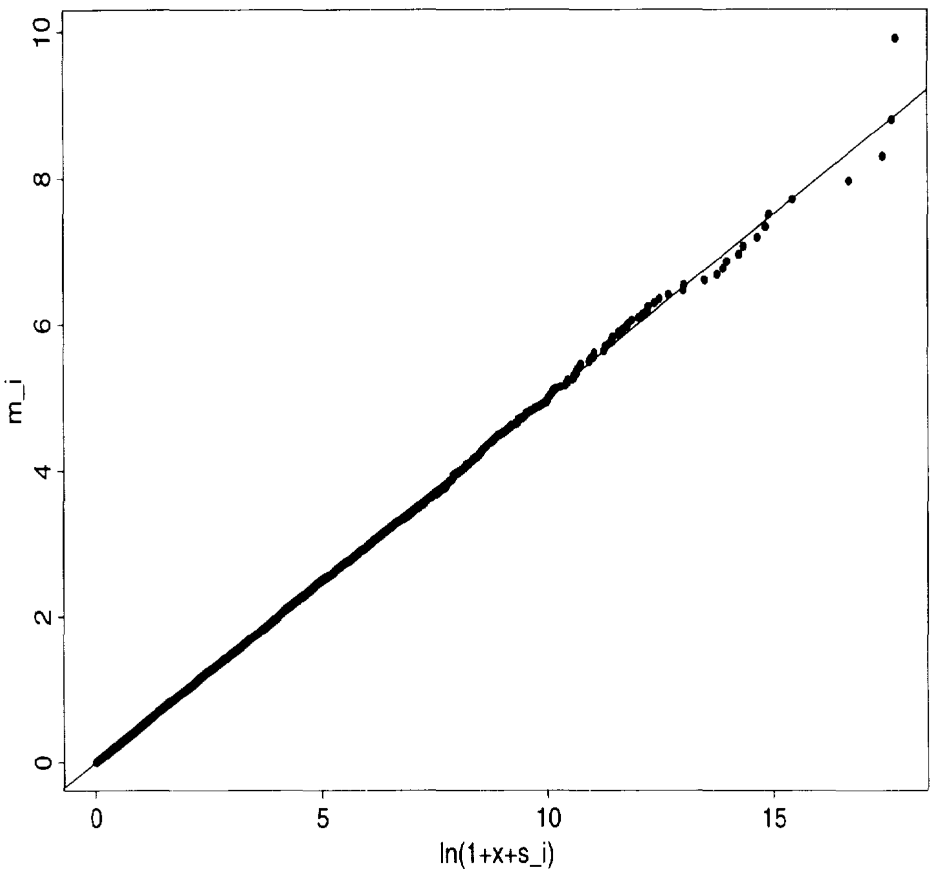

Figure 2: Result of the Q-Q plot for the de Moivre's law ($a = 0.5$, $\omega = 110$, $x = 50$) and the line $Y = a\ \ln(\omega - x) - aX$.

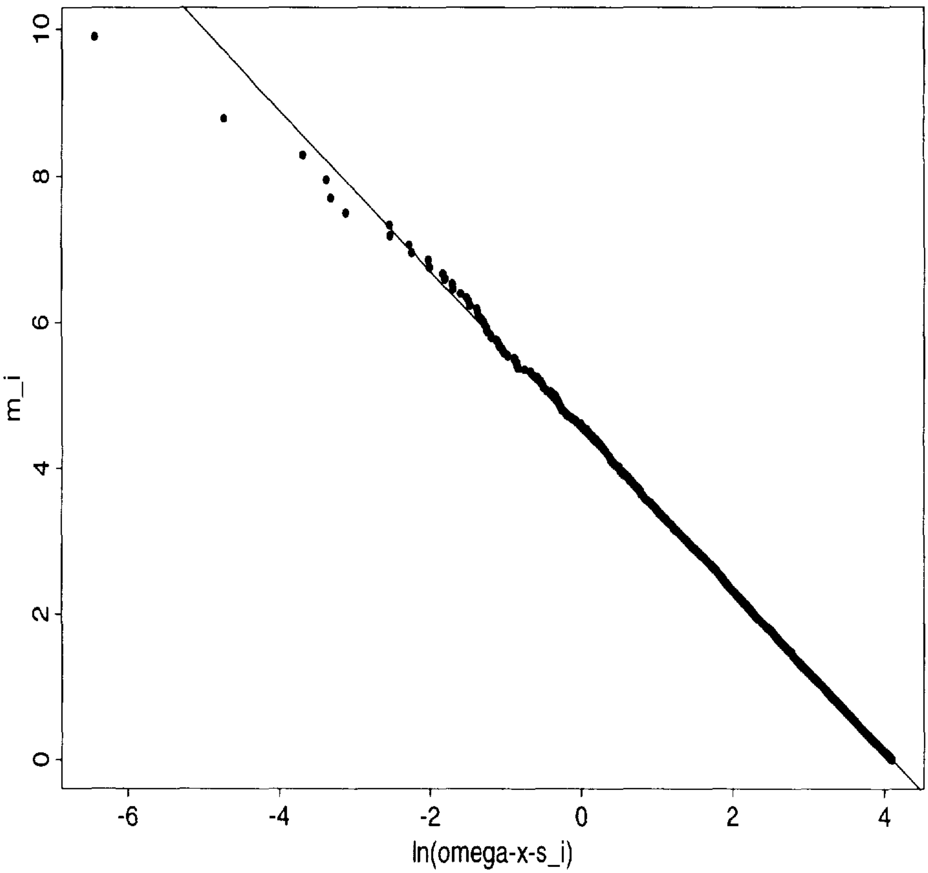

Figure 3: Result of the Q-Q plots for the Weibull's law ($a = 0.01414214$, $b = 0.5$, $x = 50$) and the line $Y = \ln(\frac{a}{b+1}) + (b+1)X$.

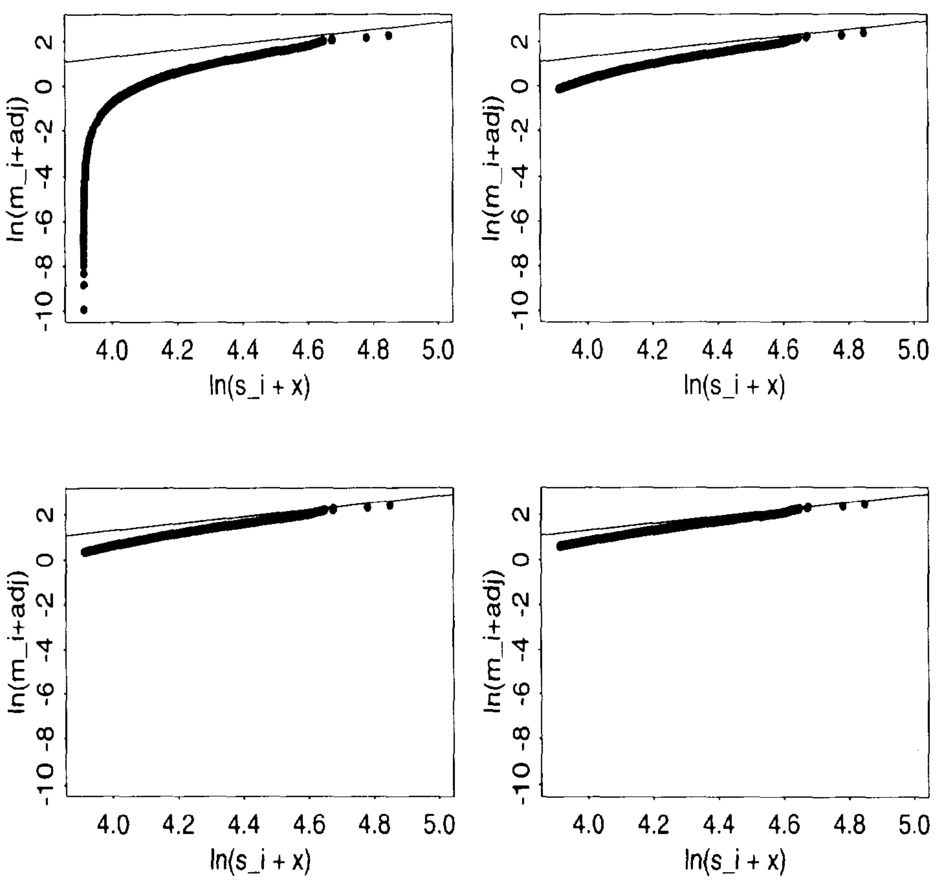

Figure 4: Result of the Q-Q plots for the Gompertz's law ($b = 0.000007$, $c = 1.12$, $x = 50$) and the line $Y = a\ \ln(bc^x/\ln(c)) - \ln(c)X$.

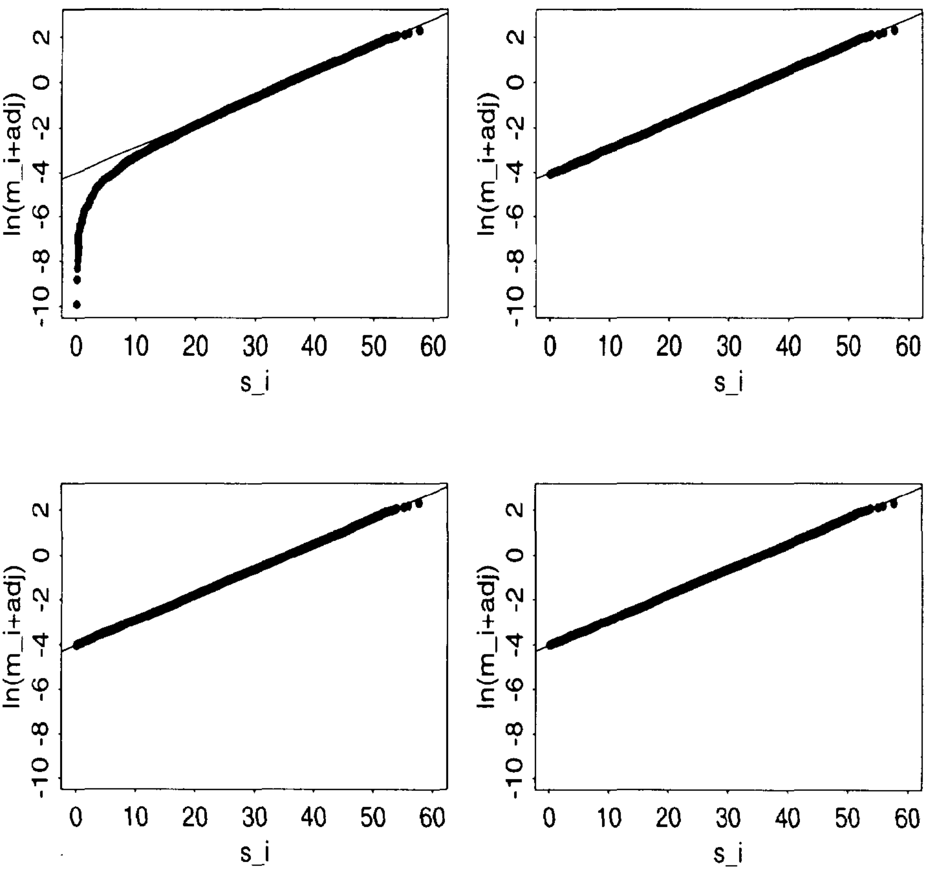

Figure 5: Result of the Q-Q plots for the Makeham's law ($a = 0.001$, $b = 0.000007$, $c = 1.12$, $x = 50$) and the line $Y = a\ \ln(bc^x/\ln(c)) - \ln(c)X$.