

Bayesian Survival Analysis with the Extended Generalized Gamma Model: Application to Demographic and Health Survey Data



Yuli Liang and Gebrenegus Ghilagaber

Abstract We extend the existing family of flexible survival models by assembling models scattered across the literature into a more knit form and under the same umbrella. New special cases are obtained not only by constraining the shape and scale parameters of the extended generalized gamma (EGG) model to fixed constants, but also by imposing relationships (such as equality, reciprocal, and negative reciprocal) between them. Apart from common parametric distributions such as exponential, Weibull, gamma, and log normal, the further extended family includes Rayleigh, inverse Rayleigh, ammag, inverse ammag, and half-normal distributions. The models are applied, in a Bayesian framework, on time to entry into first marriage among Eritrean men and women based on data from the 2010 Population and Health Survey. The application demonstrates that the further extended family of distributions provides a wide range of alternatives for a baseline distribution in the analysis of survival data. The empirical results reveal that the inverse gamma model fits best the data for men. It also performs closely as good as the EGG model in the data for women as well as in the combined sample.

Keywords Extended generalized gamma (EGG) model · Parametric survival models · Accelerated failure-time (AFT) models · Proportional hazards models · Bayesian inference · Parametric inference · Likelihood ratio test · Eritrea · Entry into marriage · Nuptiality · Demographic and Health Surveys (DHS) · Education · Residence · Exponential distribution · Inverse exponential distribution · Weibull distribution · Reciprocal Weibull distribution · Rayleigh distribution · Inverse Rayleigh distribution · Standard and generalized gamma

Y. Liang

School of Business and Economics, Department of Economics and Statistics, Linnaeus University, Växjö, Sweden
e-mail: Yuli.Liang@lnu.se

G. Ghilagaber (✉)

Department of Statistics, Stockholm University, Stockholm, Sweden
e-mail: Gebre@stat.su.se

distributions · Inverse gamma distribution · Ammag distribution · Inverse Ammag distribution · Log normal distribution · Half normal distributions · Log predictive density score (LPDS) · Censoring · Bayes Factors · Time to event data · Model comparison · Markov chain Monte Carlo (MCMC) · Random walk · Metropolis Hasting algorithm · Block sampling · Posterior distribution · Ergodic mean theorem · Inefficiency factor (IF)

1 Introduction

The usual goals in the analysis of survival data include: (a) describing the distributional shape of the time variable; (b) comparing the survival experiences of different groups in a population; and (c) modeling the relationship between explanatory variables and survival time—as measured by time to the event of interest or the rate at which the event occurs.

Two classes of models are common in the literature for investigating effects of explanatory variables on survival. In the Cox proportional hazards models, the explanatory variables act multiplicatively on a baseline hazard so that their effect is to increase or decrease the hazard relative to that of the baseline group. A second class of models, known as the accelerated failure-time models, specifies the covariates to act multiplicatively on time to event itself so that their effect is to accelerate or decelerate time to event relative to an event time for baseline group. According to Wei (1992), the accelerated failure-time model has an intuitive physical interpretation and would be a useful alternative to the Cox PH model in survival analysis.

It has been documented that covariate effects on survival time are not robust to the choice of the baseline distribution—see, for instance (Addison and Portugal, 1987; Bergström and Edin, 1992; Bergström et al., 1994; Ghilagaber, 2005). It is, therefore, of paramount importance to correctly specify the baseline distribution if results from analysis of survival data are to be utilized optimally. A number of distributions for survival data are available in the literature scattered across disciplines and application areas. Some previous works have attempted to put these scattered models in a more knit form by embedding a number of competing models under the umbrella of a general parametric framework as in Butler and McDonald (1986) and Peng et al. (1998). This enables the use of ordinary parametric inference for assessment of each competing model relative to a more comprehensive one. Among others, (Ghilagaber, 2005) shows that five parametric duration models (exponential, Weibull, gamma, log normal, and reciprocal Weibull) may be treated as special cases of a more general extended generalized gamma (EGG) model by constraining the shape and/or scale parameters of the EGG model to some fixed constants.

In this chapter, we extend the EGG model further and increase the family of flexible distributions to include 13 special cases. This is achieved by including distributions that not only constrain the shape and scale parameters to specified

constants but also impose some relationships between them. The new set of special cases include the Rayleigh and inverse Rayleigh distributions as well as the ammag and inverse ammag distributions as described in Cox et al. (2007). Further, a half-normal distribution can be obtained as a special case of ammag distribution.

A Bayesian approach is used to fit the EGG model and its 13 special cases to data on time to entry into first marriage among Eritrean men and women. Each special case model is then tested relative to a more general model using the log predictive density score (LPDS) in a Bayesian approach, see Li et al. (2010). Compared to the classical likelihood inference approach, the Bayesian approach provides three main advantages. First, we sample from a posterior density using Markov Chain Monte Carlo (MCMC), and hence, we can make exact inference for any sample size in any parametric survival models of various complexities. Second, we do not need to worry about the problem of local maximum trapping since our algorithm can go through the whole parameter spaces supported by the data. Third, it is straightforward to investigate the performance of joint posterior density, whereas in a frequentist paradigm, we need to run simulation by pre-specifying the true values of parameters when evaluating the performance of maximum likelihood estimates.

In Sect. 2, we introduce the accelerated failure-time models and demonstrate how a number of common distributions can be brought under the umbrella of the EGG model. Bayesian density estimation of the EGG model and MCMC implementation is described in Sect. 3. In Sect. 4, we illustrate the models of Sect. 2 and the methods of Sect. 3 using real-life data from the 2010 Eritrean Population and Health Survey. Section 5 concludes the chapter by way of summary and concluding remarks. A full list of the distributions used in this chapter, a proof for a lemma, and the R code used in the illustrative example are provided in Appendices.

2 Parametric Models for Survival Data

2.1 Background

Survival data contain information on durations until event or censoring (t_1, t_2, \dots, t_n) together with a censoring indicator as well as background variables or covariates (z_1, z_2, \dots, z_p) that are often socio-demographic characteristics of individuals or organizations. The distribution of survival time, T , may be described by its three equivalent functions: the survival function, $S(t) = P(T > t)$, the density function, $f(t)$, or the hazard (intensity) function, $h(t) = f(t)/S(t)$, where the last two functions require absolute continuity.

These functions can vary not only over time, but also among individuals within a population. Thus, one objective in the analysis of survival data is to draw inferences about the influence of covariates on these functions. One popular model is the Cox proportional hazards model presented in Cox (1972) where a p -dimensional vector of covariates \mathbf{z} affects the hazard function in a multiplicative manner according to

$$h(t|\mathbf{z}) = h_0(t) \exp(\mathbf{z}'\boldsymbol{\beta}), \quad (1)$$

where $h_0(t)$ is an unspecified baseline function of time and $\boldsymbol{\beta} \in R^p$ is an unknown vector of parameters representing the effect of the covariates \mathbf{z} . The factor $\exp(\mathbf{z}'\boldsymbol{\beta})$ describes the intensity (hazard) for an individual with vector \mathbf{z} relative to that of a standard individual (with $\mathbf{z} = \mathbf{0}$).

2.2 Accelerated Failure-Time Models

A second class of models, the accelerated failure-time model, specifies the covariates to act multiplicatively on the event time itself rather than on the hazard function.

If T_0 is the random time to event associated with an individual in the baseline group ($\mathbf{z} = \mathbf{0}$), then the accelerated failure-time model specifies that for an individual with a non-zero vector of covariates \mathbf{z} , the event time is given by

$$T = T_0 \exp(\mathbf{z}'\boldsymbol{\beta}) \quad (2)$$

or equivalently

$$\ln(T) = \mathbf{z}'\boldsymbol{\beta} + \ln(T_0), \quad (3)$$

where, as before, T is the event time, \mathbf{z} is a vector of covariates, and $\boldsymbol{\beta}$ is a vector of regression parameters. Since covariates alter, by a scale factor, the rate at which an individual traverses the time axis, Eq. (2) is referred to as the accelerated failure-time model. Thus, in accelerated failure-time models, the effect of the explanatory variables is to accelerate or decelerate time to event relative to T_0 .

The model in (3) is a linear model with $\ln(T_0)$ playing the role of an error term with an underlying baseline distribution. Usually, a scale parameter δ is allowed in the model to give

$$\ln(T) = \mathbf{z}'\boldsymbol{\beta} + \delta \ln(T_0) = \mathbf{z}'\boldsymbol{\beta} + \delta\epsilon, \quad (4)$$

where a more conventional notation ϵ is used for the error term.

From (4), we note that $T = e^{\mathbf{z}'\boldsymbol{\beta}} T_0^\delta$. Thus, the survival function of T may be written in terms of that of T_0 :

$$S(t) = P(T > t) = P(e^{\mathbf{z}'\boldsymbol{\beta}} T_0^\delta > t) = P(T_0^\delta > t e^{-\mathbf{z}'\boldsymbol{\beta}}) = S_0(t e^{-\mathbf{z}'\boldsymbol{\beta}}), \quad (5)$$

where $S_0(\cdot)$ is the survival function of the baseline time with scale parameter δ , T_0^δ , and $e^{-\mathbf{z}'\boldsymbol{\beta}}$ is the accelerating/decelerating factor. In other words, the probability for an individual with covariate vector \mathbf{z} surviving beyond time t is the same as the probability for an individual in the baseline group ($\mathbf{z} = \mathbf{0}$) surviving beyond time

$te^{-z'\beta}$. A positive coefficient β shifts the time $te^{-z'\beta}$ to the left of t , while a negative β shifts the time $te^{-z'\beta}$ to the right of t if all components of $z > 0$. Accordingly, the density and hazard functions can also be written in terms of the baseline density and hazard:

$$f(t) = e^{-z'\beta} f_0(te^{-z'\beta})$$

$$h(t) = e^{-z'\beta} h_0(te^{-z'\beta}).$$

The distribution of T_0 in (4) may be selected from positive-valued distributions such as Weibull or log normal that, in turn, yield extreme-value and normal distributions for the error term ϵ . Below, we demonstrate how the list may be expanded by assembling various models under the same umbrella.

2.3 The Extended Generalized Gamma (EGG) Model

Stacy (1962) introduced the generalized gamma model that is useful in embedding competing models into a single parametric framework. This model is the distribution of T such that $\ln(T) = \mu + \delta\epsilon$, where $\mu \in R, \delta > 0$, and the random error term ϵ has the density

$$f(k, \epsilon) = \frac{1}{\Gamma(k)} \exp [k\epsilon - \exp(\epsilon)], k > 0,$$

where k is an additional shape parameter. Prentice (1974) showed that a shift of parameter of the form $q = k^{-\frac{1}{2}}$ leads to a standard normal distribution for T giving an interior point for $q = 0$ in the parameter space. The final model with parameters $\mu, q \in R$ and $\delta > 0$ can be written as $\ln(T) = \mu + \delta\epsilon$, where the error density function $f(q, \epsilon)$ is given by

$$f(q, \epsilon) = \begin{cases} \frac{|q|}{\Gamma(q^{-2})} (q^{-2})^{q^{-2}} \exp \{q^{-2} [q\epsilon - \exp(q\epsilon)]\}, & q \neq 0 \\ \frac{1}{\sqrt{2\pi}} \exp(-\frac{\epsilon^2}{2}), & q = 0. \end{cases} \tag{6}$$

The distribution of T when the error term has the density given in Eq. (6) is known as the extended generalized gamma (EGG) distribution, see, for instance (Ghilagaber, 2005; Ghilagaber et al., 2014).

As can be seen from the lower part of (6), the EGG model reduces to the standard normal distribution for ϵ when the shape parameter q is equal to zero. Accordingly, T will have a log-normal distribution. When the shape parameter $q = 1$, (6) reduces to

$$f(q, \epsilon) = \exp [\epsilon - \exp(\epsilon)],$$

which is the standard (type 1) extreme-value distribution. As $\ln(T)$ is a linear function of ϵ , it has the same (extreme-value) distribution as ϵ . Hence, $T = \exp(\mathbf{z}'\boldsymbol{\beta} + \delta\epsilon)$ as defined in Eq. (4) will have a Weibull distribution. If $q = 1$ and $\delta = 1$, then T has the exponential distribution as a special case of the Weibull distribution. The case of $q = -1$ corresponds to extreme maximum-value distribution for $\ln(T)$. This, in turn, corresponds to reciprocal Weibull distribution for T .

The case of $\delta = 1$ and $q > 0$ is also of interest. Farewell and Prentice (1977) argue that this gives the ordinary gamma distribution for T . Others, (Bergström and Edin, 1992; Bergström et al., 1994, 1997), argue that this did not hold in their case illustrations. Consequently, we shall relax this special case to $\delta = 1$ and $q \in R$ and label it the “gamma” distribution in our illustrative example. Below, we further extend the above family of distributions by imposing some relationships between the scale and shape parameters.

2.4 Further Extensions of the EGG Model

We begin with a baseline distribution for time to event, $T_0 \sim EGG(0, 1, q)$, and label it as standard generalized gamma distribution with density and survival functions given by

$$f_{EGG(0,1,q)}(t_0) = \frac{|q|}{t_0\Gamma(q-2)}(q^{-2}t_0^q)^{q-2} \exp(-q^{-2}t_0^q),$$

$$S_{EGG(0,1,q)}(t_0) = \begin{cases} 1 - \Phi(\ln t_0), & q = 0 \\ 1 - \gamma(q^{-2}, t_0^q q^{-2})/\Gamma(q-2), & q > 0 \\ \gamma(q^{-2}, t_0^q q^{-2})/\Gamma(q-2), & q < 0, \end{cases}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. By transformation, $t = e^\mu t_0^\delta \sim EGG(\mu, \delta, q)$, and T is said to have the extended generalized gamma distribution with shape parameter $\mu \in R$, scale parameter $\delta > 0$, and an additional index shape parameter $q \in R$. We denote this by $T \sim EGG(\mu, \delta, q)$, with density

$$f_{EGG}(t) = \frac{|q|}{t\delta\Gamma(q-2)} \left[q^{-2}(e^{-\mu}t)^{\frac{q}{\delta}} \right]^{q-2} \exp \left[-q^{-2}(e^{-\mu}t)^{\frac{q}{\delta}} \right] \quad (7)$$

$$= \begin{cases} \frac{1}{\delta t \sqrt{2\pi}} e^{-\frac{(\ln t - \mu)^2}{2\delta^2}} & q = 0 \\ \frac{q}{\delta} t^{\frac{q}{\delta}-1} \frac{1}{\Gamma(q^{-2})} \left[q^{-2} (e^{-\mu})^{\frac{q}{\delta}} \right]^{q^{-2}} (t^{\frac{q}{\delta}})^{q^{-2}-1} \exp \left[-q^{-2} (e^{-\mu})^{\frac{q}{\delta}} t^{\frac{q}{\delta}} \right] & q > 0 \\ -\frac{q}{\delta} t^{\frac{q}{\delta}-1} \frac{1}{\Gamma(q^{-2})} \left[q^{-2} (e^{-\mu})^{\frac{q}{\delta}} \right]^{q^{-2}} (t^{\frac{q}{\delta}})^{q^{-2}-1} \exp \left[-q^{-2} (e^{-\mu})^{\frac{q}{\delta}} t^{\frac{q}{\delta}} \right] & q < 0. \end{cases}$$

The component

$$\frac{1}{\Gamma(q^{-2})} \left[q^{-2} (e^{-\mu})^{\frac{q}{\delta}} \right]^{q^{-2}} (t^{\frac{q}{\delta}})^{q^{-2}-1} \exp \left[-q^{-2} (e^{-\mu})^{\frac{q}{\delta}} t^{\frac{q}{\delta}} \right]$$

in the above equation is the density of the gamma distribution for $t^{\frac{q}{\delta}}$ with a shape parameter q^{-2} and a rate parameter $q^{-2} (e^{-\mu})^{\frac{q}{\delta}}$. The next lemma gives the r th moment and the first four central moments of the EGG density. The following definitions of skewness and excess kurtosis are used:

$$S(T) = \frac{E [T - E(T)]^3}{V(T)^{3/2}},$$

$$K(T) = \frac{E [T - E(T)]^4}{V(T)^2} - 3,$$

where $V(T)$ is the variance.

Lemma 1 *If $T \sim EGG(\mu, \delta, q)$, then*

$$E(T^r) = \begin{cases} \frac{\Gamma(q^{-2} + r\frac{\delta}{q})}{\left(q^{-\frac{2\delta}{q}} e^{-\mu} \right)^r \Gamma(q^{-2})}, & \text{if } r\delta/q > -q^{-2}, \\ \infty & \text{otherwise.} \end{cases}$$

$$E(T) = \frac{\Gamma(q^{-2} + \frac{\delta}{q})}{q^{-\frac{2\delta}{q}} e^{-\mu} \Gamma(q^{-2})},$$

$$V(T) = \frac{\Gamma(q^{-2} + \frac{2\delta}{q})\Gamma(q^{-2}) - \Gamma^2(q^{-2} + \frac{\delta}{q})}{\Gamma^2(q^{-2})(q^{-\frac{2\delta}{q}} e^{-\mu})^2},$$

$$E [T - E(T)]^3 = 2\Gamma^3 \left(q^{-2} + \frac{\delta}{q} \right) - 3\Gamma \left(q^{-2} + \frac{2\delta}{q} \right) \Gamma \left(q^{-2} + \frac{\delta}{q} \right) \Gamma(q^{-2})$$

$$+ \Gamma \left(q^{-2} + \frac{3\delta}{q} \right) \Gamma^2(q^{-2}),$$

$$E [T - E(T)]^4 = -3\Gamma^4 \left(q^{-2} + \frac{\delta}{q} \right) + 6\Gamma \left(q^{-2} + \frac{2\delta}{q} \right) \Gamma^2 \left(q^{-2} + \frac{\delta}{q} \right) \Gamma(q^{-2}) - 4\Gamma \left(q^{-2} + \frac{3\delta}{q} \right) \Gamma \left(q^{-2} + \frac{\delta}{q} \right) \Gamma^2(q^{-2}) + \Gamma \left(q^{-2} + \frac{4\delta}{q} \right) \Gamma^3(q^{-2}).$$

A simplified proof of Lemma 1 is provided in Appendix 2.

From Lemma 1, we note that $S(T)$ and $K(T)$ are the functions of q and δ/q , implying both q and δ/q are shape parameters.

The survival function of t is then given by

$$S_{EGG}(t) = \begin{cases} 1 - \Phi \left(\frac{\ln t - \mu}{\delta} \right) & q = 0 \\ 1 - \gamma \left[q^{-2}, t^{\frac{q}{\delta}} q^{-2} (e^{-\mu})^{\frac{q}{\delta}} \right] / \Gamma(q^{-2}) & q > 0 \\ \gamma \left[q^{-2}, t^{\frac{q}{\delta}} q^{-2} (e^{-\mu})^{\frac{q}{\delta}} \right] / \Gamma(q^{-2}) & q < 0, \end{cases} \tag{8}$$

where $\gamma \left[q^{-2}, t^{\frac{q}{\delta}} q^{-2} (e^{-\mu})^{\frac{q}{\delta}} \right] / \Gamma(q^{-2})$ is the corresponding cumulative distribution function of the gamma distribution for $t^{\frac{q}{\delta}}$ when $q > 0$ and $\gamma \left[q^{-2}, t^{\frac{q}{\delta}} q^{-2} (e^{-\mu})^{\frac{q}{\delta}} \right]$ is a lower incomplete gamma function with the form of $\gamma(s, r) = \int_0^r x^{s-1} e^{-x} dx$ described in Abramowitz and Stegun (1964).

The EGG model redefined in Eqs. (7) and (8) is a rich and versatile model containing many special cases based on different combinations of q and δ .

Apart from those mentioned in the previous subsection, the list may be extended to include the inverse exponential ($q = -\delta = -1$), standard gamma ($q = \delta$), inverse gamma (when $q = -\delta$), ammag ($q = 1/\delta$), inverse ammag ($q = -1/\delta$), Rayleigh ($q = 1$ and $\delta = 1/2$), inverse Rayleigh ($q = -1$ and $\delta = 1/2$), and half-normal ($q = \sqrt{2}$ and $\delta = \sqrt{2}/2$).

EGG nests more special cases such as Maxwell–Boltzmann, but we have not included this in the present chapter since our focus is on the distribution of survival time T . Further, the equivalent distributions of some special cases are excluded. For instance, the inverse gamma model is equivalent to the Levy model in some special cases: inverse gamma($q^{-2}, q^{-2}e^{-\mu}$) \leftrightarrow Levy(0, c) when $q^{-2} = 1/2$ and $q^{-2}e^{-\mu} = 2c$. The standard gamma model is also equivalent to a chi-squared model in some situations: standard gamma($q^{-2}, q^{-2}e^{-\mu}$) \leftrightarrow $\chi^2_{(v)}$ when $q^{-2} = v/2$ and $q^{-2}e^{-\mu} = 1/2$.

To sum up, the EGG model constitutes of at least 13 special cases whose relationships are depicted in Fig. 1. Each special case model can be assessed relative to a more comprehensive one using appropriate procedures for comparing nested models. A summary of the density functions, $f(t)$, and survival functions, $S(t)$, for 13 special cases is provided in Appendix 1. The corresponding hazard functions can be obtained by $h_{EGG}(t) = f_{EGG}(t)/S_{EGG}(t)$. The hazards in the EGG models

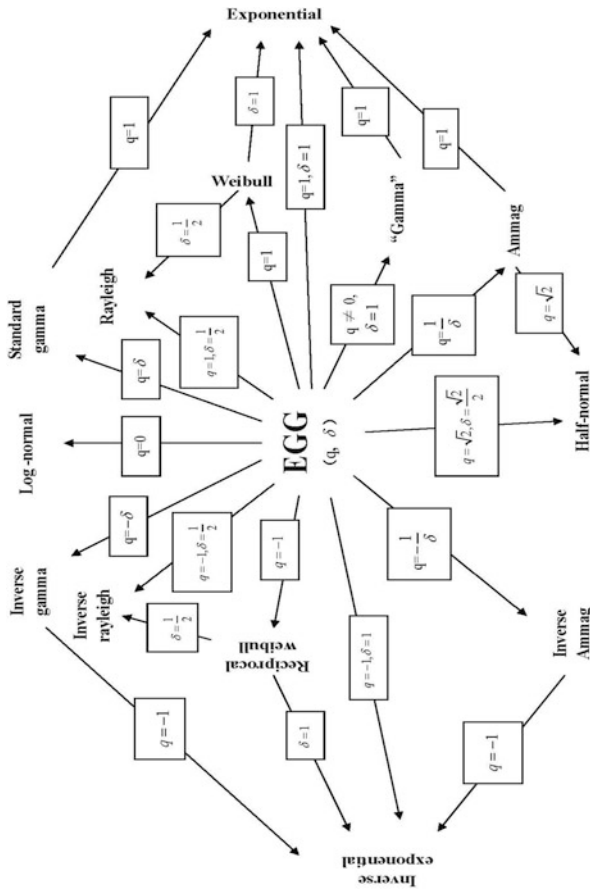


Fig. 1 Relationships among EGG family of distributions

can be of various forms—increasing, decreasing, bathtub, or arc-shaped (Cox et al., 2007).

When we adapt the generalized gamma distribution to accelerated failure-time models, the location parameter μ can be composed of a linear predictor based on p covariates $\mu = \beta_0 + \sum_{i=1}^n z_{ji} \beta_j$ ($j = 1 \cdots p$), which justifies the feasibility of the EGG in accelerated failure-time models.

The distribution of $\epsilon = \ln(T_0)$ is given in Eq. (6). When $q = 0$, ϵ is standard normal distributed; when $q \neq 0$, it can be manipulated to give

$$f(\epsilon; q) = |q| e^{q\epsilon} \frac{(q^{-2})^{q^{-2}}}{\Gamma(q^{-2})} (e^{q\epsilon})^{q^{-2}-1} \exp \left[-q^{-2} \exp(q\epsilon) \right] \quad (9)$$

with the corresponding survival functions

$$S(\epsilon, q) = \begin{cases} 1 - \Phi(\epsilon) & q = 0 \\ 1 - \gamma \left[q^{-2}, \exp(q\epsilon)q^{-2} \right] / \Gamma(q^{-2}) & q > 0 \\ \gamma \left[q^{-2}, \exp(q\epsilon)q^{-2} \right] / \Gamma(q^{-2}) & q < 0. \end{cases} \quad (10)$$

Based on the density of ϵ , Fig. 2 shows the shape of some density functions, $f(\epsilon)$, for some selected values of q . Here, we have a special case of $\ln(T) = \mathbf{z}'\boldsymbol{\beta} + \delta\epsilon = \mu + \delta\epsilon$

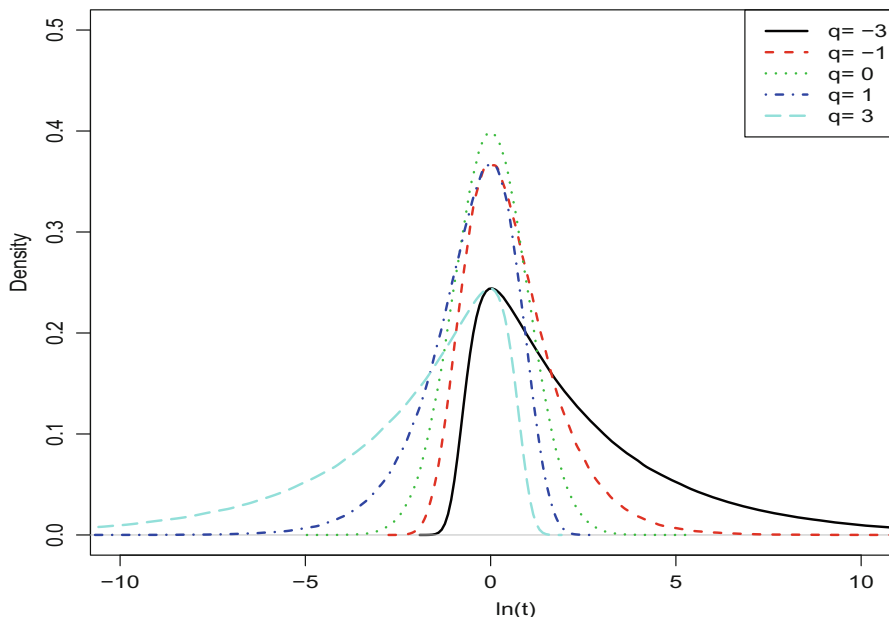


Fig. 2 Five distributions of $\ln(T)$ for $\mu = 0$, $\delta = 1$, and some values of q

in which $\mu = 0$ and $\delta = 1$. We note that the densities are positively skewed for $q < 0$ and negatively skewed for $q > 0$ with both the absolute skewness and kurtosis monotone increasing in $|q|$ —which are in accordance with those of Prentice (1974).

3 Bayesian Inference in the Extended Generalized Gamma Model

Bayesian inference for a three-parameter EGG model and four-parameter generalized gamma distribution (EGG model with one extra location parameter) is discussed in Tsonas (2001) and Van Noortwijk (2001) for situations where there is no censoring. Inference becomes more complicated in the presence of censored observations due to, for instance, difficulty to find conjugate prior or derive full conditional posterior.

Heleno and Alberto (1986) have used Bayesian approach for EGG model with censored data using Jeffrey multi-parameter prior. Ramos et al. (2017) have shown that both the Jeffreys prior and the reference priors give improper posteriors to the EGG model, and then proposed the overall reference prior in Berger et al. (2015), which provided the proper posterior. In this section, we present Bayesian inference in the EGG model that allows for any type of censoring mechanism.

3.1 Prior and Posterior

In a Bayesian framework, any prior information about the parameters of interest is combined with the data (likelihood) to derive a posterior distribution.

In our present case, we use normal priors with mean 0 and large variance σ_1^2 for each effect parameters $\beta_j (j = 0, \dots, p)$. We also assume a vague prior, a gamma distribution with hyperparameters a and b for the scale parameter δ . For the index shape parameter q , a normal prior with mean 0 and large variance σ_2^2 is assumed. These independent priors can be summarized as follows:

$$\begin{aligned} \beta_j &\sim N(0, \sigma_1^2), \quad j = 1, \dots, p \\ \delta &\sim \text{Gamma}(a, b) \\ q &\sim N(0, \sigma_2^2). \end{aligned}$$

We can use any prior that reflects our prior knowledge (if any) of the unknown parameters. In our illustration in Sect. 4, we will use $\sigma_1 = \sigma_2 = 1000$ and hyperparameters $a = b = 1$. The rationale behind this is to let the likelihood dominate the posterior so that the inferences drawn are driven by the data.

Denoting data with \mathcal{D} , the joint posterior distribution is then given by

$$\begin{aligned} f(\boldsymbol{\beta}, \delta, q|\mathcal{D}) &\propto L(\boldsymbol{\beta}, \delta, q; \mathcal{D})f(\boldsymbol{\beta}, \delta, q) \\ &= L(\boldsymbol{\beta}, \delta, q; \mathcal{D}) \prod_{j=1}^p f(\beta_j)f(\delta)f(q), \end{aligned}$$

where $L(\boldsymbol{\beta}, \delta, q; \mathcal{D})$ is the likelihood function, and $f(\cdot)$ is the prior density function of β_j , δ , and q with known hyperparameters. The above posterior can be generalized to other types of likelihood functions based on other censoring mechanisms (than the standard right censoring assumed in our present case). With right censored data, the likelihood function becomes

$$L(\boldsymbol{\beta}, \delta, q; \mathcal{D}) = \prod_{i=1}^n (\delta^{-1} f(\epsilon_i, q))^{d_i} S(\epsilon_i, q)^{1-d_i},$$

where d_i is the censoring indicator and $f(\epsilon_i, q)$ and $S(\epsilon_i, q)$ are given by Eqs. (9) and (10), respectively.

Since there is no explicit analytical form for the posterior distribution, sampling is performed using numerical methods based on Markov Chain Monte Carlo (MCMC).

3.2 MCMC: Random Walk Metropolis–Hastings Algorithm with Block Sampling

We sample all parameters sequentially from the joint posterior distribution using the Metropolis–Hastings algorithm. See Gelman et al. (2004) for more details on the Metropolis–Hastings algorithm and its nice properties. A random walk Metropolis–Hastings algorithm with block sampling is used, and the sampling procedure for the parameters $\boldsymbol{\theta} = (\boldsymbol{\beta}, \delta, q)'$ can be summarized as follows:

- (1) Set the initial values for the parameters $\boldsymbol{\theta}_0 = (\boldsymbol{\beta}_0, \delta_0, q_0)'$.
- (2) Construct the proposal distribution $J(\boldsymbol{\theta}_p|\boldsymbol{\theta}_c) \sim N(\boldsymbol{\theta}_c, c^2\boldsymbol{\Sigma})$, where $\boldsymbol{\theta}_p$ is the candidate value, $\boldsymbol{\theta}_c$ is the current value, and c is the scaling constant and $\boldsymbol{\Sigma}$ is a known covariance matrix. Here we choose $\boldsymbol{\Sigma} = -H^{-1}(\widehat{\boldsymbol{\theta}})$, where $H(\widehat{\boldsymbol{\theta}})$ is the Hessian matrix evaluated at $\widehat{\boldsymbol{\theta}}$, which is obtained by Newton's method. Following Gelman et al. (2004), we choose a value of $c = 2.4/\sqrt{k}$, where k is the length of the vector $\boldsymbol{\theta}$.
- (3) Generate $\boldsymbol{\theta}^*$ from $J(\boldsymbol{\theta}_p|\boldsymbol{\theta}_c)$ and U from $U(0, 1)$.
- (4) If

$$U < \frac{f(\boldsymbol{\theta}^*|\mathcal{D})f(\boldsymbol{\theta}^*)J(\boldsymbol{\theta}_c|\boldsymbol{\theta}_p)}{f(\boldsymbol{\theta}_c|\mathcal{D})f(\boldsymbol{\theta}_c)J(\boldsymbol{\theta}_p|\boldsymbol{\theta}_c)},$$

the candidate vector $\boldsymbol{\theta}^*$ is accepted and $\boldsymbol{\theta}_c = \boldsymbol{\theta}^*$; otherwise, we keep $\boldsymbol{\theta}_c$.

- (5) Return to step (2).

3.3 Posterior Statistics and Convergence Diagnostics

We summarize our posterior distribution by way of posterior means and highest posterior density (hpd). Since MCMC is based on ergodic mean theorem (Markov chain law of large numbers), convergence can be verified using diagnostic plots such as a plot of the cumulative mean against the number of iterations. In addition, inefficiency factors (IF) can be computed as a measure of the efficiency of the MCMC scheme.

3.4 Bayesian Model Comparisons

The common way of comparing models in the Bayesian framework is the use of Bayes factor that is the ratio of marginal likelihood of two competing models.

Suppose we have a set of candidate models $\mathcal{M}_m, m = 1, \dots, M$ and the corresponding model parameters θ_m . The posterior model probability is then given by

$$P(\mathcal{M}_m|Y) \propto P(\mathcal{M}_m)P(Y|\mathcal{M}_m),$$

where Y represents the data at hand. The posterior odds $P(\mathcal{M}_m|Y)/P(\mathcal{M}_l|Y)$ can be used to compare two models, and it can be written in terms of the Bayes factor:

$$\frac{P(\mathcal{M}_m)}{P(\mathcal{M}_l)} BF_{ml},$$

where BF_{ml} is the Bayes factor between \mathcal{M}_m and \mathcal{M}_l with the form

$$BF(Y) = \frac{P(Y|\mathcal{M}_m)}{P(Y|\mathcal{M}_l)} = \frac{\int P(Y|\theta_m, \mathcal{M}_m)P(\theta_m|\mathcal{M}_m)d\theta_m}{\int P(Y|\theta_l, \mathcal{M}_l)P(\theta_l|\mathcal{M}_l)d\theta_l}.$$

The marginal likelihood is a conditional expectation for the likelihood given the prior

$$E_{P(\theta_m|\mathcal{M}_m)}(P(Y|\theta_m, \mathcal{M}_m)).$$

It is sensitive to the choice of the prior, especially when the prior is not very informative (Villani et al., 2009). For instance, if $P(\theta_m|\mathcal{M}_m)$ is far from $P(Y|\theta_m, \mathcal{M}_m)$, while $P(\theta_l|\mathcal{M}_l)$ is close to $P(Y|\theta_l, \mathcal{M}_l)$, it is possible that $P(Y|\mathcal{M}_m)$ is less than $P(Y|\mathcal{M}_l)$ even though \mathcal{M}_l is a sub-model of \mathcal{M}_m .

To avoid such sensitivity to the choice of prior, we compare our models in the illustration on the basis of their predictive performance. The data is split randomly into B folds, and $B-1$ fold is used as a training data \tilde{y}_{-b} , while the rest one-fold is

used as a testing data \tilde{y}_b . The B-fold cross-validation of the log predictive density score (LPDS) is then formed as

$$B^{-1} \sum_{b=1}^B \ln p(\tilde{y}_b | \tilde{y}_{-b}, x).$$

In other words, part of the observations are used to update the flat (non-informative) prior and the sensitivity to the prior can be reduced substantially. According to Villani et al. (2009), the Bayes factor is roughly B times more discriminatory than the LPDS. For selecting models in Sect. 4, the LPDS was computed using B = 5 folds of the data.

4 Application: Educational and Residential Differences in Marriage Timing Among Eritrean Men and Women

We now illustrate the models and methods described in the previous sections with real-life data—entry into marriage among Eritrean men and women based on its 2010 Population and Health Survey (EPHS2010).

The main goals with the illustration are to study the distributional shapes of the times to marriage, model the effects of covariates on these event times, and examine if inferences regarding covariate effects are robust to the choice of distributional shape.

The study of marriage timing (age at marriage) is also of substantive interest in its own because of its strong negative association with women's health directly (Raj, 2010) or indirectly through its negative impact on health care utilization (Godha et al., 2016).

4.1 Data and Variables

The data used for illustration in this chapter come from the 2010 Eritrea Population and Health Survey, EPHS2010 (National-Statistics-Office-Eritrea and Faf0-AIS, 2013). The EPHS2010 was designed as a follow-up to its predecessors—the 1995 and 2002 Demographic and Health Surveys (National-Statistics-Office-Eritrea and Macro-International-Inc., 1997, 2003), and to update the information from the previous surveys as well as provide findings on some new topics of interest.

The EPHS2010 was conducted between January and July 2010 and gathered information from 30224 women aged 15–49 and 5021 men aged 15–59. For the purpose of this paper, only respondents with known values on marital status at the time of the survey are used in the analyses. This resulted in 10238 usable records for women and all 5021 records for men. Detailed tabulations for the entire survey

Table 1 Summary statistics of the data sets used in the illustration

Covariate	Levels	Women			Men			Combined Sample		
		n	Events	% Event	n	Event	% Event	n	Event	% Event
	No Educ	4186	3799	90.75	1050	892	84.95	5236	4691	89.59
	Primary	2055	1634	79.51	803	543	67.62	2858	2177	76.17
Education	Middle	1827	1006	55.06	1209	455	37.63	3036	1461	48.12
	Secon	1894	886	46.78	1516	461	30.41	3410	1347	39.50
	PostSec	276	96	34.78	442	218	49.32	718	314	43.73
	Capital	1819	969	53.27	931	344	36.95	2750	1313	47.75
Residence	Other Towns	2504	1739	69.45	1257	565	44.95	3761	2304	61.26
	Rural Areas	5915	4713	79.68	2833	1660	58.60	8748	6373	72.85
	Total	10,238	7421	72.49	5021	2569	51.17	15,259	9990	65.47

may be found in the EPHS2010 Final Report (National-Statistics-Office-Eritrea and Fafo-AIS, 2013). Summary statistics for the subset of data used in the present chapter are shown in Table 1.

By the survey time (January–July 2010), 7421 of the 10238 women (72 %) and 2569 of the 5021 men (51 %) have responded they were ever married (this includes those who might have been separated or widowed after). The rest, 2817 women and 2452 men (28 % and 49 %, respectively), have responded that they were still single at the time of interview. The distribution of the women across educational level shows that 4186 (41 %) had no education at all, 2055 (20 %) had primary-level education, 1827 (18 %) had middle-level education, 1894 (18 %) had secondary-level education, while the rest 276 (3 %) had post-secondary education. The corresponding figures for men are 1051 (21 %), 803 (16 %), 1209 (24 %), 1516 (30 %), and 442 (9 %), respectively. Further, 1819 (18 %) of the women respondents were from the capital (Asmara), 2504 (24 %) were from other towns, while the majority 5915 (58 %) were from rural areas. The corresponding figures for men are 931 (19 %), 1257 (25 %), and 2833 (56 %), respectively.

The columns of percentage married in Table 1 reveal clear differentials across both educational levels and residence for both women and men. For instance, while women with no education constitute 41 % of the entire sample, they constitute 51 % of the marriages (3799 of 7421). Women with post-secondary education, on the other hand, constitute only 1 % of the marriages (96 of 7421). The pattern is similar but less dramatic for men—those with no education constitute 35 % of the marriages, while those with post-secondary education constitute only 8 % of the marriages. Differentials across residence show that women from rural areas constitute 58 % of the sub-sample but 64 % of the marriages. Women from the capital, on the other hand, constitute 19 % of the sub-sample but only 13 % of the marriages. The contribution of men from the capital to the sub-sample is 18 %, while their contribution to the total marriage is 15 %. Men from rural areas constitute 56 % of the sub-sample but 65 % of the marriages.

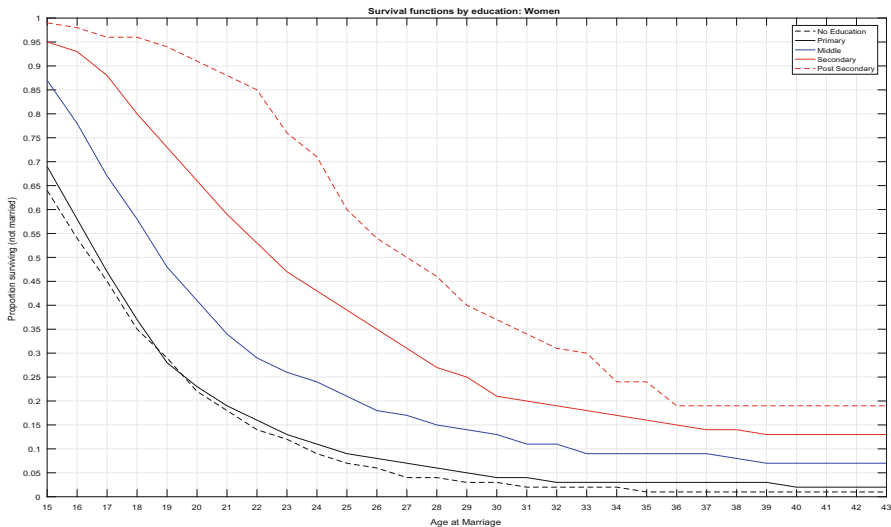


Fig. 3 Survival functions by education: Women

Plots of survival functions by education and residence for women and men are shown in Figs. 3, 4, 5, 6, and 7. Figures 3 and 4 show plots for women by education and residence, respectively, while Figs. 5 and 6 show the corresponding plots for men. Figure 7 shows gender differences in entry to first marriage among all men and women.

The plots depict what we already noted in Table 1—that there are differentials across education and residence and that the educational differences are more pronounced in the women data than in men data. The last figure shows that women enter marriage at faster rates than men.

The summary in Table 1 and Figs. 3, 4, 5, 6, and 7 provides a good description of the data at hand, but in order to make sound inferences based on the sub-sample, we need deeper analyses of the data and formal statistical tests. Ghilagaber (2018) has analyzed the data sets using frequentist statistical methods ranging from elementary measures of association between an event of interest and background variables to more complex and advanced methods that utilize the data more efficiently. Elsewhere in this book, (Munezero and Ghilagaber, 2022b) analyze the data sets using dynamic Bayesian approach where covariate effects are allowed to vary over time.

In the next sub-section, we present and discuss results from fitting the further EGG model of Sect. 2 to the above data sets in the Bayesian framework of Sect. 3.

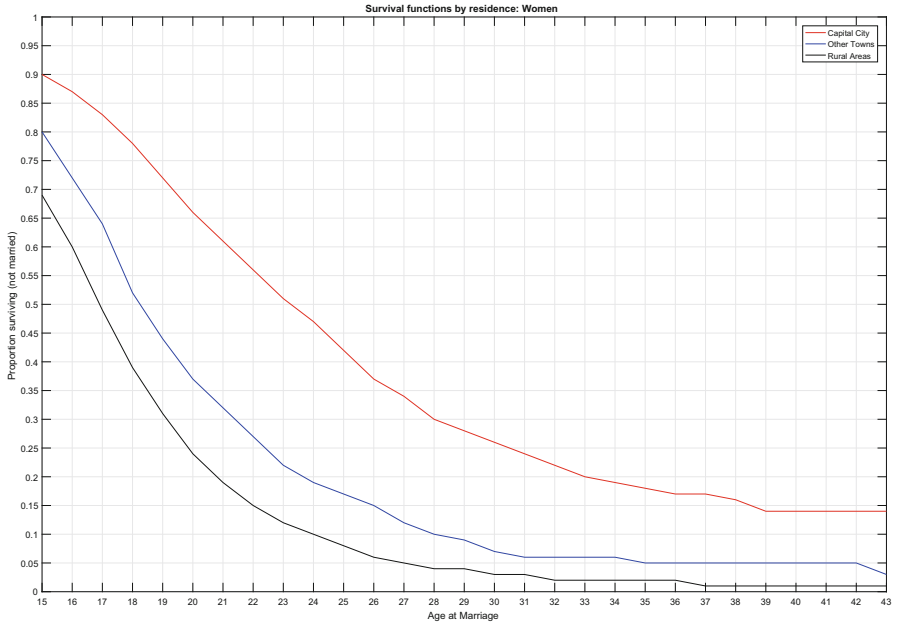


Fig. 4 Survival functions by residence: Women

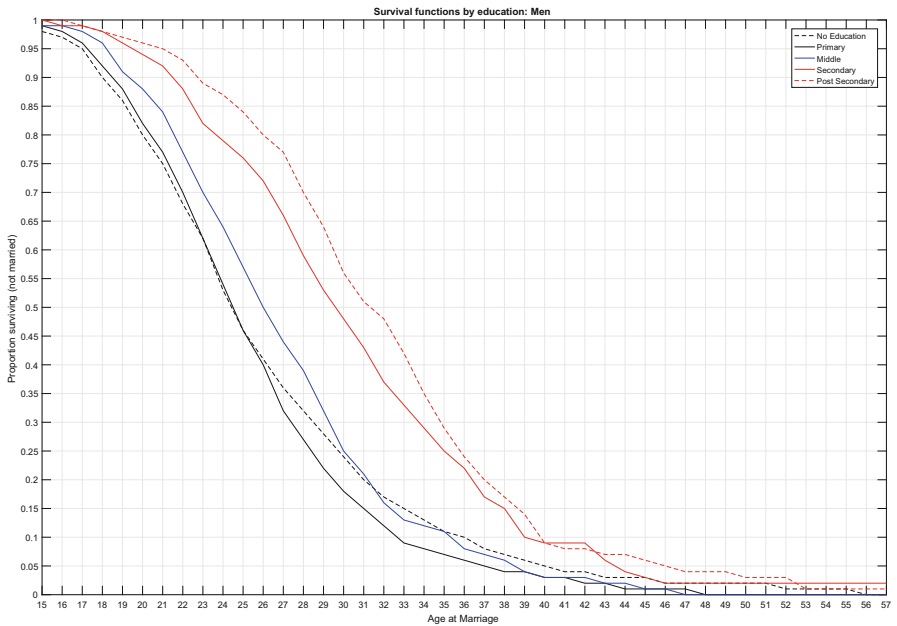


Fig. 5 Survival functions by education: Men

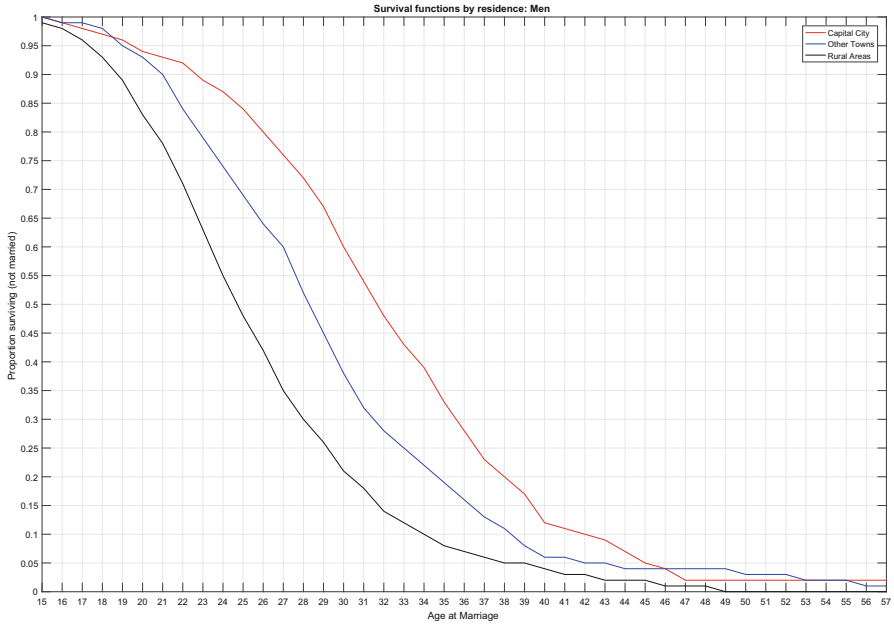


Fig. 6 Survival functions by residence: Men

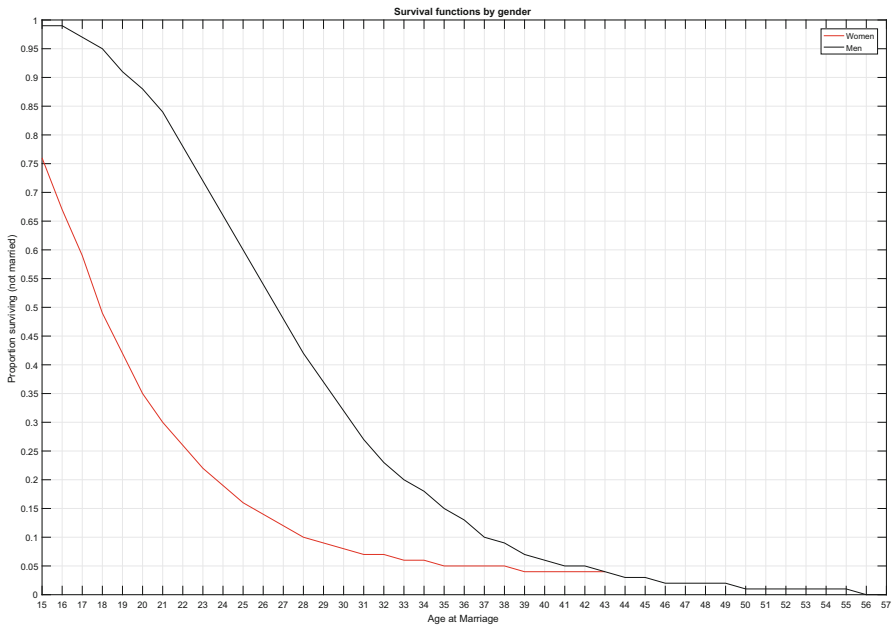


Fig. 7 Survival functions by gender

4.2 Results from Bayesian Analysis of the Data Using the EGG Model

Table 2 contains a summary of our results to which we will return at the end of this section. Results from fitting the extended generalized gamma (EGG) model and its 13 special cases to the data for women, men, and the combined sample are shown in Tables 3, 4, and 5, respectively.

In Table 3, the results from the unconstrained EGG model show that the scale and shape parameters (which are freely estimated from the data) are $\delta = 0.246$ and $q = -0.526$, respectively.

These estimates give early indications of the constants to which the scale and shape parameters are close as well as the relationship between them. For instance, the estimated shape parameter (-0.526) is much closer to -1 and 0 than it is to 1 . This, in turn, means the reciprocal Weibull distribution (which constrains the shape parameter to -1) and the log-normal distribution (which constrains the shape parameter to 0) are more plausible candidate distributions than the Weibull distribution (which constrains the shape parameter to 1).

With regard to the relationships between the scale and shape parameters, a model that constrains negative equality is $\delta = -q$ that seems to be more plausible compared to, for instance, a model that constrains reciprocal or negative reciprocal relationship. This is so because a reciprocal relationship would give a scale parameter of $1/(-0.526) = -1.90$, while a negative reciprocal relationship would yield $-(1/(-0.526)) = 1.90$ both of which are far from the freely estimated scale parameter 0.246 . This, in turn, excludes models such as ammag and inverse ammag in favor of the inverse gamma model.

The above closeness of the special case models to the more general EGG model is also reflected in the values of log predictive density scores (LPDS) given in the last columns of each model. For instance, the LPDS of the EGG model is -4584 , while that of the closest model (the inverse gamma) is -4594 . On the other hand, the LPDS for ammag and inverse ammag are -5705 and -5184 , respectively, which are far from that of the EGG.

Table 2 Posterior means (and 95 % hpd) of estimated effects in the selected models

Covariate	Levels	Women (EGG model)			Men (Inverse gamma)			Combined (EGG model)		
		Estimate	2.5 %	97.5 %	Estimate	2.5 %	97.5 %	Estimate	2.5 %	97.5 %
Education	No Educ	0 (ref.)	–	–	0 (ref.)	–	–	0 (ref.)	–	–
	Primary	0.015	0.002	0.028	-0.016	-0.04	0.009	0.042	0.029	0.055
	Middle	0.138	0.123	0.153	0.023	-0.001	0.047	0.186	0.007	0.2
Education	Secon	0.251	0.234	0.269	0.097	0.071	0.122	0.315	0.008	0.331
	PostSec	0.425	0.39	0.461	0.138	0.105	0.171	0.518	0.493	0.466
Residence	Capital	0 (ref.)	–	–	0 (ref.)	–	–	0 (ref.)	–	–
	Other	-0.086	-0.069	-0.052	-0.061	d	-0.035	-0.052	-0.068	-0.036
	Rural	-0.079	-0.094	-0.062	-0.137	-0.161	-0.111	-0.042	-0.058	-0.026

Table 3 Posterior means (and 95 % hpd) of covariate effects on log time to event: Women

Model	Educational Level (ref: No Educ)				Residence (ref: Capital)			Scale & Shape		GoF
	Primary	Middle	Seco	PostSec	Other Towns	Rural	Scale (δ)	Shape (q)	LPDS	
EGG	0.015	0.138	0.251	0.425	-0.069	-0.079	0.246	-0.526	-4584	
	(0.002, 0.028)	(0.123, 0.153)	(0.234, 0.269)	(0.39, 0.461)	(-0.086, -0.052)	(-0.094, -0.062)	(0.242, 0.25)	(-0.574, -0.476)		
Exponential	0.121	0.526	0.774	1.187	-0.126	-0.112	1	1	-6201	
	(0.062, 0.179)	(0.454, 0.599)	(0.691, 0.853)	(0.984, 1.398)	(-0.202, -0.040)	(-0.189, -0.031)	1	1		
Inv. Exp	0.066	0.321	0.488	0.755	-0.087	-0.079	1	-1	-5985	
	(0.013, 0.119)	(0.264, 0.38)	(0.425, 0.55)	(0.623, 0.895)	(-0.155, -0.021)	(-0.141, -0.015)				
Weibull	0.007	0.116	0.228	0.42	-0.146	-0.184	0.267	1	-4892	
	(-0.009, 0.022)	(0.096, 0.136)	(0.206, 0.25)	(0.364, 0.478)	(-0.168, -0.124)	(-0.206, -0.162)	(0.263, 0.271)	1		
Rec. Weibull	0.021	0.146	0.247	0.419	-0.054	-0.053	0.243	-1	-4614	
	(0.008, 0.032)	(0.132, 0.16)	(0.231, 0.263)	(0.382, 0.452)	(-0.07, -0.037)	(-0.069, -0.038)	(0.239, 0.247)			
Rayleigh	0.045	0.258	0.416	0.673	-0.129	-0.145	0.5	1	-5315	
	(0.016, 0.075)	(0.222, 0.295)	(0.374, 0.459)	(0.57, 0.778)	(-0.171, -0.088)	(-0.186, -0.106)				
Inv. Rayleigh	0.032	0.194	0.322	0.524	-0.071	-0.073	0.5	-1	-5101	
	(0.005, 0.06)	(0.165, 0.222)	(0.291, 0.353)	(0.453, 0.593)	(-0.102, -0.039)	(-0.105, -0.041)				
Stand. Gamma	0.008	0.125	0.242	0.414	-0.109	-0.134	q	0.252	-4665	
	(-0.007, 0.022)	(0.109, 0.142)	(0.222, 0.262)	(0.371, 0.458)	(-0.129, -0.09)	(-0.152, -0.114)		(0.248, 0.256)		
Gamma	0.081	0.371	0.555	0.848	-0.094	-0.087	$1, q \neq 0$	-0.388	-5939	
	(0.027, 0.135)	(0.311, 0.429)	(0.489, 0.622)	(0.705, 0.995)	(-0.162, -0.029)	(-0.155, -0.02)		(-0.444, -0.33)		
Inv. Gamma	0.014	0.135	0.25	0.423	-0.081	-0.097	$-q$	-0.251	-4594	
	(0, 0.027)	(0.119, 0.152)	(0.233, 0.267)	(0.383, 0.463)	(-0.099, -0.063)	(-0.114, -0.079)		(-0.255, -0.247)		
Ammag	0.049	0.268	0.448	0.808	-0.175	-0.206	$1/q$	2.361	-5705	
	(0.021, 0.076)	(0.225, 0.314)	(0.395, 0.502)	(0.66, 0.972)	(-0.221, -0.127)	(-0.25, -0.158)		(2.31, 2.416)		
Inv. Ammag	0.031	0.165	0.221	0.373	-0.045	-0.016	$-1/q$	-2.986	-5184	
	(0.013, 0.05)	(0.145, 0.185)	(0.199, 0.243)	(0.326, 0.415)	(-0.065, -0.023)	(-0.037, 0.005)		(-3.062, -2.921)		
Log normal	0.012	0.131	0.248	0.421	-0.094	-0.114	0.252	0	-4621	
	(-0.003, 0.026)	(0.114, 0.146)	(0.229, 0.266)	(0.381, 0.461)	(-0.112, -0.077)	(-0.133, -0.097)	(0.249, 0.257)			
Half-normal	0.087	0.415	0.63	1.006	-0.136	-0.139	$\sqrt{2}/2$	$\sqrt{2}$	-5898	
	(0.044, 0.13)	(0.358, 0.475)	(0.563, 0.694)	(0.836, 1.185)	(-0.199, -0.07)	(-0.2, -0.075)				

Table 4 Posterior means (and 95 % hpd) of covariate effects on log time to event: Men

Model	Educational Level (ref: No Educ)				Residence (ref: Capital)		Scale & Shape		GoF
	Primary	Middle	Seco	PostSec	Other Towns	Rural	Scale (δ)	Shape (q)	
EGG	-0.018	0.022	0.096	0.138	-0.062	-0.137	0.235	-0.199	-1754
	(-0.042, 0.005)	(-0.002, 0.049)	(0.069, 0.122)	(0.106, 0.171)	(-0.089, -0.035)	(-0.164, -0.11)	(0.228, 0.241)	(-0.289, -0.106)	
Exponential	0.123	0.604	0.861	0.558	-0.099	-0.135	1	1	-2441
	(0.014, 0.229)	(0.49, 0.722)	(0.726, 0.992)	(0.389, 0.72)	(-0.238, 0.038)	(-0.272, 0)			
Inv. Exp	0.074	0.307	0.44	0.344	-0.078	-0.126	1	-1	-2303
	(-0.025, 0.168)	(0.212, 0.396)	(0.345, 0.535)	(0.212, 0.475)	(-0.18, 0.019)	(-0.22, -0.022)			
Weibull	-0.062	-0.037	0.037	0.066	-0.055	-0.156	0.22	1	-1829
	(-0.085, -0.038)	(-0.061, -0.011)	(0.006, 0.067)	(0.027, 0.106)	(-0.086, -0.023)	(-0.191, -0.126)	(0.214, 0.225)		
Rec. Weibull	-0.001	0.037	0.108	0.153	-0.038	-0.11	0.241	-1	-1782
	(-0.024, 0.024)	(0.012, 0.06)	(0.083, 0.132)	(0.123, 0.184)	(-0.064, -0.014)	(-0.135, -0.086)	(0.235, 0.247)		
Rayleigh	0.01	0.193	0.341	0.265	-0.074	-0.142	0.5	1	-2083
	(-0.045, 0.06)	(0.133, 0.254)	(-0.14, -0.005)	(-0.209, -0.077)	(0.272, 0.405)	(0.18, 0.349)			
Inv. Rayleigh	0.02	0.115	0.206	0.215	-0.057	-0.117	0.5	-1	-1970
	(-0.029, 0.07)	(0.065, 0.162)	(0.157, 0.255)	(0.15, 0.277)	(-0.108, -0.004)	(-0.167, -0.069)			
Stand. Gamma	-0.031	0.008	0.081	0.12	-0.066	-0.15	q	0.228	-1763
	(-0.054, -0.008)	(-0.017, 0.031)	(0.054, 0.107)	(0.086, 0.153)	(-0.096, -0.041)	(-0.177, -0.123)		(0.223, 0.234)	
Gamma	0.084	0.344	0.492	0.376	-0.081	-0.128	1, $q \neq 0$	-0.615	-2297
	(-0.013, 0.177)	(0.252, 0.432)	(0.391, 0.594)	(0.241, 0.514)	(-0.184, 0.015)	(-0.227, -0.034)		(-0.707, -0.527)	
Inv. Gamma	-0.016	0.023	0.097	0.138	-0.061	-0.137	$-q$	-0.235	-1754
	(-0.04, 0.009)	(-0.001, 0.047)	(0.071, 0.122)	(0.105, 0.171)	(-0.087, -0.035)	(-0.161, -0.111)		(-0.241, -0.229)	
Ammag	-0.042	0.179	0.438	0.214	-0.026	-0.142	1/ q	2.86	-2253
	(-0.084, 0)	(0.107, 0.253)	(0.351, 0.524)	(0.115, 0.313)	(-0.122, 0.068)	(-0.237, -0.049)		(2.7, 3.02)	
Inv. Ammag	0.008	0.049	0.113	0.147	-0.002	-0.075	-1/ q	-2.923	-1998
	(-0.024, 0.039)	(0.017, 0.08)	(0.082, 0.141)	(0.104, 0.189)	(-0.035, 0.031)	(-0.106, -0.046)		(-3.033, -2.815)	
Log normal	-0.023	0.016	0.091	0.131	-0.065	-0.143	0.233	0	-1756
	(-0.049, 0.002)	(-0.009, 0.04)	(0.065, 0.117)	(0.097, 0.165)	(-0.093, -0.037)	(-0.171, -0.117)	(0.227, 0.24)		
Half-normal	0.064	0.435	0.673	0.436	-0.084	-0.139	$\sqrt{2}/2$	$\sqrt{2}$	-2335
	(-0.012, 0.146)	(0.343, 0.527)	(0.573, 0.778)	(0.309, 0.561)	(-0.192, 0.031)	(-0.253, -0.028)			

Table 5 Posterior means (and 95 % hpd) of covariate effects on log time to event: combined sample

Model	Educational Level (ref: No Educ)			Residence (ref: Capital)		Scale & Shape		GoF
	Primary	Middle	Secun	PostSec	Other Towns	Rural	Scale (δ)	
EGG	0.042	0.186	0.315	0.493	-0.052	-0.042	0.285	-0.525
	(0.029, 0.055)	(0.172, 0.2)	(0.299, 0.331)	(0.466, 0.518)	(-0.068, -0.036)	(-0.058, -0.026)	(0.281, 0.289)	(-0.575, -0.475)
Exponential	0.162	0.643	0.936	1.021	-0.078	-0.03	1	1
	(0.112, 0.21)	(0.582, 0.702)	(0.865, 1.008)	(0.905, 1.141)	(-0.146, -0.009)	(-0.097, 0.04)		
Inv. Exp	0.102	0.403	0.591	0.742	-0.056	-0.024	1	-1
	(0.057, 0.151)	(0.356, 0.449)	(0.544, 0.642)	(0.658, 0.828)	(-0.111, -0.005)	(-0.076, 0.027)		
Weibull	0.02	0.147	0.278	0.412	-0.077	-0.102	0.294	1
	(0.005, 0.035)	(0.128, 0.165)	(0.257, 0.298)	(0.371, 0.451)	(-0.098, -0.057)	(-0.122, -0.082)	(0.290, 0.298)	
Rec. Weibull	0.042	0.189	0.308	0.491	-0.039	-0.025	0.279	-1
	(0.029, 0.055)	(0.175, 0.202)	(0.293, 0.322)	(0.468, 0.518)	(-0.055, -0.025)	(-0.039, -0.01)	(0.275, 0.283)	
Rayleigh	0.068	0.305	0.488	0.611	-0.078	-0.073	0.5	1
	(0.041, 0.094)	(0.275, 0.338)	(0.454, 0.524)	(0.55, 0.673)	(-0.113, -0.041)	(-0.108, -0.038)		
Inv. Rayleigh	0.058	0.246	0.389	0.563	-0.047	-0.03	0.5	-1
	(0.036, 0.082)	(0.22, 0.27)	(0.362, 0.413)	(0.519, 0.605)	(-0.075, -0.021)	(-0.056, -0.004)		
Stand. Gamma	0.035	0.173	0.307	0.464	-0.07	-0.075	q	0.288
	(0.021, 0.049)	(0.157, 0.189)	(0.29, 0.324)	(0.433, 0.495)	(-0.09, -0.053)	(-0.094, -0.057)		(0.285, 0.292)
Gamma	0.119	0.452	0.66	0.8	-0.062	-0.027	$1, q \neq 0$	-0.454
	(0.075, 0.164)	(0.405, 0.497)	(0.608, 0.715)	(0.712, 0.891)	(-0.113, -0.008)	(-0.077, 0.025)		(-0.499, -0.407)
Inv. Gamma	0.041	0.184	0.318	0.491	-0.057	-0.05	- q	-0.289
	(0.027, 0.055)	(0.17, 0.199)	(0.301, 0.333)	(0.464, 0.518)	(-0.074, -0.04)	(-0.067, -0.034)		(-0.293, -0.285)
Ammag	0.035	0.288	0.517	0.595	-0.07	-0.108	$1/q$	2.304
	(0.01, 0.062)	(0.251, 0.326)	(0.472, 0.563)	(0.515, 0.682)	(-0.113, -0.027)	(-0.15, -0.069)		(2.256, 2.353)
Inv. Ammag	0.046	0.2	0.286	0.46	-0.034	0.003	- $1/q$	-2.738
	(0.029, 0.063)	(0.182, 0.218)	(0.268, 0.303)	(0.428, 0.49)	(-0.052, -0.016)	(-0.015, 0.021)		(-2.785, -2.688)
Log normal	0.039	0.181	0.316	0.48	-0.065	-0.063	0.291	0
	(0.025, 0.054)	(0.166, 0.197)	(0.298, 0.333)	(0.451, 0.508)	(-0.083, -0.048)	(-0.079, -0.046)	(0.287, 0.295)	
Half-normal	0.114	0.495	0.754	0.841	-0.08	-0.057	$\sqrt{2}/2$	$\sqrt{2}$
	(0.079, 0.153)	(0.446, 0.542)	(0.7, 0.808)	(0.743, 0.936)	(-0.138, -0.027)	(-0.118, -0.006)		

Another point worth noting is that the estimates of the covariate effects and their associated 95% hpd are much alike in the models that are close to each other (in terms of estimated scale and/or shape parameters or in terms of LPDS) than those estimates that are far apart.

Thus, for the women data, it would not make much difference if we base our conclusions on the estimates from the EGG model or the inverse gamma model though a formal test would favor the larger EGG model.

The results for men shown in Table 4 can be interpreted similarly. Here, the scale and shape parameters estimated freely from the data in the EGG model are $\delta = 0.235$ and $q = -0.199$, respectively. Again, the inverse gamma model that imposes a negative relationship between the scale and shape parameters ($\delta = -q$) seems to be much more plausible than any other model. In fact, a closer look at the LPDS values shows that it even outperforms the larger EGG model though the difference in LPDS is marginal.

Hence, for the men data, we have a very strong evidence to base our conclusions on the results from the inverse gamma model that, of course, are identical to those from the EGG model.

Last, the results for the combined sample are shown in Table 5. Similar reasoning as in the above leads to the choice of EGG model or the inverse gamma model though a formal test would favor the larger EGG model. That the results for the combined sample reflected those for women are not surprising because women constitute about two-third of the combined sample.

The final estimates of covariate effects and their associated 95% hpd from our chosen models for respective data sets are summarized in Table 2.

The results in Table 2 show that there are significant differentials in entry to first marriage across women's educational level and residence where lower education and rural residence are associated with higher intensities of marriage. For men, the educational differences are less pronounced as there is no significant difference in the intensities of entry to first marriage between those with no education and those with primary or middle education. The residential differential is, however, still significant. The results for the combined sample follow those of women because, as mentioned before, women constitute majority in the combined sample.

5 Summary and Concluding Remarks

In this chapter, we presented the extended generalized gamma (EGG) model for survival data with censored observations. Previous works have shown that five known models can be treated as special cases of the EGG model by constraining the scale parameter, shape parameter, or both to some constants. In the present chapter, we extended the EGG model further to include 13 special case models. This was achieved by imposing relationships between the scale and shape parameters in addition to constraining them to some constants.

The issues were illustrated with data on entry into first marriage among Eritrean men and women based on data from the 2010 Eritrean Population and Health Survey

(EPHS 2010). Inference was fully Bayesian using a random walk Metropolis–Hastings algorithm to sample from the posterior distribution, and we compared the models with each other and relative to the more general EGG model using the log predictive density score (LPDS).

The application demonstrates that the further extended family of distributions provides a wide range of alternatives for a baseline distribution in the analysis of survival data with censored observations. For instance, we found that the inverse gamma model, where we impose the scale parameter to be the negative of the shape parameters ($\delta = -q$), fits the men data best and outperforms the EGG model. It also performs well in the women data and the combined sample though the evidence is not as strong as in the men data. This was in accordance with the freely estimated values of the scale and shape parameters in the EGG model.

The empirical results in the final selected models reveal significant differentials in the pace of entry to first marriage across women’s educational levels and residence. As would be expected, lower education and rural residence is associated with higher intensities of marriage. Educational differentials are, however, less pronounced for men as there was no significant difference in the intensities of entry to first marriage between those with no education (the baseline group) and those with primary or middle education. The residential differential was still significant in the men’s data. When we analyzed the combined data, the results followed those of women due, mainly, to the fact that women constitute about two-third in the combined sample.

It may be worth noting that the educational level of the individuals refers to what is achieved by the survey time. As such, it is anticipatory in the sense that the reported educational level might have been achieved after the event of interest. But, our aim here is to demonstrate the models and methods empirically, and the anticipatory nature of education does not affect our purpose. Ghilagaber and Koskinen (2009), Ghilagaber and Larsson (2019), and Munezero and Ghilagaber (2022a) study potential biases due to the use of anticipatory covariates and how to account for that.

Our analysis was based on the tacit assumption that the survivor function $S(t)$ tends to 0 as the study period gets longer. This, in turn, means that we have assumed all individuals will experience the event of interest sooner or later. This may not be true for the event in our illustrative example (marriage) as there may be some individuals who may never marry for various reasons. Future works may, therefore, consider accounting for such long-term survivors (those who may never experience the event of interest). This can be achieved by using, for instance, a mixture model consisting of a hazard/intensity model for those who experienced the event or may experience it in the future and a logistic model for the probability of being long-term survivor (never experiencing the event).

Appendix 1: Density Functions, $f(t)$, and Survival Functions, $S(t)$, of Special Cases in the Extended Generalized Gamma Model

Distribution	$f(t)$	$S(t)$
Exponential	$e^{-\mu} e^{-e^{-\mu} t}$	$e^{-e^{-\mu} t}$
I. Exponential	$e^{\mu} e^{-\frac{e^{\mu}}{t}} / t^2$	$1 - e^{-\frac{e^{\mu}}{t}}$
Weibull	$\frac{1/\delta}{e^{\mu}} \left(\frac{t}{e^{\mu}}\right)^{1/\delta-1} e^{-\left(\frac{t}{e^{\mu}}\right)^{1/\delta}}$	$e^{-t^{1/\delta} e^{-\mu/\delta}}$
R. Weibull	$\frac{1/\delta}{e^{\mu}} \left(\frac{t}{e^{\mu}}\right)^{-1/\delta-1} e^{-\left(\frac{t}{e^{\mu}}\right)^{-1/\delta}$	$1 - e^{-t^{-1/\delta} e^{\mu/\delta}}$
Rayleigh	$\frac{2t}{e^{2\mu}} e^{-\frac{t^2}{2\mu}}$	$e^{-\frac{t^2}{2\mu}}$
I. Rayleigh	$\frac{2e^{2\mu}}{t^3} e^{-\frac{e^{2\mu}}{t^2}}$	$1 - e^{-\frac{e^{2\mu}}{t^2}}$
S. Gamma	$1/\Gamma(q^{-2}) (q^{-2} e^{-\mu})^{q^{-2}-1} e^{-q^{-2} e^{-\mu} t}$ [a]	$1 - \gamma(q^{-2}, tq^{-2} e^{-\mu}) / \Gamma(q^{-2})$ [b]
Gamma	$(q > 0) qt^{q-1} \frac{1}{\Gamma(q^{-2})} (q^{-2} e^{-q\mu})^{q^{-2}-1} e^{-q^{-2} e^{-q\mu} t}$	$1 - \gamma(q^{-2}, tq^{-2} e^{-q\mu}) / \Gamma(q^{-2})$
	$(q < 0) -qt^{q-1} \frac{1}{\Gamma(q^{-2})} (q^{-2} e^{-q\mu})^{q^{-2}-1} e^{-q^{-2} e^{-q\mu} t}$	$\gamma(q^{-2}, tq^{-2} e^{-q\mu}) / \Gamma(q^{-2})$
I. Gamma	$t^{-2} \frac{1}{\Gamma(q^{-2})} (q^{-2} e^{\mu})^{q^{-2}-1} e^{-q^{-2} e^{\mu} t^{-1}}$	$\gamma(q^{-2}, t^{-1} q^{-2} e^{\mu}) / \Gamma(q^{-2})$
Ammag	$q^2 t^{q^2-1} \frac{1}{\Gamma(q^{-2})} (q^{-2} e^{-q^2 \mu})^{q^{-2}-1} e^{-q^{-2} e^{-q^2 \mu} t}$	$1 - \gamma\left(q^{-2}, \frac{t^{q^2}}{q^2} e^{-q^2 \mu}\right) / \Gamma(q^{-2})$
I. Ammag	$q^2 t^{-q^2-1} \frac{1}{\Gamma(q^{-2})} (q^{-2} e^{q^2 \mu})^{q^{-2}-1} e^{-q^{-2} e^{q^2 \mu} t^{-q^2}}$	$\gamma\left(q^{-2}, \frac{t^{-q^2}}{q^2} e^{q^2 \mu}\right) / \Gamma(q^{-2})$
Log normal	$\frac{1}{t\delta\sqrt{2\pi}} e^{-\frac{(\log t - \mu)^2}{2\delta^2}}$	$\frac{1}{2} (1 + \operatorname{erf}(\frac{x - \mu}{\sqrt{2}\delta}))$
Half-normal	$\frac{1}{e^{\mu}} \sqrt{\frac{2}{\pi}} e^{-\frac{x^2}{2e^{2\mu}}}$	$1 - \operatorname{erf}\left(\frac{x}{\sqrt{2}e^{\mu}}\right)$ [c]

^a $\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$

^b $\gamma(s, r) = \int_0^r x^{s-1} e^{-x} dx$

^c $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-z^2} dz$

Appendix 2: Proof of Lemma 1

Proof:

$$E(T^r) = \int_0^\infty t^r \frac{|q|}{\delta} t^{\frac{q}{\delta}-1} \frac{[q^{-2}(e^{-\mu})^{\frac{q}{\delta}}]^{q^{-2}}}{\Gamma(q^{-2})} (t^{\frac{q}{\delta}})^{q^{-2}-1} e^{-q^{-2}(e^{-\mu})^{\frac{q}{\delta}} t^{\frac{q}{\delta}}} dt.$$

Let $x = q^{-2}(e^{-\mu})^{\frac{q}{\delta}} t^{\frac{q}{\delta}}$; then $t = (q^2 e^{\frac{q\mu}{\delta}} x)^{\frac{\delta}{q}}$,

$$\begin{aligned} E(T^r) &= \frac{[(q^2 e^{\frac{q\mu}{\delta}})^{\frac{\delta}{q}}]^r}{\Gamma(q^{-2})} \int_0^\infty x^{\frac{r\delta}{q}+q^{-2}-1} e^{-x} dx \\ &= \frac{[(q^2 e^{\frac{q\mu}{\delta}})^{\frac{\delta}{q}}]^r}{\Gamma(q^{-2})} \Gamma(q^{-2} + \frac{r\delta}{q}) \quad (q^{-2} + \frac{r\delta}{q} > 0) \end{aligned}$$

When $r = 1$, we get $E(T)$, and when $r = 2$, we get $E(T^2)$. Using these, we have $V(T) = E(T^2) - E^2(T)$. ■

Appendix 3: R Program Codes for Bayesian Inference

```
#####
#####Random Walk Metropolis algorithm#####
#####
library(MASS)
library(mvtnorm)

#Calculate log-likelihood and the Hessian evaluated at the mode
calculate.loglike <- function(beta) {

  beta1 <- beta[1]+beta[2]*Prim+beta[3]*Middle+beta[4]*Secon+
    beta[5]*PostSec+beta[6]*Other+beta[7]*Rural
  delta <- beta[8]
  q <- beta[9]

  error <- (log(y)-beta1)/delta
  s <- q^(-2)
  r <- exp(q*error)
  #Probability density function
  pdf <- function(r,s,q) {
    if (q!=0) pdf <- abs(q)*r*dgamma(r,s,s)
    else pdf <- dnorm(error,0,1)
  }
  #Survival function
```

```

surv <- function(r,s,q){
  if (q!=0) surv <- (q>0)*(1-pgamma(r,s,s))+
                    (q<0)*(pgamma(r,s,s))
  else surv <- 1-pnorm(error,0,1)
surv
}
#log-likelihood function
loglike <- function(r,s,q){
  loglike <- sum(d*(log(pdf(r,s,q))-log(delta))+
                (1-d)*log(surv(r,s,q)))
  loglike
}

list(loglike=loglike(r,s,q))
}

m=20000
beta0 <- c(5,0,0,0,0,0,0)
delta0 <- 1
q0 <- 0.12

theta0 <- c(beta0,delta0,q0)

mu.beta <- c(0,0,0,0,0,0,0) #prior
s.beta <- c(1000,1000,1000,1000,1000,1000,1000) #prior

theta <- matrix(nrow=m,ncol=9)
acc.prob_theta <- 0
current.theta <- theta0

for (t in 1:m){
  cur <- calculate.loglike(current.theta)

  prop.theta <- mvrnorm(1,current.theta,Sigma=vcov_mode)
  jump_1 <- -as.numeric(0.5*log(det(vcov_mode)))-
            0.5*t(prop.theta-current.theta)%%solve(vcov_mode)
            %%(prop.theta-current.theta)
  prop <- calculate.loglike(prop.theta)

  ll_1 <- prop$loglike
  jump_2 <- -as.numeric(0.5*log(det(vcov_mode)))-
            0.5*t(current.theta-prop.theta)%%solve(vcov_mode)
            %%(current.theta-prop.theta)
  ll_2 <- cur$loglike
  loga <- ll_1-ll_2+sum(dnorm(prop.theta[1:7],mu.beta,s.beta,
                             log=T))-
            sum(dnorm(current.theta[1:7],mu.beta,s.beta,log=T))+
            dgamma(prop.theta[8],1,1,log=T)-dgamma(current.theta[8],
            1,1,log=T)+
            dnorm(prop.theta[9],0,1000,log=T)-dnorm(current.theta[9],
            0,1000,log=T)+
            jump_2-jump_1

  u <- runif(1)

```

```

u <- log(u)
  if (u < loga){
    current.theta <- prop.theta
    acc.prob_theta <- acc.prob_theta+1
  }
theta[t,] <- current.theta
}

#####
#####Log predictivedensity score (LPDS)#####
#####

#5-fold cross validation
cv <- matrix(cbind(Age,1,Prim,Middle,Secon,PostSec,Other,Rural,
  Marr),ncol=9, nrow=length(Age))

index <- sample(length(y),size=length(y),replace=F)
index_1 <- index[1:2047]
index_2 <- index[2048:4095]
index_3 <- index[4096:6143]
index_4 <- index[6144:8191]
index_5 <- index[8192:10238]

new <- list(as.matrix(cv[index_1,]),as.matrix(cv[index_2,]),as.
  matrix(cv[index_3,]),
  as.matrix(cv[index_4,]),as.matrix(cv[index_5,]))

calculate.loglike <- function(b,delta,q,data){

  x<- data[,2:8]
  y <- data[,1]
  d <- data[,9]
  mu <- x%*%b

  #Probability density function
  pdf <- function(b,delta,q){
    if (q!=0) pdf <- abs(q)/delta*y^(q/delta-1)*
      dgamma(y^(q/delta),q^(-2),q^(-2)*exp
        (-mu)^(q/delta))
    else pdf <- dnorm(log(y),mu,delta^2)
  }
  pdf
}
#Survival function
surv <- function(b,delta,q){
  if (q!=0) surv <- (q>0)*(1-pgamma(y^(q/delta),q^(-2),q^(-2)
    *exp(-mu)^(q/delta)))+
    (q<0)*pgamma(y^(q/delta),q^(-2),q^(-2)
    *exp(-mu)^(q/delta))
  else surv <- 1-pnorm(log(y),mu,delta^2)
}
surv
}
loglike <- sum(d*(log(pdf(b,delta,q)))+(1-d)*log(surv(b,delta,q)))
return(list(loglike=loglike))
}

```

```

m<-20000

#initial values
beta0 <- mode[1:7]
delta0 <- mode[8]
q0 <- -0.15
theta0 <- c(beta0,delta0)

mu.beta <- c(0,0,0,0,0,0,0,0) #prior
s.beta <- c(1000,1000,1000,1000,1000,1000,1000,1000) #prior

theta <- matrix(nrow=m,ncol=8)
q <- c()
pre <- c()
logpre <- c()
loglike <- c()
Cb <- c()
acc.prob_theta <- 0
acc.prob_q <- 0
current.theta <- theta0
current.q <- q0

  for (k in 1:5){
    for (t in 1:m){
      prop.theta <- mvrnorm(1,current.theta,cov[1:8,1:8])
      ll_1 <- calculate.loglike(prop.theta[1:7],prop.theta[8],
        current.q,data=rbind(new[-k][[1]],new[-k][[2]],
          new[-k][[3]],new[-k][[4]]))$loglike
      ll_2 <- calculate.loglike(current.theta[1:7],current.theta[8],
        current.q,data=rbind(new[-k][[1]],new[-k][[2]],new[-k][
          [3]],new[-k][[4]]))$loglike
      jump_1 <- -as.numeric(0.5*log(det(cov[1:8,1:8]))) -
        0.5*t(prop.theta-current.theta)%*%solve(cov[1:8,1:8])%*(prop.
        theta-current.theta)
      jump_2 <- -as.numeric(0.5*log(det(cov[1:8,1:8]))) -
        0.5*t(current.theta-prop.theta)%*%solve(cov[1:8,1:8])
        %*(current.theta-prop.theta)

      loga <- ll_1-ll_2+sum(dnorm(prop.theta[1:7],mu.beta,s.beta,
        log=T))-sum(dnorm(current.theta[1:7],mu.beta,s.beta,
        log=T))+dgamma(prop.theta[8],1,1,log=T)-dgamma(current.
        theta[8],1,1,log=T)+jump_2-jump_1

      u <- runif(1)
      u <- log(u)
      if (u < loga){
        current.theta <- prop.theta
        acc.prob_theta <- acc.prob_theta+1
      }
      theta[t,] <- current.theta

      prop.q <- rnorm(1,current.q,cov[9,9])
      ll_1 <- calculate.loglike(current.theta[1:7],current.theta[8],

```

```

      prop.q, data=rbind(new[-k][[1]], new[-k][[2]], new[-k][[3]],
      new[-k][[4]])$loglike
ll_2 <- calculate.loglike(current.theta[1:7], current.theta[8],
      current.q, data=rbind(new[-k][[1]], new[-k][[2]], new[-k]
      [[3]], new[-k][[4]]))$loglike
loga <- ll_1-ll_2+dnorm(prop.q, 0, 1000, log=T) -dnorm(current.q, 0,
      1000, log=T)
u <- runif(1)
u <- log(u)
if (u < loga){
  current.q <- prop.q
  acc.prob_q <- acc.prob_q+1
}
q[t] <- current.q
loglike[t] <- calculate.loglike(theta[t, 1:7], theta[t, 8], q[t],
      data=new[[k]])$loglike
}
Cb[k] <- mean(loglike)
pre <- exp(loglike-Cb[k])
burnin <- 5000
logpre[k] <- Cb[k]+log(mean(pre[(burnin+1):m]))
}
lpds <- mean(logpre)

```

Acknowledgments Yuli Liang's research is partly supported by the strategic funding of Örebro University. Liang carried out part of the work while affiliated with Stockholm University. The data sets analyzed in this chapter were provided to the co-author (GG) by the National Statistics Office, Eritrea. He is grateful to its director Mr. Ainom Berhane and its senior statisticians Mr. Hagos Ahmed and Mr. Samuel Tesfamariam. The views expressed in the chapter are solely of the authors and do not express the views or opinions of the data source or its employees.

References

- Abramowitz, M., & Stegun, I. (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. New York: Dover publications.
- Addison, J., & Portugal, P. (1987). On the distributional shape of unemployment duration. *The Review of Economics and Statistics*, 69(3), 520–526.
- Berger, J. O., Bernardo, J. M., Sun, D. et al. (2015). Overall objective priors. *Bayesian Analysis*, 10, 189–221.
- Bergström, R., & Edin, P. (1992). Time aggregation and the distributional shape of unemployment duration. *Journal of Applied Econometrics*, 7(1), 5–30.
- Bergström, R., Engwall, L., & Wallerstedt, E. (1994). Organizational foundations and closures in a regulated environment: Swedish commercial banks 1831–1990. *Scandinavian Journal of Management*, 10(1), 29–48.
- Bergström, R., Engwall, L., & Wallerstedt, E. (1997). The importance of flexible hazard functions in the analysis of organizational survival data—experiences from a cohort of Swedish commercial banks. *Quality and Quantity*, 31(1), 15–35.
- Butler, R., & McDonald, J. (1986). Trends in unemployment duration data. *The Review of Economics and Statistics*, 68(4), 545–557.

- Cox, C., Chu, H., Schneider, M., & Muñoz, A. (2007). Parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution. *Statistics in Medicine*, 26(23), 4352–4374.
- Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2), 187–202.
- Farewell, V., & Prentice, R. (1977). A study of distributional shape in life testing. *Technometrics*, 19(1), 69–75.
- Gelman, A., Carlin, J., Stern, H., & Rubin, D. (2004). *Bayesian data analysis*. Boca Raton, FL: Chapman & Hall/CRC.
- Ghilagaber, G. (2005). The extended generalized gamma model and its special cases: Applications to modeling marriage durations. *Quality and Quantity*, 39(1), 71–85.
- Ghilagaber, G. (2018). Statistical analysis of educational and urban-rural differences in marriage-timing: Evidence from the 2010 Eritrean population and health survey for women and men. In *Proceedings of the International Conference on Eritrean Studies, 20–22 July 2016* (pp. 1097–1124).
- Ghilagaber, G., Elisa, W., & Gyimah, S. O. (2014). A family of flexible parametric duration functions and their applications to modeling child-spacing in sub-Saharan Africa. In *Advanced Techniques for Modelling Maternal and Child Health in Africa. Springer Series on Demographic Methods and Population Analysis* (Vol. 34, pp. 185–209).
- Ghilagaber, G., & Koskinen, J. H. (2009). Bayesian adjustment of anticipatory covariates in analyzing retrospective data. *Mathematical Population Studies*, 16(2), 105–130.
- Ghilagaber, G., & Larsson, R. (2019). Maximum likelihood adjustment of anticipatory covariates in the analysis of retrospective data. Research report 2019–1. Stockholm: Department of Statistics, Stockholm University.
- Godha, D., Gage, A. J., Hotchkiss, D. R., & Cappa, C. (2016). Predicting maternal health care use by age at marriage in multiple countries. *Journal of Adolescent Health*, 58(5), 504–511.
- Heleno, A., & Alberto, J. (1986). The log-linear model with a generalized gamma distribution for the error: a Bayesian approach. *Statistics and Probability Letters*, 4(6), 325–332.
- Li, F., Villani, M., & Kohn, R. (2010). Flexible modeling of conditional distributions using smooth mixtures of asymmetric student t densities. *Journal of Statistical Planning and Inference*, 140, 3638–3654.
- Munezero, P., & Ghilagaber, G. (2022a). Dynamic Bayesian adjustment of anticipatory covariates in retrospective data: application to the effect of education on divorce risk. *Journal of Applied Statistics*, 49(6), 1382–1401.
- Munezero, P., & Ghilagaber, G. (2022b). Dynamic Bayesian modelling of educational and residential differences in family initiation among Eritrean men and women. In *Modern Biostatistical Methods for Evidence-Based Global Health Research*. Springer Series on Emerging Topics in Statistics and Biostatistics.
- National-Statistics-Office-Eritrea & Fafo-AIS (2013). *Eritrea Population and Health Survey 2010*. New York: National Statistics Office and Fafo Institute for Applied International Studies.
- National-Statistics-Office-Eritrea & Macro-International-Inc. (1997). *Eritrea Demographic and Health Survey 1995*. New York: National Statistics Office and Macro International.
- National-Statistics-Office-Eritrea & Macro-International-Inc. (2003). *Eritrea Demographic and Health Survey 2002*. New York: National Statistics and Evaluation Office and ORC Macro.
- Peng, Y., Dear, K., & Denham, J. (1998). A generalized F mixture model for cure rate estimation. *Statistics in Medicine*, 17(8), 813–830.
- Prentice, R. (1974). A log gamma model and its maximum likelihood estimation. *Biometrika*, 61(3), 539.
- Raj, A. (2010). When the mother is a child: the impact of child marriage on the health and human rights of girls. *Archives of Disease in Childhood*, 95(11), 931–935.
- Ramos, P. L., Achcar, J. A., Moala, F. A., Ramos, E., & Louzada, F. (2017). Bayesian analysis of the generalized gamma distribution using non-informative priors. *Statistics*, 51, 824–843.
- Stacy, E. (1962). A generalization of the gamma distribution. *The Annals of Mathematical Statistics*, 33(3), 1187–1192.

- Tsionas, E. (2001). Exact inference in four-parameter generalized gamma distributions. *Communications in Statistics-Theory and Methods*, 30(4), 747–756.
- Van Noortwijk, J. (2001). Bayes estimates of flood quantiles using the generalised gamma distribution. In *System and Bayesian Reliability; Essays in Honor of Professor Richard E. Barlow on his 70th Birthday* (pp. 351–374).
- Villani, M., Kohn, R., & Giordani, P. (2009). Regression density estimation using smooth adaptive Gaussian mixtures. *Journal of Econometrics*, 153, 155–173.
- Wei, L. J. (1992). The accelerated failure time model: a useful alternative to the Cox regression model in survival analysis. *Statistics in Medicine*, 11(14–15), 1871–1879.