Sociedad de Estadística e Investigación Operativa
Test (2003) *Vol.* **12**, *No.* 2, pp. 413–426

Bayesian Prediction Under a Mixture of Two-Component Gompertz Lifetime Model

Zeinhum F. Jaheen*
Department of Mathematics,
University of Assiut, Egypt.

Abstract

A heterogeneous population can be represented by a finite mixture of two component Gompertz lifetime model. Based on type I censored samples, prediction bounds for the s^{th} future observable from a heterogeneous population are obtained from a Bayesian approach. A numerical example is given to illustrate the procedures and the accuracy of the prediction intervals is investigated via extensive Monte Carlo simulation.

Key Words: Heterogeneous population, type I censored sample, mixture of two Gompertz, Bayesian prediction, Monte Carlo simulation.

AMS subject classification: 62F15, 62M20

1 Introduction

In many practical problems of Statistics, one wishes to use the results of a previous data to predict the results of a future data from the same population. One way to do this is to construct an interval, which will contain these results with a specified probability. This interval is called the prediction interval. Prediction has been applied in medicine, engineering, business and other areas as well. For details on the history of statistical prediction, analysis and applications, see for example, Aitchison and Dunsmore (1975) and Geisser (1993).

The Gompertz model was formulated by Gompertz (1825) to fit mortality tables. It has been used as a growth model, especially in epidemiological

*Correspondence to: Zeinhum F. Jaheen, Department of Mathematics, Umm Al-Qura University (TAIF), P.O. Box: 3970, Taif Kingdom of Saudi Arabia. Email: jaheen_64@yahoo.com

Received: January 2001; Accepted: October 2002

and biomedical studies. The probability density function (pdf) of the Gompertz distribution (denoted by Gomp(λ)) is given by

$$f(t \mid \lambda) = \exp\left\{\lambda t - \frac{1}{\lambda}[\exp(\lambda t) - 1]\right\}, \quad t > 0, \quad (\lambda > 0). \tag{1.1}$$

The pdf of the Gompertz distribution is unimodal. It has positive skewness and an increasing hazard rate function. It can be shown that the Gompertz distribution is a truncated form of the type I extreme value distributions, (Johnson et al., 1995). Osman (1987) derived a compound Gompertz model by assuming that one of the parameters of the Gompertz distribution is a random variable following the gamma distribution. He studied the properties of compound Gompertz distribution and suggested its use for modelling lifetime data and analyzing the survivals in heterogeneous populations.

Mixtures of life distributions occur when two different causes of failure are present each with the same parametric form of life distributions. Finite mixtures of distributions have been used as models throughout the history of modern statistics. For more details about finite mixture of distributions see, Everitt and Hand (1981), Titterington et al. (1985), McLachlan and Basford (1988), Lindsay (1995), and McLachlan and Peel (2000), among others.

A random variable T is said to follow a finite mixture distribution with k components, if the density function of T can be written in the form:

$$f(t) = \sum_{j=1}^{k} p_j f_j(t), \tag{1.2}$$

where p_j is a non-negative real number (known as the j^{th} mixing proportion) such that $\sum_{j=1}^{k} p_j = 1$ and f_j is a density function (known as the j^{th} component), $j = 1, 2, ..., k$.

The property of identifiability is an important consideration on estimating the parameters in a mixture of distributions, testing hypotheses, classification of random variables, etc., can be meaningfully discussed only if the class of all finite mixtures is identifiable. Discussions of the identifiability of finite mixtures may be found in several papers, among others, by Teicher (1961, 1963), AL-Hussaini and Ahmad (1981) and Ahmad (1988).

A finite mixture of two-component Gompertz lifetime model may describe a heterogeneous population. The pdf of a finite mixture of two Gomp(λ_j), $j = 1, 2$, components is given by

$$f_T(t \mid p, \lambda_1, \lambda_2) = pf_1(t \mid \lambda_1) + (1-p)f_2(t \mid \lambda_2), \tag{1.3}$$

where, $p \in [0, 1]$ and for $j = 1, 2$, $f_j(t)$ are given by (1.1) after indexing λ by j.

Gordon (1990a,b) assumed that the survival function of treated cancer patients is a mixture of two subpopulations, one of which dies of other causes with a given proportion and the members of the other subpopulation die of their disease with the complementary proportion. The survival time distributions for both subpopulations are modelled by the Gompertz distribution. Wu and Lee (1999) characterized the mixtures of Gompertz distributions by using conditional expectation of order statistics. AL-Hussaini et al. (2000) showed that a finite mixture of k Gompertz components is identifiable. They considered estimation of the reliability and hazard rate functions of a mixture of two Gompertz components based on type I and II censored samples, using maximum likelihood and Bayesian methods.

In this paper, Bayesian prediction bounds for the s-th future observable from a heterogeneous population represented by a finite mixture of two component Gompertz lifetime model are obtained based on type I censored samples. A numerical example is given to illustrate the computations in the two particular cases of predicting the minimum and maximum of future observables, and the accuracy of prediction intervals is investigated via extensive Monte Carlo simulation

2 Bayesian two-sample prediction

Suppose that n units from a population with pdf (1.3) are subjected to a life testing experiment and that the test is terminated after a predetermined time t_0 (type I censoring). It is assumed that an item can be attributed to the appropriate subpopulation after it had failed. Suppose that r units have failed during the interval $(0, t_0)$: r_1 from the first and r_2 from the second subpopulation, such that $r = r_1 + r_2$ and $n - r$ units which cannot be identified as to subpopulation are still functioning. Naturally, neither

r_1 nor r_2 (and consequently r) is fixed, but they are rather determined by knowing t_0. Let t_{ij} denote the failure time of the j^{th} unit that belongs to the i^{th} subpopulation and that $t_{ij} \leq t_0, j = 1, 2, ..., r_i, i = 1, 2$. Such scheme of sampling was suggested by Mendenhall and Hader (1958).

Based on such scheme of sampling, the likelihood function (see Mendenhall and Hader, 1958) is given by

$$L(p, \lambda_1, \lambda_2 \mid \underline{t}) \\ = \frac{n!}{(n-r)!}\left[\prod_{j=1}^{r_1} p f_1(t_{1j} \mid \lambda_1)\right]\left[\prod_{j=1}^{r_2}(1-p) f_2(t_{2j} \mid \lambda_2)\right][R(t_0)]^{n-r}, \quad (2.1)$$

where $R(t)$ is the reliability function of the mixture model given by

$$R(t) = pR_1(t) + (1-p)R_2(t), \quad (2.2)$$

$R_j(t)$ is the reliability function corresponding to component j, $j = 1, 2$.

It then follows, by substituting in (2.1) and expanding $[R(t_0)]^{n-r}$ using the binomial expansion, that

$$\begin{aligned} L(p, \lambda_1, \lambda_2 \mid \underline{t}) &= \frac{n!}{(n-r)!}\sum_{j_1=0}^{n-r}\binom{n-r}{j_1} p^{r_1+n-r-j_1}(1-p)^{r_2+j_1} \\ &\quad \times \exp\left[\sum_{i=1}^{2}\{\lambda_i T_i - B_{ij_1}(\lambda_i)\}\right], \end{aligned} \quad (2.3)$$

where $\underline{t} = (t_{11}, ..., t_{1r_1}, t_{21}, ..., t_{2r_2})$, and for $i = 1, 2$,

$$\begin{aligned} T_i &= \sum_{j=1}^{r_i} t_{ij}, \\ B_{1j_1}(\lambda_1) &= \frac{1}{\lambda_1}\{(n-r-j_1)\phi(\lambda_1; t_0) + \omega_1\}, \\ B_{2j_1}(\lambda_2) &= \frac{1}{\lambda_2}\{j_1\phi(\lambda_2; t_0) + \omega_2\}, \\ \omega_i &= \sum_{j=1}^{r_i}\phi(\lambda_i; t_{ij}), \qquad \phi(\lambda_i; t) = e^{\lambda_i t} - 1. \end{aligned} \quad (2.4)$$

Suppose that p, λ_1 and λ_2 are independent random variables. The joint prior pdf of the random vector $(p, \lambda_1, \lambda_2)$ is thus given by

$$g(p, \lambda_1, \lambda_2) = g_1(\lambda_1) g_2(\lambda_2) g_3(p),$$

where, for $i = 1, 2$, $g_i(\lambda_i)$ is a prior pdf of Λ_i and $g_3(p)$ is a prior pdf of p. Let $p \sim Beta(a, b)$ and we choose the random variable Λ_i, $i = 1, 2$, to follow a left truncated exponential density with parameter γ_i, as used by AL-Hussaini et al. (2000). Its pdf is

$$g_i(\lambda_i) = \frac{1}{\gamma_i} \exp\left\{ \frac{-(\lambda_i - 1)}{\gamma_i} \right\}, \quad \lambda_i > 1, \tag{2.5}$$

where, for $i = 1, 2$, $\gamma_i > 0$.

Based on these considerations, the joint prior density of p, λ_1 and λ_2 is given by

$$g(p, \lambda_1, \lambda_2) = \frac{1}{B(a,b)} p^{a-1} (1-p)^{b-1} \frac{1}{\gamma_1 \gamma_2} \exp\left\{ - \sum_{i=1}^{2} \frac{(\lambda_i - 1)}{\gamma_i} \right\}, \tag{2.6}$$

where a, b and $\gamma_i, i = 1, 2$ are real numbers and $B(.,.)$ stands for the beta function. The prior (2.6) is so chosen that it would be rich enough to cover the prior belief of the experimenter.

The posterior pdf of p, λ_1 and λ_2 given $\underline{t}$ is obtained by combining (2.3) and (2.6), and written as

$$\begin{aligned} q(p, \lambda_1, \lambda_2 \mid \underline{t}) &= A \sum_{j_1=0}^{n-r} \binom{n-r}{j_1} p^{\delta_1 - 1} (1-p)^{\delta_2 - 1} \\ &\quad \times \exp\left[\sum_{i=1}^{2} \left\{ \lambda_i \left(T_i - \frac{1}{\gamma_i} \right) - B_{ij_1}(\lambda_i) \right\} \right], \end{aligned} \tag{2.7}$$

where

$$\begin{aligned} \delta_1 &= n - r_2 - j_1 + a, \\ \delta_2 &= r_2 + j_1 + b, \end{aligned} \tag{2.8}$$

A is a normalizing constant given by

$$\begin{aligned}
A^{-1} &= \int_0^1 \int_0^\infty \int_0^\infty q(p, \lambda_1, \lambda_2 \mid \underline{t}) d\lambda_1 d\lambda_2 dp \\
&= \sum_{j_1=0}^{n-r} \binom{n-r}{j_1} \left(\int_0^1 p^{\delta_1 - 1} (1-p)^{\delta_2 - 1} dp \right) \\
&\times \prod_{i=1}^{2} \int_0^\infty \exp\left[\lambda_i \left(T_i - \frac{1}{\gamma_i} \right) - B_{ij_1}(\lambda_i) \right] d\lambda_i \\
&= \sum_{j_1=0}^{n-r} \binom{n-r}{j_1} B(\delta_1, \delta_2) \prod_{i=1}^{2} I_{ij_1}(\underline{t}), \qquad (2.9)
\end{aligned}$$

and for $i = 1, 2$,

$$I_{ij_1}(\underline{t}) = \int_0^\infty \exp\left[\lambda_i \left(T_i - \frac{1}{\gamma_i} \right) - B_{ij_1}(\lambda_i) \right] d\lambda_i. \qquad (2.10)$$

A future sample of size m is assumed to be independent of the past (informative) sample of size n and is obtained from the same population with pdf (1.3). Let Y_s be the ordered lifetime of the s^{th} components to fail in a future sample of size m, $1 \leq s \leq m$. The s^{th} order statistic in a sample of size m represents the life length of a $(m - s + 1)$ out of m system which is an important technical structure in reliability theory. The distribution function of Y_s is given (see Arnold et al., 1992) by

$$\begin{aligned}
&F_{Y_s}(y_s \mid p, \lambda_1, \lambda_2) \\
&= \sum_{\ell=s}^{m} \binom{m}{\ell} [F_T(y_s \mid p, \lambda_1, \lambda_2)]^\ell [1 - F_T(y_s \mid p, \lambda_1, \lambda_2)]^{m-\ell}, \\
&= \sum_{\ell=s}^{m} \sum_{j_2=0}^{\ell} (-1)^{j_2} \binom{m}{\ell} \binom{\ell}{j_2} [R(y_s)]^{m-\ell+j_2}, \qquad (2.11)
\end{aligned}$$

where $F_T(y_s \mid p, \lambda_1, \lambda_2) = 1 - R(y_s)$ is the distribution function of the mixture model and $R(y_s)$ is as given by (2.2), after replacing t by y_s.

Expanding $[R(y_s)]^{m-\ell+j_2}$, by using the binomial expansion, and rearranging, it follows that

$$\begin{aligned}
&F_{Y_s}(y_s \mid p, \lambda_1, \lambda_2) \\
&= \sum_{\ell=s}^{m} \sum_{j_2=0}^{\ell} (-1)^{j_2} \binom{m}{\ell} \binom{\ell}{j_2} \\
&\times \left[p e^{-\frac{1}{\lambda_1}\phi(\lambda_1; y_s)} + (1-p) e^{-\frac{1}{\lambda_2}\phi(\lambda_2; y_s)} \right]^{m-\ell+j_2} \\
&= \sum_{\ell=s}^{m} \sum_{j_2=0}^{\ell} (-1)^{j_2} \binom{m}{\ell} \binom{\ell}{j_2} p^{m-\ell+j_2} e^{-\frac{m-\ell+j_2}{\lambda_1}\phi(\lambda_1; y_s)} \\
&\times \sum_{j_3=0}^{m-\ell+j_2} \binom{m-\ell+j_2}{j_3} \left(\frac{(1-p) e^{-\frac{1}{\lambda_2}\phi(\lambda_2; y_s)}}{p e^{-\frac{1}{\lambda_1}\phi(\lambda_1; y_s)}} \right)^{j_3}.
\end{aligned}$$

Or, equivalently,

$$\begin{aligned}
&F_{Y_s}(y_s \mid p, \lambda_1, \lambda_2) \\
&= \sum \Omega p^{\delta_3} (1-p)^{j_3} \exp\left[-\frac{\delta_3}{\lambda_1}\phi(\lambda_1; y_s) - \frac{j_3}{\lambda_2}\phi(\lambda_2; y_s) \right],
\end{aligned} \tag{2.12}$$

where

$$\begin{aligned}
\delta_3 &= m - \ell + j_2 - j_3, \\
\sum &= \sum_{\ell=s}^{m} \sum_{j_2=0}^{\ell} \sum_{j_3=0}^{m-\ell+j_2}, \\
\Omega &= (-1)^{j_2} \binom{m}{\ell} \binom{\ell}{j_2} \binom{m-\ell+j_2}{j_3}.
\end{aligned} \tag{2.13}$$

The Bayes predictive pdf of y_s given $\underline{t}$ (see Aitchison and Dunsmore, 1975), is defined by

$$f^*(y_s \mid \underline{t}) = \int_0^1 \int_0^\infty \int_0^\infty f(y_s \mid p, \lambda_1, \lambda_2) q(p, \lambda_1, \lambda_2 \mid \underline{t}) d\lambda_1 d\lambda_2 dp, \tag{2.14}$$

where $q(p, \lambda_1, \lambda_2 \mid \underline{t})$ is the posterior pdf, given by (2.7) and $f(y_s \mid p, \lambda_1 \lambda_2)$ is the pdf of the s-th component in a future sample which can be obtained from (2.12).

Bayesian prediction bounds for y_s can be obtained, for a given value of ν, by computing

$$\begin{aligned} Pr[Y_s \geq \nu \mid \underline{t}] &= \int_{\nu}^{\infty} f^*(y_s \mid \underline{t}) dy_s, \\ &= 1 - \int_0^1 \int_0^{\infty} \int_0^{\infty} F_{Y_s}(\nu \mid p, \lambda_1 \lambda_2) \\ &\quad \times q(p, \lambda_1 \lambda_2 \mid \underline{t}) d\lambda_1 d\lambda_2 dp, \end{aligned} \tag{2.15}$$

where $F_{Y_s}(. \mid p, \lambda_1\lambda_2)$ is the cumulative distribution function (cdf) of the s-th component in a future sample given by (2.12).

Substituting (2.7) and (2.12) in (2.15), we obtain

$$\begin{aligned} &Pr[Y_s \geq \nu \mid \underline{t}] \\ &= 1 - A \sum_{j_1=0}^{n-r} \sum \Omega \binom{n-r}{j_1} \int_0^1 \int_0^{\infty} \int_0^{\infty} p^{\delta_1+\delta_3-1} (1-p)^{\delta_2+j_3-1} \\ &\quad \times \prod_{i=1}^{2} \exp\left[\lambda_i \left(T_i - \frac{1}{\gamma_i}\right) - B^*_{ij_1}(\lambda_i; \nu)\right] d\lambda_1 d\lambda_2 dp \\ &= 1 - A \sum^{*} \Omega^* B(\delta_1+\delta_3, \delta_2+j_3) \prod_{i=1}^{2} I^*_{ij_1}(\nu), \end{aligned} \tag{2.16}$$

where

$$\begin{aligned} \sum^{*} &= \sum_{j_1=0}^{n-r} \sum, \quad \Omega^* = \Omega \binom{n-r}{j_1}, \\ B^*_{1j_1}(\lambda_1; \nu) &= B_{1j_1}(\lambda_1) + \frac{\delta_3}{\lambda_1} \phi(\lambda_1; \nu), \\ B^*_{2j_1}(\lambda_2; \nu) &= B_{2j_1}(\lambda_2) + \frac{j_3}{\lambda_2} \phi(\lambda_2; \nu), \end{aligned} \tag{2.17}$$

for $i = 1, 2$,

$$I^*_{ij_1}(\nu) = \int_0^{\infty} \exp\left[\lambda_i (T_i - \frac{1}{\gamma_i}) - B^*_{ij_1}(\lambda_i; \nu)\right] d\lambda_i, \tag{2.18}$$

and δ_3, $\sum$ and Ω are given by (2.13).

It can be shown, see Appendix A, that $Pr[Y_s \geq 0 \mid \underline{t}] = 1$.

A $100\tau\%$ prediction interval for Y_s is then given by

$$P[L(\underline{t}) < Y_s < U(\underline{t})] = \tau, \tag{2.19}$$

where $L(\underline{t})$ and $U(\underline{t})$ are obtained, respectively, by solving the following two equations

$$P[Y_s > L(\underline{t})] = \frac{1+\tau}{2}, \qquad P[Y_s > U(\underline{t})] = \frac{1-\tau}{2}. \tag{2.20}$$

3 Numerical computations

In this section, a numerical example is given to illustrate the results and the accuracy of the prediction intervals is investigated via Monte Carlo simulation.

3.1 Example

A numerical example is given to illustrate the computations in the two particular cases of predicting the minimum and maximum of future observables. We compute the lower and upper prediction bounds for Y_1 and Y_m, the first and last failure times in a future sample of size $m = 10$. The failure times are assumed to follow a finite mixture of two-component Gompertz lifetime model with *pdf* given by (1.3). We assumed the prior parameters $a = 4.5$, $b = 3.2$, $\gamma_1 = 3.0$ and $\gamma_2 = 2.0$. The prior information about p, λ_1 and λ_2 suggests that $p = 0.579$, $\lambda_1 = 3.274$ and $\lambda_2 = 1.589$. For $n = 15$ and $t_0 = 0.5$, the following sample is obtained

$$\begin{array}{llllllll} .0377, & .1198, & .2374, & .3201, & .3492, & .3527, & .4882, & (r_1 = 7), \\ & & & & & .1240, & .1345, & (r_2 = 2). \end{array}$$

This sample is generated from the mixture of Gompertz components in such a way that $t_{ij} < t_0, j = 1, 2, ..., r_i, i = 1, 2$. If for $i = 1, 2$, $U_i \sim \text{Uniform}(0, 1)$, then $T_i = \frac{1}{\lambda_i} \log[1 - \lambda_i \log(1 - U_i)] \sim Gomp(\lambda_i)$. An observation t_{ij} belongs to subpopulation 1 if $u_{ij} \leq p$ and to subpopulation 2 if $u_{ij} > p$, where u_{ij} and t_{ij} are the j^{th} values of the i^{th} random variables U_i and T_i, respectively, $j = 1, 2, ..., r_i, i = 1, 2$.

Using these data in Equation (2.20) with $\tau = 0.95$, the lower and upper 95% prediction bounds for Y_1, the first failure time, are 0.01349 and 0.15271,

		Y_1		Y_m	
n	m	AW	CP	AW	CP
15	10	0.1408	0.9383	0.6921	0.9396
	13	0.1413	0.9427	0.6983	0.9437
20	15	0.1392	0.9443	0.6582	0.9455
	22	0.1399	0.9468	0.6545	0.9474
30	25	0.1285	0.9482	0.6507	0.9485
	35	0.1306	0.9487	0.6496	0.9489

Table 1: Simulation results for $\tau = 0.95$

respectively. Whereas the 95% prediction bounds for Y_{10}, the last failure time, are given by 0.20183 and 0.64319, respectively.

3.2 Monte Carlo simulation

The behavior of the Bayes prediction bounds derived in Section 2 is investigated via Monte Carlo simulations according to the following steps:

1. For given values of $a = 4.5$, $b = 3.2$, $\gamma_1 = 3.0$ and $\gamma_2 = 2.0$, 1000 pairs of samples $(T_{11}, ..., T_{1r_1}, T_{21}, ..., T_{2r_2})$ and $(Y_1, Y_2, ..., Y_m)$ of different sizes n and m were generated from the mixture model of two Gompertz with pdf (1.3) for given $t_0 = 0.5$.

2. For each pair of samples, the $100\tau\%$ Bayesian prediction interval for Y_1 and Y_m was computed by solving numerically equations (2.20). The subroutine "ZSPOW" from the IMSL is used for this purpose. The lengths of the intervals and the number of intervals containing Y_1 and Y_m were obtained.

3. The average interval lengths from the 1000 pairs of samples were computed and the proportions of intervals containing Y_1 (and Y_m) were obtained as estimates of the mean interval lengths (or average width **AW**) and coverage probabilities (**CP**), respectively.

4. Steps 1-3 were performed for $\tau = 0.95$ and 0.99.

		Y_1		Y_m	
n	m	AW	CP	AW	CP
15	10	0.1546	0.9721	0.8385	0.9735
	13	0.1574	0.9767	0.8394	0.9793
20	15	0.1501	0.9811	0.7409	0.9834
	22	0.1521	0.9837	0.7516	0.9848
30	25	0.1473	0.9858	0.7251	0.9875
	35	0.1495	0.9877	0.7263	0.9883

Table 2: Simulation results for $\tau = 0.99$

4 Concluding remarks

In this paper, the finite mixture of two-component Gompertz is proposed to be the underlying lifetime model from which observables are to be predicted by using a Bayesian approach. The two-sample prediction technique is used and type-I censored data from the mixture model is considered. It has been noticed, from Tables 1 and 2, that the prediction intervals are affected by increasing n and in this case the coverage probability's are quite close to the confidence levels 95% and 99%, and therefore the intervals tend to perform very well in terms of simulated coverage probabilities. The average interval width tends to decrease as n increases and then the prediction intervals become better as n gets larger.

Different values of the prior parameters a, b, γ_1 and γ_2 rather than those appearing in the above tables have been considered but did not change the previous conclusion. On the other hand, when the prior parameters are unknown, the empirical Bayes approach may be used to estimate such parameters, (see, for example, Maritz and Lwin, 1989).

Appendix A

To show that $Pr[Y_s \geq 0 \mid \underline{t}] = 1$, we notice, from (2.4) and (2.17), that for $\nu = 0$,

$$\begin{aligned} B^*_{1j_1}(\lambda_1; 0) &= B_{1j_1}(\lambda_1), \\ B^*_{2j_1}(\lambda_2; 0) &= B_{2j_1}(\lambda_2). \end{aligned} \tag{A.1}$$

It follows from (2.10) and (2.18), that for $i = 1, 2$ and $\nu = 0$,

$$I^*_{ij_1}(0) = I_{ij_1}(\underline{t}). \tag{A.2}$$

By using (A.2) in (2.16), we obtain

$$Pr[Y_s \geq 0 \mid \underline{t}] \;=\; 1 - A \sum^{*} \Omega^* B(\delta_1 + \delta_3, \delta_2 + j_3) \prod_{i=1}^{2} I_{ij_1}(\underline{t}). \tag{A.3}$$

Substituting from (2.17) in (A.3), we have

$$\begin{aligned}
&Pr[Y_s \geq 0 \mid \underline{t}] \\
&= 1 - A \sum_{j_1=0}^{n-r} \sum_{\ell=s}^{m} \sum_{j_2=0}^{\ell} \sum_{j_3=0}^{m-\ell+j_2} (-1)^{j_2} \binom{n-r}{j_1} \binom{m}{\ell} \binom{\ell}{j_2} \\
&\quad \times \binom{m-\ell+j_2}{j_3} B(\delta_1 + \delta_3, \delta_2 + j_3) \prod_{i=1}^{2} I_{ij_1}(\underline{t}), \\
&= 1 - A \sum_{j_1=0}^{n-r} \binom{n-r}{j_1} B(\delta_1, \delta_2) \\
&\quad \times \prod_{i=1}^{2} I_{ij_1}(\underline{t}) \Bigg\{ \sum_{\ell=s}^{m} \sum_{j_2=0}^{\ell} (-1)^{j_2} \binom{m}{\ell} \binom{\ell}{j_2} \Bigg\} \\
&\quad \times \left[\sum_{j_3=0}^{m-\ell+j_2} \binom{m-\ell+j_2}{j_3} p^{m-\ell+j_2-j_3} (1-p)^{j_3} \right].
\end{aligned}$$

Or, equivalently,

$$Pr[Y_s \geq 0 \mid \underline{t}] = 1 - \sum_{\ell=s}^{m} \binom{m}{\ell} \Bigg\{ \sum_{j_2=0}^{\ell} (-1)^{j_2} \binom{\ell}{j_2} \Bigg\}. \tag{A.4}$$

Making use of the combinatorial identity $\sum_{i=0}^{j} (-1)^i \binom{j}{i} = 0$ in Equation (A.4), we obtain

$$Pr[Y_s \geq 0 \mid \underline{t}] = 1. \tag{A.5}$$

Acknowledgements

The author thanks the associate editor and referee for their helpful remarks and comments, which improved the original manuscript.

References

AHMAD, K. E. (1988). Identifiability of finite mixtures using a new transform. *Ann. Inst. Statist. Math.*, 40(2):261–265.

AITCHISON, J. and DUNSMORE, I. R. (1975). *Statistical Prediction Analysis*. Cambridge University Press, Cambridge.

AL-HUSSAINI, E. K. and AHMAD, K. E. (1981). On the identifiability of finite mixtures of distributions. *IEEE Trans. Inform. Theory*, IT-27(5):664–668.

AL-HUSSAINI, E. K., AL-DAYIAN, G. R., and ADHAM, S. A. (2000). On finite mixture of two-component Gompertz lifetime model. *J. Statist. Comput. Simul.*, 67(1):1–14.

ARNOLD, B. C., BALAKRISHNAN, N., and NAGARAJA, H. N. (1992). *A First Course in Order Statistics*. John Wiley and Sons, New York.

EVERITT, B. S. and HAND, D. J. (1981). *Finite Mixture Distributions*. University Press, Cambridge.

GEISSER, S. (1993). *Predictive Inference: An Introduction*. Chapman and Hall, London.

GOMPERTZ, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new model of determining the value of life contingencies. *Philos. Trans. Roy. Soc. London*, 115:513–585.

GORDON, N. H. (1990a). Application of the theory of finite mixtures for the estimation of cure rates of treated cancer patients. *Statistics in Medicine*, 9:397–407.

GORDON, N. H. (1990b). Maximum likelihood estimation for mixtures of two Gompertz distributions when censoring occurs. *Commun. Statist. Simula. Comput.*, 19(2):733–747.

IMSL (1984). *IMSL, Reference Manual*. IMSL, Inc., Houston.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1995). *Continuous Univariate Distributions*, vol. 2. John Wiley and Sons, New York, 2nd ed.

Lindsay, B. G. (1995). *Mixture Models: Theory, Geometry and Applications*, vol. 5 of *NSF-CBMS Regional Conference Series in Probability and Statistics*. Institute of Mathematical Statistics, California.

Maritz, J. S. and Lwin, T. (1989). *Empirical Bayes Methods*. Chapman and Hall, Bocaraton, 2nd ed.

McLachlan, G. J. and Basford, K. E. (1988). *Mixture Models: Inferences and Applications to Clustering*. Marcel Dekker, New York.

McLachlan, G. J. and Peel, D. (2000). *Finite Mixture Models*. John Wiley and Sons, New York.

Mendenhall, W. and Hader, R. J. (1958). Estimation of parameters of mixed exponentially distributed failure time distributions from censored life test data. *Biometrika*, 45:504–520.

Osman, M. I. (1987). *A new model for analyzing the survival of heterogeneous data*. Ph. D. thesis, Case Western Reserve University, USA.

Teicher, H. (1961). Identifiability of mixtures. *Ann. Math. Statist.*, 33:244–248.

Teicher, H. (1963). Identifiability of finite mixtures. *Ann. Math. Statist.*, 34:1265–1269.

Titterington, D. M., Smith, A. F. M., and Makov, U. E. (1985). *Statistical Analysis of Finite Mixture Distributions*. John Wiley and Sons, New York.

Wu, J. W. and Lee, W. C. (1999). Characterization of the mixtures of Gompertz distributions by conditional expectations of order statistics. *Biometrical Journal*, 41(3):371–381.