

# Estimating hazard rates using appropriate statistical distribution

Nermin Saad Mohamed Fahmy \*

Faculty of King Abdul-Aziz University

\*Corresponding author E-mail: [dr.nerminfahmy@yahoo.com](mailto:dr.nerminfahmy@yahoo.com)

## Abstract

This research deals primarily with the hazard rate function of a class of distributions, and discusses the relation between the hazard rate function, and the density function. It was found that the Makeham-Gompertz mortality distribution and the truncated extreme value distribution had the same hazard rate function. We used the hazard rate function derived from these distributions to find the actuarial functions.

**Keywords:** Density Function; Hazard Rate Function; Makeham Gompertz Distribution; Survival Function; Truncated Extreme Value Distribution.

## 1. Introduction

The statistical and actuarial studies indicate that the mortality rates improve continuously in the twentieth century. Also, all of these studies indicate that; there is an observable improvement in the life expectation in the developing countries.

Generally, the analyst of mortality examines changes in mortality  $\mu(x,t)$  as a function of both age  $x$  and time  $t$ . It was found in the mathematical demographic literature, and in the survival functions studies, that there is a very strong relation between the hazard rate function (H.R.F), which is a conditional density function in mortality at age  $x$  given living until the age  $x$ , mortality rates and statistical distributions. There are several mathematical relations between H.R.F and the survival function.

The survival function and the mortality rates functions were derived, consequently, the mortality probability functions, from the simple statistical distributions, (i.e.) from the Exponential distribution and the Uniform distribution. Also there are other statistical distributions more complicated and need complex derivations to these functions like Makeham-Gompertz.

This study is going to try to present how to obtain actuarial functions which, depend on statistical distributions which, could enable the company to calculate equitable premiums, and make suitable reserves which do not affect liquidity or the company profit. It means that, that the suitable premium must not be very low; because this result in losses to the company and also, must not be very high; because this will result in losses in the insurance market in the case of competition. Also, it must be considerable that, the whole policy reserve for a group of policies must not be very high or very low, if the reserve is higher than desirable, there will be access in the money considerable as reserves, consequently, prevent the company from investing this access in investments may be with high profit, whereas the reserve is less than the desirable, the company may not be able to face all the future.

## 2. Literature review

[12] Introduced Gumbel distribution which has application in the missiles, rainfall, floods, general phenomenology data, shipping the aircraft and survival period.

It is a unimodel with mode at  $x = \mu$  and has points of inflection at:

$$X = \mu \pm \sigma \ln \left\{ \frac{1}{2} (3 + \sqrt{5}) \right\} \quad (1)$$
$$= \mu \pm 0.96242 \sigma$$

The truncated distribution considered in that paper is obtained by truncating all values of  $x$  less than  $s$  and normalizing. As a result, the truncated pdf is unimodel with mode  $\mu$  if  $s < \mu$  (which  $s$  is the observed date and it is usually the date at death or the date of withdrawals) and is a strictly decreasing function if  $s \geq \mu$ .

[15] investigated the Gompertz distribution with hazard function (2) over the entire range  $(-\infty, \infty)$  and found that the standard deviation of the distribution is  $(\pi/\sqrt{6})/k$ . Since very little of the density of this distribution over  $(-\infty, \infty)$  lies over the negative part of the axis for values of  $B$  and  $k$  (where the constant  $B$  indicates the base level of mortality in life table,  $K$  shows the increase in mortality with age) conflicting with human life tables, the standard deviation of the time to death under the Gompertz law of a life aged zero is close to  $(\pi/\sqrt{6})/k$ . The mode  $m$  and ageing parameter  $k$  of the following Gompertz law

$$\mu_x = k \exp\{(x - m)k\} \quad (2)$$

Are therefore the two essential parameters summarizing recent mortality, providing measures of location and spread respectively.

They obtained the cumulant function, exact moments and percentiles for the Gompertz distribution with hazard function (3)

$$\mu_x = B e^{kx} \tag{3}$$

Over the range  $(-\infty, \infty)$ . Their moment and percentile formulae can be re-arranged as follows in terms of the mode  $m$  and ageing constant  $k$ :

$$\text{Mean} = m - \gamma / k = m - 0.5772157 / k; \tag{4}$$

$$\text{Standard deviation} = (\pi / \sqrt{6}) / k = 1.2825498 / k \tag{5}$$

$$100 \text{ percentile} = m + [\ln(-\ln(1-p))] / k; \tag{6}$$

( $\gamma$  is the Euler constant). Comparison of (4) and (6) reveals that the mean should correspond to the 43.0% point of the distribution, whilst (6) reveals that the mode must lie at the 63.2% point of the distribution.

The distribution for which equations (4)-(6) are exact, lies over the range  $(-\infty, \infty)$ . For values of the parameters  $m$  and  $k$  typical of recent human mortality, however, very little of the density function lies above the negative axis. Formulae (4)-(6) should therefore provide accurate approximations for the expectation of life at birth, the standard deviation of the time to death from birth, and the percentiles of the age at death random variable. The fact that the Gompertz distribution over  $(-\infty, \infty)$  has a very small density on the negative part of the axis will compensate to some extent for higher-than-Gompertz mortality in a normal life table below age 30.

They found that the Gompertz law still provides the simplest most accurate representation of old age mortality, and its parameters, the mode and force of mortality at the mode, provide respectively the measures of location and spread. They found also, that the Gompertz law of mortality can be a very useful model for obtaining quick accurate approximations for many life table functions.

[6] discussed the problem of estimating the parameters of the left and the right truncated Gompertz distribution when truncation points are unknown. He stated that truncated distributions arise when sample selection and observation is not possible in some subregion of the sample space. This occurs as a consequence of actual elimination of part of the original data.

The Gompertz distribution is applicable as a model for survival distributions which has an increasing hazard rate for the life of the creatures and systems, and it is widely used in actuarial works.

[14] showed that the mortality of a recent developed population is largely the mortality of old age, and most of the information about it may be summarized by two parameters, the first indicating the age near which most of the deaths occur and the second the spread of deaths about that age. The Gompertz law of mortality, still provides a good general representation of the age pattern of mortality at these ages, and its parameters, a model age at death and aging parameter (equal to the force of mortality at that model age) provide direct measures of location and spread. He calculated the two Gompertz parameters if the upper and lower quartiles of the curve of deaths are known, and accurate approximations to numerous life table functions (e.g.  $e_x$ ) are then immediately available.

[16] developed methods that allow the calculation of generational life annuities and life expectancies from the initial period table, obviating the need for sets of generational tables and revision of these tables.

[5] discussed the problem of estimating the parameters of the truncated extreme value distribution. The method of Maximum Likelihood Estimating is used to estimate the parameter. Necessary and sufficient conditions for the existence and uniqueness of the roots of the likelihood equations of the model and censoring data are given.

[17] used methods dependent on four important observations concerning a life table of the Gompertz form.

- a) The force of mortality at the mode of the curve of deaths is equal to  $k$ , (where  $k$  shows the increase in mortality with age).
- b) The effect on life table of multiplying the force of mortality at every point throughout the life span by the same factor  $r^x$  (where  $r$ , the annual mortality improvement factor, is typically in the neighborhood of 1.0) to produce a new Gompertz life table in which a life aged  $x$  has mortality equivalent to a life in the original table aged  $x + N \ln(r) / k$ .
- c) If a population has a cross sectional (period) life table of the Gompertz form with ageing parameter  $k$ , and all its members are subject to a constant on-going annual improvement factor  $r$  at every age, then a member of the population aged  $x$  with force of mortality  $\mu_x$  under the period table has a Gompertz generational life table with ageing parameter  $k + \ln(r)$ .
- d) If the force of the mortality at exact age  $x$  is fixed at  $\mu_x$  and the ageing parameter  $k$  is allowed to change, then the change in the complete expectation of life at age  $x$  corresponding to the change in  $k$  is given by the derivative. Finally, he found that the derivatives essential to calculate generational life annuities are easy to compute.

[1] investigated the customer credit scoring aimed at distinguished good payers from bad payers at the time of the loan application. Their aim of survival analysis is to estimate the distribution of the event times  $f(t)$  of a group of subjects. This usually is done by specifying two other mathematically equivalent functions, the survival function  $S(t)$  and the hazard function  $h(t)$  defined as follows [1], [10]:

$$S(t) = P(T > t) = \int_t^\infty f(u) du = 1 - F(t)$$

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \tag{7}$$

Whereby  $F(t)$  represents the cumulative distribution function. The survival function models are the probability that a subject will survive at least time period  $t$ .

This paper deals primarily with a class of distributions with hazard rate  $h(x) = a(x) b(\theta)$  and the parameters  $(\theta_1 \dots \theta_\ell)$  where  $\theta_j > 0, j = 1 \dots \ell$  of the class of distributions with hazard rate

$$h(x) = \sum_{j=1}^{\ell} h_j(x) \tag{8}$$

Where  $h_j(x) = a_j(x) b_j(\theta)$  is the cause of death specific hazard rate associated with the  $j^{\text{th}}$  cause. The functions  $b(\theta)$  and  $b_j(\theta)$  are positive, differentiable and strictly monotonic functions of  $\theta$ . This class of distributions includes the following distributions as well as their corresponding compound distributions, i.e., distributions with several hazard rates corresponding to independent causes, see Dick London (1988).

- The relation between the hazard rate function  $\mu_x$ , the survival function  $S(x)$  and the density function  $f(x)$ :

$$\mu_x = \frac{f(x)}{S(x)} \tag{9}$$

$$\mu_x = \frac{-d}{dx} \ln S(x) = \frac{-d}{dx} S(x) \tag{10}$$

Where

$$S(x) = \exp \left\{ -\int_0^x \mu_y dy \right\} \tag{11}$$

$$= 1 - F(x)$$

$$f(x) = \frac{d}{d(x)} F(x) = \frac{d}{d(x)} (1 - S(x)) = \frac{-d}{d(x)} S(x) \quad (12)$$

### 3. Gompertz distribution

Consider the law of human mortality described by [7] where the force of mortality has the form  $\mu_x = B C^x$  with  $B > 0$ ,  $C \geq 1$ . It is well known that this model is an excellent description of the pattern of mortality at the adult ages. Nevertheless, actuaries rarely use it to calculate the value of annuities and insurances, as evidenced by the textbooks written by [9], [2]. The probable reason for this is the intractability of the mathematical expectations that emerge from the analysis.

The purpose of this section is to elevate the Gompertz law and compile a set of facts that are accessible to the practicing actuary. It was found that many of the actuarial functions can be expressed in terms of the left-truncated gamma function, as [13] demonstrated. Hoping that future actuaries will adopt the model for calculating the net single premiums of life insurances and annuities.

First, a few definitions and background facts about the Gompertz law are introduced. A parameterization of the law that is statistically informative are present. Next, this law readily explains the pattern of mortality for a valuation mortality table are showed and the Gompertz parameters with the valuation mortality rates are estimated. Next, explicit expressions are derived for continuous life insurances, annuities, net level premiums and reserves by using the left-truncated gamma function which can easily be approximated.

### 4. Definitions and basic results

This section presents some notation and basic results. It is started by giving an informative representation of the Gompertz law. [3] expresses the force of mortality as:

$$\mu_x = \frac{1}{\sigma} \exp\{(x - m)/\sigma\}, \sigma > 0, m \in \mathfrak{R} \quad (13)$$

Note that  $c = e^{1/\sigma}$  and  $B = \sigma^{-1} e^{-m/\sigma}$ . This representation is informative because  $m$  is a location parameter approximately equal to the mean and  $\sigma$  is a scale parameter proportional to the standard deviation. In human populations we usually find that  $m > \sigma > 0$ . For the United States population, Carriere (1992) found that these parameters were  $m = 82.3$  and  $\sigma = 11.4$ .

Using the relation  $S(x) = \exp\left\{-\int_0^x \mu_t dt\right\}$ , we find that the survival function for the Gompertz law is equal to

$$S(x) \equiv \exp\left\{e^{-m/\sigma} - e^{(x-m)/\sigma}\right\} \quad (14)$$

The new parameter  $\sigma$  can be interpreted as a dispersion parameter because if  $m > 0$ , then

$$\lim_{\sigma \rightarrow 0} \{S(m - \varepsilon) - S(m + \varepsilon)\} = 1, \forall \varepsilon > 0 \quad (15)$$

This limit suggests that all the mass concentrates about  $m$  when  $\sigma$  is small and so  $m$  can also be interpreted as a location parameter when  $m > 0$ . Using the relation  $f(x) = -\frac{d}{dx} S(x) = \mu_x S(x)$  we find that the density is equal to

$$f(x) = \frac{1}{\sigma} \exp\left\{e^{-m/\sigma} - e^{(x-m)/\sigma} + (x - m)/\sigma\right\}, x \geq 0 \quad (16)$$

It is easy to verify that the mode of the density is equal to when  $m \leq 0$  and that the mode is  $m$  when  $m > 0$ .

[15] defined a Gompertz law for all  $x \in \mathfrak{R}$  and found explicit formulas for the moments. Using our notation, they analyzed the distribution function

$$G(x) \equiv 1 - \exp\left\{-e^{(x-m)/\sigma}\right\}, x \in \mathfrak{R} \quad (17)$$

[15] fail to mention that this extended Gompertz distribution is simply the Gumbel distribution for minima, an extreme value distribution. For an extensive discussion about extreme-value distribution, consult [8]. For a specific discussion about the Gumbel distribution, see [10].

### 5. Estimation of the gompertz parameters

In this section, we estimate the parameters of the Gompertz law using the AM92 and EG96 ultimate mortality rates. The Gompertz law is only applicable at the older ages and so we focus our analysis on the rates  $\hat{q}_x$ , for the ages  $x = 40, \dots, 100$ .

#### 5.1. Estimation using the crude rates

Using the relation  $q_x = 1 - \frac{S(x+1)}{S(x)}$ , we find that Gompertz law yields the identity

$$q_x(m, \sigma) \equiv 1 - \exp\left\{e^{(x-m)/\sigma} (1 - e^{-1/\sigma})\right\}. \quad (18)$$

Let  $D_x$  denote the total amount of death claims associated with the crude rate  $\hat{q}_x$ . [4] suggests that a good way of estimating  $m$  and  $\sigma$  is to minimize the robust loss function

$$L(m, \sigma) = \sum_{x=40}^{100} \sqrt{D_x} \left| 1 - \frac{q_x(m, \sigma)}{\hat{q}_x} \right|. \quad (19)$$

The function  $L(m, \sigma)$  was minimized using the nonlinear module of the statistical computer software Mathcad 2000. We found that the nonlinear simplex or polytope algorithm was very successful at minimizing the non-differentiable loss function  $L(m, \sigma)$  using the AM92 data we found that the parameters  $\hat{m}_x = 82.153$  and  $\hat{\sigma} = 10.304$ , minimized (19) the EG96 data yielded the parameters  $\hat{m}_x = 87.281$  and  $\hat{\sigma} = 10.478$

#### 5.2 Estimation using the graduated rates

Using the relation  ${}_x q_0 = S(x) - S(x+1)$ , we find that the Gompertz law yields the identity

$${}_x q_0 \equiv \exp\left\{e^{-m/\sigma}\right\} \cdot \left\{\exp\{-\sigma \mu_x\} - \exp\{-\sigma \mu_{x+1}\}\right\} \quad (20)$$

Let  ${}_x \hat{q}_0$  denote a graduated probability from the Basic Tables. Another way of estimating  $m$  and  $\sigma$  is to minimize the function

$$L(m, \sigma) = \sum_{x=40}^{99} ({}_x q_0(m, \sigma) - {}_x \hat{q}_0)^2 \quad (21)$$

Using the AM92 data we found that the parameters  $\hat{m} = 81.022$  and  $\hat{\sigma} = 10.379$  minimized the previous formula. The EG96 data yielded the parameter estimates  $\hat{m} = 86.235$  and  $\hat{\sigma} = 9.593$ . Note that these parameter estimates are similar to those given in part A.

### 6. Continuous insurance and annuities

In this section, we give explicit expressions for continuous life insurances, annuities, net level premiums and reserves by using the left-truncated gamma function. Some of these expressions may be found in London (1988). Let  $T(x) \geq 0$  denote the time at death of a life aged  $x \geq 0$  and let  ${}_t p_x \equiv \frac{S(x+1)}{S(x)}$  denote the probability of surviving to time  $t \geq 0$ . For the Gompertz law, we can write

$${}_t p_x = \exp \{ e^{(x-m)\sigma} (1 - e^{t/\sigma}) \} = \exp \{ \sigma \mu_x (1 - e^{t/\sigma}) \} \tag{22}$$

Let  $M_x(u) \equiv E[\exp\{uT(x)\}]$  denote the moment generating function of  $T(x)$ . the density of  $T(x)$  is  ${}_t p_x \mu_{x+t}$  and so  $M_x(u)$  is equal to

$$M_x(u) = \int_0^\infty e^{ut} {}_t p_x \mu_{x+t} dt = \int_0^\infty e^{ut} \exp \{ \sigma \mu_x (1 - e^{t/\sigma}) \} \mu_x e^{t/\sigma} dt. \tag{23}$$

Applying the transformation  $z = \sigma \mu_x e^{t/\sigma}$  yields

$$M_x(u) = \exp \{ \sigma \mu_x - u(x-m) \} \int_{\sigma \mu_x}^\infty z^{-u/\sigma} e^{-z} dz \tag{24}$$

It is fairly easy to show that  $M_x(u) < \infty$  fall  $u \in \mathfrak{R}$ . Let us denote the integral in (12) as  $\Gamma(\sigma \mu_x, 1+u\sigma)$  and introduce the left-truncated gamma function. This is defined as:

$$\Gamma(t, \alpha) \equiv \int_t^\infty z^{\alpha-1} e^{-z} dz, t > 0, \alpha \in \mathfrak{R}. \tag{25}$$

Note that  $\Gamma(t, \alpha) < \infty$ . Next, the cumulant generating function, denoted as  $\psi_x(u)$ , is defined as

$$\psi_x(u) \equiv \ln \{ M_x(u) \} = um - ux + \sigma \mu_x + \ln \{ \Gamma(\sigma \mu_x, 1+u\sigma) \}. \tag{26}$$

Let  $\delta \geq 0$  be the force of interest and let  $\bar{A}_x$  denote the net single premium for a continuous whole-life insurance. Let  $u = -\delta$  in (12), then we get

$$\bar{A}_x = \exp \{ \sigma \mu_x + \delta(x - m) \} \cdot \Gamma(\sigma \mu_x, 1 - \delta\sigma). \tag{27}$$

Consider the continuous life annuity, expressed as a Stieltjes integral,

$$\bar{a}_x \equiv \int_0^\infty \int_0^t e^{-\delta x} dx d({}_t q_x) = \int_0^\infty e^{-\delta t} {}_t p_x dt \tag{28}$$

The last identity for  $\bar{a}_x$  is true regardless of any continuity assumptions on  ${}_t p_x$ . Calculating further, we find that

$$\bar{a}_x = \sigma \exp \{ \sigma \mu_x + \delta(x - m) \} \cdot \Gamma(\sigma \mu_x, -\delta\sigma). \tag{29}$$

Once again, we get an expression that is based on the left-truncated gamma function. Next, consider the level premium  $\bar{P}_x \equiv \bar{A}_x / \bar{a}_x$ . Using (29) and (27) we find that

$$\bar{P}_x = \Gamma(\sigma \mu_x, 1 - \delta\sigma) / \sigma \Gamma(\sigma \mu_x, -\delta\sigma). \tag{30}$$

Also, consider the net level premium reserve  ${}_V \bar{V}(\bar{A}_x) = 1 - \frac{\bar{a}_{x+t}}{\bar{a}_x}$ .

Using (29) we find that

$${}_V \bar{V}(\bar{A}_x) = 1 - \exp \{ \sigma(\mu_{x+t} - \mu_x) - \delta t \} \frac{\Gamma(\sigma \mu_{x+t}, -\delta\sigma)}{\Gamma(\sigma \mu_x, -\delta\sigma)}. \tag{31}$$

Let us prove that  $\bar{A}_x \rightarrow 1$  as  $x \rightarrow \infty$ . This will imply that  $\bar{P}_x \rightarrow 0$  and  ${}_V \bar{V}(\bar{A}_x) \rightarrow 1$  as  $x \rightarrow \infty$ . It is sufficient to prove that  $\bar{a}_x \rightarrow 0$  as  $x \rightarrow \infty$  because  $\bar{A}_x = 1 - \delta \bar{a}_x$ . Note that

$$0 \leq \bar{a}_x = \int_0^\infty e^{-\delta t} {}_t p_x dt \leq \int_0^\infty {}_t p_x dt$$

And that  ${}_t p_x = \exp \{ e^{(x-m)\sigma} (1 - e^{t/\sigma}) \} \rightarrow 0$  as  $x \rightarrow \infty$  for all  $t > 0$ . Therefore, by the monotone Convergence Theorem (Royden, 1986),  $\int_0^\infty {}_t p_x dt \rightarrow 0$  as  $x \rightarrow \infty$  and the result follows.

### 7. Makeham-gompertz distribution

The Makeham hazard rate (force of mortality) as presented in [9] is a generalization of the Gompertz hazard rate and is written as

$$h(x) = \alpha + Bc^x = \alpha + \frac{1}{\sigma} \exp\left(\frac{x - \mu}{\sigma}\right), \alpha, \sigma > 0, -\infty < \mu < \infty \tag{32}$$

It has wide usage in actuarial applications, the constant  $\alpha$ , adds a hazard that acts with equal intensity at all ages under consideration. [5], [11] studied a variety of methods for estimating the Makeham parameters including the method of Maximum Likelihood.

In studies of insured lives it may be possible to identify a single cause of death associated with the parameter  $\alpha$ . One such cause is accidental means. For lives insured with accidental death benefits only or in conjunction with regular insurance, accidental death is established for each death claim.

When proposing an analytic function that would closely reproduce the typical  $\ell_x$  curve, [7] stated, "It is possible that death may be the consequence of two generally coexisting causes: the one, chance, without previous disposition to death or deterioration; the other, a deterioration, or increased inability to withstand destruction." However, his law of mortality only takes account of the second cause. [7] combined the two causes additively and derived the Makeham-Gompertz of mortality,

$$\mu_x = A + B C^x \tag{33}$$

Where  $\mu_x$  is the force of mortality (a measure of the mortality at the precise moment of attaining age  $x$ ).

$$\mu_x = -(D\ell_x) / \ell_x \tag{34}$$

Where  $D\ell_x$  denotes the derivative of  $\ell_x$  with respect to  $x$ , so an expression for  $\ell_x$  may be derived from the expression for  $\mu_x$

### 8. Extreme value distribution

Extreme value distributions are generally considered to consist of the three following families:

1) Type 1:

$$\text{pr}[X \leq x] = \exp \{ - e^{-(x-\epsilon)/\theta} \} \tag{35}$$

2) Type 2:

$$\text{pr}[X \leq x] = \begin{cases} 0 & x < \varepsilon \\ \exp\{-\left(\frac{x-\varepsilon}{\theta}\right)^k\} & x \geq \varepsilon \end{cases} \quad (36)$$

3) Type 3:

$$\text{pr}[X \leq x] = \begin{cases} \exp\{-\left(\frac{\varepsilon-x}{\theta}\right)^k\} & x \leq \varepsilon \\ 1 & x > \varepsilon \end{cases} \quad (37)$$

Where  $(\varepsilon, \theta > 0)$  and  $(k > 0)$  are parameters.

The corresponding distributions of  $(-x)$  are also called extreme value distributions.

Of these three families of distributions, Type 1 is by far the one most commonly referred to in discussions of ‘extreme value’ distributions. Indeed, some authors call (1) “the” extreme value distribution. In view of this, and the fact that distributions (2) and (3) can be transformed to type 1 distributions by the simple transformations

$$Z = \log(x - \varepsilon), Z = -\log(\varepsilon - x)$$

Respectively, we will confine ourselves to discussion of Type 1 distributions. We may also note that the Type 3 distribution of  $(-x)$  is a Weibull distribution. Of course, Types 1 and 2 are also closely related to the Weibull distribution Type 1 is sometimes called the Log-Weibull distribution.

The (p.d.f) of this distribution is given by:

$$f(x) = \frac{1}{\sigma} \exp\left\{\frac{x-\mu}{\sigma} - \exp\left(\frac{x-\mu}{\sigma}\right)\right\} \quad (38)$$

$$-\infty < x < \infty, \sigma > 0, -\infty < \mu < \infty$$

It is called The Second Double Exponential Distribution, and the type 1 extreme value distribution of the smallest value by [8]. Thus we could refer the distribution whose pdf is given above as the extreme value distribution. It is unimodal with mode at  $x = \mu$  and has points of inflection at

$$x = \mu \pm \sigma \ln\left(\frac{1}{2}(3 + \sqrt{5})\right) \quad (39)$$

$$= \mu \pm 0.9624 \sigma$$

The p.d.f of  $y = \frac{x-\mu}{\sigma}$  is shown in the following figure F(y)

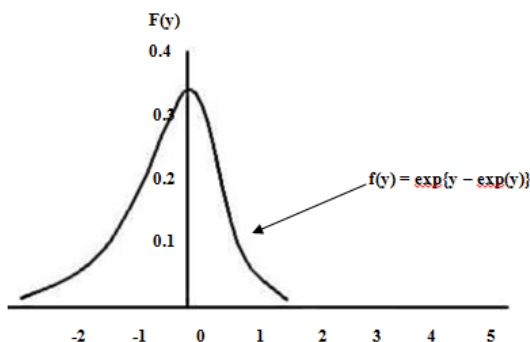


Fig. 1: Probability Density Function.

The truncated distribution considered here is obtained by truncating all values of  $x$  less than  $s$  and normalizing.

As a result, the truncated p.d.f is unimodal with mode  $\mu$  if  $s < \mu$  and is a strictly decreasing function if  $s > \mu$ . The inflection points occur at the same points as truncated distribution if such points

have been truncated. The (pdf) of the truncated random variable  $x$  is:

$$f(x) = \frac{1}{\sigma} \exp\left\{\frac{x-\mu}{\sigma} + \exp\left(\frac{S-\mu}{\sigma}\right) - \exp\left(\frac{x-\mu}{\sigma}\right)\right\} \quad (40)$$

$$X > S, \sigma > 0, -\infty < \mu < \infty$$

And the cumulant distribution function (c.d.f) of the truncated random variable  $x$  is

$$F(x) = 1 - \exp\left\{\exp\left(\frac{S-\mu}{\sigma}\right) - \exp\left(\frac{x-\mu}{\sigma}\right)\right\} \quad X > S$$

$$= 0 \quad X \leq S$$

The hazard rate function (or the force of mortality) is

$$h(x) = \frac{f(x)}{1 - F(x)} = \frac{1}{\sigma} \exp\left(\frac{x-\mu}{\sigma}\right) \quad X > S$$

$$= 0 \quad X \leq S$$

[6] Discussed the problem of estimating the parameters of the left and the right Truncated Gompertz distribution when truncation points are unknown. Maximum likelihood estimators and estimators involving expected values of appropriate order statistics are derived. Asymptotic sampling errors of estimates are also given. [5] Proved that the MLE for the truncated extreme value distribution exists and unique.

### Conclusions

Finally we conclude in this research that Gompertz distribution or modified Gompertz distribution or truncated extreme value distribution has the same hazard rate,  $\mu_x = A + B C^x$  (33).

### References

- [1] Beasens, B., Gestel, V., Stepanova, M., Sunkens, J., and Vanthienen, J., (2004), "Neural Network Survival Analysis for Personal Loan Data", Research Society, No. 8.
- [2] Bowers, N.L., Gerber, H.U., Hickman, J.C., D.A. and Nesbitt, C. J. (1986), "Actuarial Mathematics". Societies of Actuaries.
- [3] Carriere, J.F. (1992), "Parametric Models for Life Tables". Of the society of Actuaries, Vol. XLIV.
- [4] Carriere, J.F. (1994), "A Select and Ultimate Parametric Model". Transactions of the society of Actuaries, Vol. XLIV.
- [5] El-Bolkiny, M., (1999), "The Existence and Uniqueness of the MLE of The Truncated Extreme Value Distribution", 14th International Conference for Statistics, Computer Science, Social and Demographic Research, Cairo, Ain Shams, University, Vol.1.
- [6] El-Youssef, M.H., (1993), "Estimation of Parameters and Truncation Point for the Truncated Gompertz Distribution", Journal of King Saudi University, Vol.5.
- [7] Gompertz, B. (1825), "On the Nature of the Function Expressive of the Law of Human Mortality". Philosophical Transactions of the Royal Society, 115.
- [8] Johnson, N.L. and Kotz, S. (1970), "Continuous Univariate Distributions", Vol. 1. Wiley.
- [9] Jordan, C.W. (1975), "Society of Actuaries Textbook on Life Contingencies", Second Edition, Society of Actuaries, Chicago.
- [10] Kotz, S. and Johnson, N.L. (1983), "Encyclopedia of Statistical Sciences", Vol. 3. Wiley.
- [11] London, D., (1988), "Survival Models and Their Estimation", Winsted and New Britain, Connecticut.
- [12] Menn, H., and Nandi, K. (1974), "On Some Properties of Roy's Union Intersection Tests", Calcutta Statistical Association Bulltin, Vol.14.
- [13] Mereu, J.A. (1962), "Annuity Values Directly from Makeham Constants". Transactions of the Society of Actuaries, Vol. XXXII.
- [14] Parr, N., (1998), "Proportional Hazards Analysis Behavior", Genus, Vol.47.

- [15] Pollard, J.H. and Valkovics, E.J. (1993) "The Gompertz Distribution and its Applications Genus".
- [16] Pollard, J.H. (1998), "An Old Tool Moder Applications", Actuarial Studies and Research Papers Series Research No. 001/98, Macquarie University.
- [17] Pollard, J.H. (2000), "Keeping A Breast of Mortality Change", Genus, No.1

$$f(x) = \sum_{j=1}^{\ell} \mu_{jx} \exp \left\{ -\sum_{j=1}^{\ell} \int_s^x \mu_{jt} dt \right\}$$

Which includes combinations of the above distributions. As an example, the p.d.f

$$f(x) = (\theta_1 x + \theta_2 / x) \left( \frac{s}{x} \right)^{\theta_2} \exp \{ -\theta_1 (x^2 - s^2) / 2 \}, x > s > 0,$$

## Appendix

Relation between the hazard rate function  $\mu_x$ , the survival function  $S(x)$  and the density function  $f(x)$ .

$$\mu_x = \frac{f(x)}{S(x)}$$

$$\mu_x = \frac{-\frac{d}{dx} S(x)}{S(x)} = -\frac{d}{dx} \ln S(x)$$

Where

$$S(x) = \exp \left\{ -\int_0^x \mu_y dy \right\}$$

$$= 1 - F(x)$$

$$f(x) = \frac{d}{d(x)} F(x) = \frac{d}{d(x)} (1 - S(x)) = -\frac{d}{d(x)} S(x)$$

Some easy hazard rate functions:

- 1) Exponential:

$$f(x) = \theta \exp \{ -\theta (x-s) \}, x > s$$

$$\mu_x = \theta$$

- 2) Rayleigh:

$$f(x) = \theta x \exp \{ -\theta (x^2 - s^2) / 2 \}, x > s > 0,$$

$$\mu_x = \frac{\theta}{x}$$

- 3) Extreme Value (when  $\sigma$  is known):

$$f(x) = \frac{1}{\sigma} \exp \left\{ \frac{x-\theta}{\sigma} - \exp \left( \frac{x-\theta}{\sigma} \right) + \exp \left( \frac{s-\theta}{\sigma} \right) \right\}, x > s$$

$$\mu_x = \frac{1}{\sigma} \exp \left( \frac{x-\theta}{\sigma} \right),$$

- 4) Pareto

$$f(x) = \theta s^\theta / x^{\theta+1}, x > s > 0,$$

$$\mu_x = \theta / x.$$

Any pdf with hazard rate  $\mu_x$  is of the form

$$f(x) = \mu_x \exp \left\{ -\int_s^x \mu_t dt \right\}, x > s.$$

The corresponding compound pdf is of the form

Is a combination of the Pareto and the Rayleigh densities where  $\mu_{1x} = \theta_1 x$  and  $\mu_{2x} = \theta_2 / x$  are the cause specific hazard rates.

The distributions described above are not generally used in actuarial mortality studies. However, the Exponential, Rayleigh and Extreme-Value Distributions have been used to describe failure-time distributions in life testing, the study of inanimate system life length. The Pareto Distribution has been used to describe the distribution of incomes over a certain level, an application where the data is often coarsely grouped.