DIGITALCOMMONS
—@WAYNESTATE—

Journal of Modern Applied Statistical Methods

11-2014

Objective Priors for Estimation of Extended Exponential Geometric Distribution

Pedro L. Ramos
Universidade de São Paulo, São Paulo, Brazil, pedrolramos@hotmail.com

Fernando A. Moala
Universidade Estadual Paulista, São Paulo, Brazil, femoala@fct.unesp.br

Jorge A. Achcar
Universidade de São Paulo, São Paulo, Brazil, achcar@fmrp.usp.br

Follow this and additional works at: http://digitalcommons.wayne.edu/jmasm

Part of the Applied Statistics Commons, Social and Behavioral Sciences Commons, and the Statistical Theory Commons

Objective Priors for Estimation of Extended Exponential Geometric Distribution

Pedro L. Ramos
Universidade de São Paulo
São Paulo, Brazil

Fernando A. Moala
Universidade Estadual Paulista
São Paulo, Brazil

Jorge A. Achcar
Universidade de São Paulo
São Paulo, Brazil

A Bayesian analysis was developed with different noninformative prior distributions such as Jeffreys, Maximal Data Information, and Reference. The aim was to investigate the effects of each prior distribution on the posterior estimates of the parameters of the extended exponential geometric distribution, based on simulated data and a real application.

Keywords: Extended exponential geometric distribution, Jeffreys, MDIP, Reference, Bayesian, noninformative, prior

Introduction

Adamidis & Loukas (2005) introduced an extension of the exponential geometric distribution (Adamidis & Loukas, 1998), naming it as an extended exponential geometric (EEG) distribution, to analyze lifetime data. This distribution provides increasing or decreasing hazard functions, depending on the values of its parameters. In this way, EEG gives a great flexibility of fit for the data.

If *T* is a random variable denoting the lifetime of a component with an extended exponential geometric (EEG) distribution, then the probability density is given by:

$$f\left(t \mid \gamma, \lambda\right) = \frac{\lambda \gamma e^{-\lambda t}}{\left(1 - \left(1 - \gamma\right) e^{-\lambda t}\right)^2}, \tag{1}$$

with $t > 0$ and parameters $\gamma > 0$ and $\lambda > 0$. Let us denote this distribution as EEG(γ, λ).

Dr. Ramos is in the Faculty of Medicine of Ribeirão Preto. Email him at: pedrolramos@hotmail.com. Dr. Moala is a Professor in the Faculty of Science and Technology. Email him at: femoala@fct.unesp.br. Dr. Achcar is a Professor in the Institute of Mathematics and Computer Science. Email him at: achcar@fmrp.usp.br.

The survival and hazard functions of EEG(γ, λ) distribution, for a fixed time t, is given by

$$S(t;\gamma,\lambda) = \frac{\gamma e^{-\lambda t}}{1-(1-\gamma)e^{-\lambda t}} \text{ and } h(t;\gamma,\lambda) = \frac{\lambda}{1-(1-\gamma)e^{-\lambda t}}, \tag{2}$$

respectively.

The mean and variance of the EEG distribution are given, respectively, by

$$E(T) = \frac{\gamma}{\lambda}\Psi(1-\gamma,1,1) \text{ and}$$

$$\text{var}(T) = \frac{\gamma}{\lambda^2}\left(2\Psi(1-\gamma,2,1)-\gamma\,\Psi^2(1-\gamma,1,1)\right) \tag{3}$$

where $\Psi(z, s, a)$ is known as Lerch transcendental function (Erdelyi et al., 1953), given by

$$\Psi(z,s,a) = \frac{1}{\Gamma(s)}\int_0^\infty \frac{u^{s-1}e^{-au}}{1-ze^{-u}}\,du = \sum_{k=0}^\infty \frac{z^k}{(a+k)^s}; z<1; a,s>0.$$

Adamidis et al. (2005) and Kitidamrongsuk (2010) gave additional properties of the EEG distribution.

Figures 1 and 2 present different forms for the density, survival and hazard functions for the EEG distribution considering different values of γ and λ.

The motivation here is to present a Bayesian analysis when there is little prior knowledge available or that reflects mainly the information from the sample.. In this situation, it is important to use noninformative priors, however, it can be difficult to choose a prior distribution that represent one of this situations. Thus, the main aim of this paper is to choose a noninformative prior distribution is for the parameters parameters λ and γ of the EEG distribution and to study the effects of these different priors in the resulting posterior distributions, especially in situations of small sample sizes, a common situation in applications.
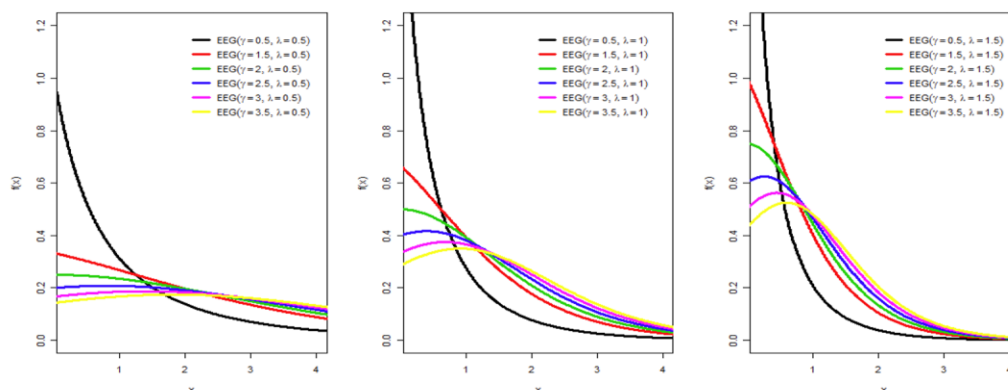
Figure 1. Probability density functions for the EEG distribution with values for the scale and shape parameters given respectively, by λ = 0.5, 1.0, 1.5 and γ = 0.5, 1.5, 2.0, 2.5, 3.0, 3.5.
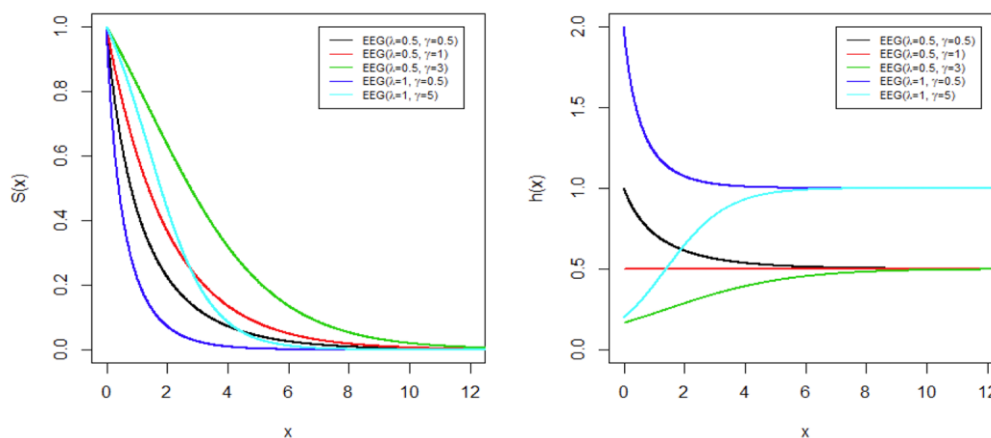
Figure 2. Survival functions and hazard functions for the EEG distribution with values for the scale and shape parameters given respectively, by λ = 0.5, 1.0 and γ = 0.5, 1, 3, 5.

Commonly used noninformative prior distributions are derived, such as uniform (Bayes, 1763; Laplace, 1774), Jeffreys (1967), reference priors (Bernardo, 1979; Berger & Bernardo, 1992), and the uncommon MDIP prior (Zellner, 1977, 1984). A simulation study is conducted comparing their performance in terms of their summaries and coverage rates of credible intervals.

Numerical integration based on stochastic simulation methods as the Markov Chain Monte Carlo (MCMC) will be used to simulate samples of the marginal posterior distribution of interest. In particular, we will be using the Metropolis-Hastings algorithm to obtain the posterior summaries of interest (see Gelfand & Smith, 1990 or Chib & Greenberg, 1995).

Methodology

Maximum Likelihood Estimation

Let $X_1, \ldots, X_n$ be a random sample from EEG(γ, λ) distribution with density (1). The likelihood function for the parameters γ and λ is given by

$$L(x;\gamma,\lambda) = (\lambda\gamma)^n \exp\left(-\lambda\sum_{i=1}^{n} x_i\right)\prod_{i=1}^{n}\left\{\left(1-(1-\gamma)e^{-\lambda x_i}\right)^{-2}\right\}, \tag{4}$$

where $\gamma > 0$ and $\lambda > 0$.

The logarithm of the likelihood function (4) is given by

$$l(x;\gamma,\lambda) = n\log(\lambda\gamma) - \lambda\sum_{i=1}^{n} x_i - 2\sum_{i=1}^{n}\log\left(1-(1-\gamma)e^{-\lambda x_i}\right). \tag{5}$$

By setting $\partial l(x;\gamma,\lambda)/\partial\gamma = 0$ and $\partial l(x;\gamma,\lambda)/\partial\lambda = 0$ and after some algebraic manipulations, we obtain the likelihood equations

$$\frac{n}{\lambda} - \sum_{i}^{n} x_i - 2(1-\lambda)\sum_{i}^{n}\frac{x_i e^{-\lambda x_i}}{1-(1-\gamma)e^{-\lambda x_i}} = 0 \text{ and}$$

$$\frac{n}{\lambda} - 2\sum_{i}^{n}\frac{e^{-\lambda x_i}}{1-(1-\gamma)e^{-\lambda x_i}} = 0 \tag{6}$$

whose solutions provide the maximum likelihood estimators of the parameters γ and λ. Note that the solutions of the likelihood equations (6) cannot be obtained analytically and hence numerical approaches need to be used.

Adamidis et al. (2005) propose to use the EM algorithm (Dempster et al., 1977) to solve the nonlinear equations (6) and find the MLE of γ and λ. The EM iterations are given by

$$\hat{\gamma}^{(j+1)} = n \left[\sum_{i=1}^{n} \frac{1+\left(1-\hat{\gamma}^{(j)}\right)\exp\left(-\hat{\lambda}^{(j)}x_i\right)}{1-\left(1-\hat{\gamma}^{(j)}\right)\exp\left(-\hat{\lambda}^{(j)}x_i\right)} \right]^{-1}$$

$$\hat{\lambda}^{(j+1)} = n \left[\sum_{i=1}^{n} \frac{x_i\left(1+\left(1-\hat{\gamma}^{(j)}\right)\exp\left(-\hat{\lambda}^{(j)}x_i\right)\right)}{1-\left(1-\hat{\gamma}^{(j)}\right)\exp\left(-\hat{\lambda}^{(j)}x_i\right)} \right]^{-1}.$$

(7)

Kitidamrongsuk (2010) shows in detail the computations of the expected Fisher information matrix $\mathbf{I}(\gamma,\lambda)$ of the EEG distribution, given by

$$\mathbf{I}(\gamma,\lambda) = \begin{bmatrix} \mathbf{I}_{11}(\gamma,\lambda) & \dfrac{n\left(\gamma-1-\gamma^2\log(\gamma)\right)}{3\lambda\gamma(\gamma-1)^2} \\[4mm] \dfrac{n\left(\gamma-1-\gamma^2\log(\gamma)\right)}{3\lambda\gamma(\gamma-1)^2} & \dfrac{n}{3\gamma^2} \end{bmatrix},$$

(8)

with

$$\mathbf{I}_{11}(\gamma,\lambda) \begin{cases} \dfrac{n\left[3(1-\gamma)-2\left((1-\gamma)-\gamma L_2(1-\gamma)\right)\right]}{3\lambda^2(1-\gamma)} & ,0<\gamma<1 \\[4mm] \dfrac{n}{3\lambda^2(1-\gamma)}\left[1-\gamma\left(1+\dfrac{\pi^2}{3}+\log^2(\gamma)-2L_2\left(\dfrac{1}{\gamma}\right)\right)\right] & ,\gamma>1 \end{cases}$$

(9)

where

$$L_r(p) = \sum_{i=1}^{\infty} \frac{p^i}{i^r}, (0<p<1)$$

is a polygarithmica function (Erdelyi et al., 1953). The maximum likelihood estimates for γ and λ are biased for small sample problems. In the case of large samples they become unbiased and asymptotically efficient. Such estimates are asymptotically Normal distributed with joint distribution given by

$$\left(\hat{\gamma},\hat{\lambda}\right) \sim N_2\left[\left(\gamma,\lambda\right), I^{-1}\left(\gamma,\lambda\right)\right], \text{ when } n \to \infty. \tag{10}$$

Bayesian Analysis

In this section we consider the Bayesian estimation of the unknown parameters λ and γ.

First, a prior distribution which expresses little information on the parameters γ and λ can be obtained from uniform densities, which do not favor any particular value of λ and γ. In this case, the joint prior distribution for λ and γ is given by

$$\pi_U\left(\gamma,\lambda\right) \propto \text{ constant.} \tag{11}$$

Another widely-used method to specify prior information is through the product of independent gamma distributions for each parameter λ and γ, since $\gamma > 0$ and $\lambda > 0$, that is, $\gamma \sim \text{Gamma}(\alpha_1, \beta_1)$ and $\lambda \sim \text{Gamma}(\alpha_2, \beta_2)$, where *Gamma(a,b)* denotes a gamma distribution with mean a/b and variance a/b^2; and $\alpha_1, \alpha_2, \beta_1$ and β_2 are known hyperparameters. Thus, the joint prior distribution for λ and γ is given by

$$\pi_G\left(\gamma,\lambda\right) \propto \gamma^{\alpha_1-1}\lambda^{\alpha_2-1}\exp\left(-\beta_1\gamma - \beta_2\lambda\right). \tag{12}$$

Assume $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0.01$, that is, a non-informative prior given by (12).

An another well-known existing non-informative prior, which represents a situation with little a priori information on the parameters was introduced by Jeffreys (1967), also known as the Jeffreys rule. The Jeffreys prior has been widely used due to the invariance property for one to one transformations of the parameters. The Jeffreys prior is defined as

$$\pi_J\left(\gamma,\lambda\right) \propto \sqrt{\det \mathbf{I}\left(\gamma,\lambda\right)}. \tag{13}$$

where $\mathbf{I}(\gamma, \lambda)$ is the Fisher information matrix defined in (8) and (9).

From the equation (13), we get

$$\pi_J(\gamma,\lambda) \propto \sqrt{\frac{-2\left((\gamma-1)^3 L_2(1-\gamma)-(-\gamma+\gamma\log(\gamma)+1)(\gamma^2+\gamma^2\log(\gamma)-3\gamma+2)\right)}{\gamma(\gamma-1)^4}} \times$$

$$\times\frac{1}{\lambda}; \quad 0<\gamma<1,$$

$$\pi_J(\gamma,\lambda) \propto \sqrt{\frac{(\gamma-1)^3\left(1-\frac{\gamma}{3}\left(-6L_2\left(\frac{1}{\gamma}\right)+3\log^2(\gamma)+\pi^2+3\right)\right)-\left(\gamma^2\log(\gamma)-\gamma+1\right)^2}{\gamma(\gamma-1)^4}} \times$$

$$\times\frac{1}{\lambda}; \quad \gamma>1.$$

(14)

It is interesting to observe it was found in (14) independent priors for the parameters λ and γ, but this joint prior has a dependence structure.

Zellner (1977, 1984) proposed a non-informative prior based on the Shannon's entropy (1948). The idea is to maximize the information from the data in relation to the prior information on the parameters. This non-informative prior distribution known as "Maximal Data Information Prior" is obtained from the solution of the equation

$$\pi_Z(\gamma,\lambda) \propto \exp\left(\int_0^\infty \log\left(f(x\mid\gamma,\lambda)\right)f(x\mid\gamma,\lambda)dx\right).$$

(15)

For the EEG distribution given in (1) the resulting non-informative prior is given by

$$\pi_Z(\gamma,\lambda) \propto (\lambda\gamma)\exp\left(-\gamma\Psi(1-\gamma,1,1)+\frac{2\log(\gamma)+2-2\gamma}{(\gamma-1)}\right),$$

(16)

where $\Psi(z,s,a)$ is defined in (5).

The proposed Zellner prior distribution (15) has limited invariance properties, where invariance is only verified under linear transformations of the vector (γ, λ) and not for all differentiable one by one transformations. Bernardo (1979) and Berger & Bernardo (1992) use the Kullback-Liebler distance between the posterior distribution $p(\theta\mid x)$ the prior distribution $\pi(\theta)$ to maximize the information from the data in relation to the known prior information for the parameters to find a non-informative prior. Additional information about the reference prior can be found in Bernardo (2005).

An important feature in this approach is the different treatment for interest and nuisance parameters when $\boldsymbol{\theta}$ is a vector of parameters. In the presence of nuisance parameters, a typical case in this paper, one must establish an ordered parameterization with the parameter of interest singled out and then follow the procedure below. The algorithm of Berger and Bernardo (1992) to derive the reference prior can be described in four steps, as follows. We will present here the two-parameters case in details.

Let $\boldsymbol{\theta} = (\theta_1, \theta_2)$ be the two parameters vector; θ_1 will be considered the parameter of interest and θ_2 is the nuisance parameter. The algorithm used to obtain the reference prior is given by

Step 1: Find the conditional reference prior $\pi_2(\theta_2 \mid \theta_1)$, assuming that θ_1 is given by

$$\pi_2\left(\theta_2 \mid \theta_1\right) = \sqrt{\mathbf{I}_{22}\left(\theta_1, \theta_2\right)}, \tag{17}$$

where $\mathbf{I}_{22}(\theta_1, \theta_2)$ is the term of order (2,2) of the information Fisher matrix.

Step 2: Normalize $\pi_2(\theta_2 \mid \theta_1)$. If $\pi_2(\theta_2 \mid \theta_1)$ is improper, choose a sequence of sets $\Omega_1 \subseteq \Omega_2 \subseteq \cdots \to \Omega$, where $\pi_2(\theta_2 \mid \theta_1)$ is proper. Find

$$c_m\left(\theta_1\right) = \frac{1}{\int_{\Omega_m} \pi_2\left(\theta_2 \mid \theta_1\right) d\theta_2} \tag{18}$$

$$p_m\left(\theta_2 \mid \theta_1\right) = c_m\left(\theta_1\right)\pi_2\left(\theta_2 \mid \theta_1\right)1_{\Omega_m}\left(\theta_2\right). \tag{19}$$

Step 3: Find the reference prior for θ_1. The result is given by the solution of

$$\pi_m\left(\theta_1\right) \propto \exp\left\{\frac{1}{2}\int_{\Omega_m} p_m\left(\theta_2 \mid \theta_1\right)\log\left\|\frac{\det \mathbf{I}\left(\theta_1, \theta_2\right)}{\mathbf{I}_{22}\left(\theta_1, \theta_2\right)}\right\| d\theta_2\right\}. \tag{20}$$

Step 4: Find the prior distribution for (θ_1, θ_2), when θ_2 is the nuisance parameter

$$\pi_2(\theta_1, \theta_2) = \lim_{m \to \infty} \left(\frac{c_m(\theta_1) \pi_m(\theta_1)}{c_m(\theta_1^*) \pi_m(\theta_1^*)} \right) \pi_2(\theta_2 \mid \theta_1), \tag{21}$$

where θ_1^* is any fixed point within the positive density for all π_m.

For the EEG distribution given in (1), the reference prior when λ is the parameter of interest is given by

$$\pi_\lambda(\lambda, \gamma) \propto \begin{cases} \sqrt{\dfrac{(1-\gamma) + 2\gamma L_2(1-\gamma)}{(1-\gamma)}} & 0 < \gamma < 1 \\[4mm] \sqrt{\dfrac{1}{(1-\gamma)} \left[1 - \gamma \left(1 + \dfrac{\pi^2}{3} + \log^2(\gamma) - 2L_2\left(\dfrac{1}{\gamma}\right) \right) \right]} & \gamma > 1 \end{cases} \tag{22}$$

The reference prior when γ is the parameter of interest is given by

$$\pi_\gamma(\lambda, \gamma) \propto \sqrt{\frac{\gamma\left(-2\left((\gamma-1)^3 L_2(1-\gamma) - (-\gamma + \gamma \log(\gamma) + 1)(\gamma^2 + \gamma^2 \log(\gamma) - 3\gamma + 2) \right) \right)}{(\gamma-1)^4}}$$

for $0 < \gamma < 1$ and $\tag{23}$

$$\pi_\gamma(\lambda, \gamma) \propto \sqrt{\frac{\gamma\left((\gamma-1)^3 \left[1 - \dfrac{\gamma}{3}\left(-6L_2\left(\dfrac{1}{\gamma}\right) + 3\log^2(\gamma) + \pi^2 + 3 \right) \right] - \left(\gamma^2 \log(\gamma) - \gamma + 1 \right)^2 \right)}{(\gamma-1)^4}}$$

for $\gamma > 1$

Finally, derive the prior distributions for the parameters the resulting joint posterior distributions for γ and λ is proportional to the product of the likelihood function (4) and the prior distributions $\pi(\gamma, \lambda)$ given in (11), (12), (14), (16), (22) and (23), that is

$$p(\gamma, \lambda \mid x) \propto \pi(\gamma, \lambda)(\lambda \gamma)^n \exp\left(-\lambda \sum_{i=1}^n x_i \right) \prod_{i=1}^n \left\{ \left(1 - (1-\gamma)e^{-\lambda x_i} \right)^{-2} \right\}. \tag{24}$$

By using any prior distribution proposed is not possible to derive the marginal posterior distributions in an analytical form for the parameters γ and λ. Thus, to obtain the posterior information on the parameters of interest as the point estimator and Bayes credibility intervals, we use MCMC algorithms to simulate samples of the values of γ and λ from the joint posterior distributions.

Results

Two applications of the theoretical results discussed in the previous sections are presented. The first involves a comparison of the estimation methods based on simulated data; the second shows an application of the EEG distribution to real data.

Analysis via numerical simulation

In this example, some simulations are performed via the Monte Carlo method. The goal is to study the effect of different non-informative prior distributions on the posterior summaries and also to compare these results with the obtained results using classical inference analysis. Posterior summaries of interest are evaluated using Monte Carlo Markov Chain (MCMC) methods. The influence of sample size on the accuracy of the obtained estimators is also examined. The following procedure was adopted:

1. Determine the values of γ and λ.
2. Specify the sample size n.
3. Generate values of a distribution EEG(γ,λ) with size n.
4. Using the data obtained in Step 3, calculate the estimates for the parameters γ and λ using MCMC in the Bayesian approach and MLE in the classical approach.
5. Repeat the steps 3 and 4 N times.

Consider two set of the true values for the parameter (γ, λ) given by $(\gamma, \lambda) = (0.5, 2)$ and $(\gamma, \lambda) = (2, 4)$ representing decreasing and increasing hazard functions, respectively. The simulated data are generate from EEG distribution with the parameter values above for different sample sizes, as $n = 10, 25$ and 50.
Tables 1 and 2 show the posterior mean and median, respectively, by considering the non-informative priors proposed in this paper for the parameters. The maximum likelihood estimates (MLE) are also available.

Table 1. Posterior medians and MLE for $\lambda = 2$ and $\gamma = \frac{1}{2}$ for 1000 samples of sizes 10, 25 and 50.

$\lambda = 2$	Jeffreys	MDIP	Ref. λ	Ref. γ	Uniform	Gamma	MLE
n=10	2.98(1.62)	2.80(0.84)	2.96(0.87)	2.88(0.81)	3.61(1.54)	1.15(1.06)	3.44(2.61)
n=25	2.62(1.10)	3.00(1.05)	2.93(0.88)	2.92(0.83)	2.95(1.00)	2.01(1.06)	2.54(1.26)
n=50	2.17(0.76)	2.52(0.70)	2.58(0.67)	2.71(0.67)	2.56(0.72)	2.07(0.77)	2.26(0.80)

$\gamma = \frac{1}{2}$	Jeffreys	MDIP	Ref. λ	Ref. γ	Uniform	Gamma	MLE
n=10	1.01(0.43)	1.31(0.77)	1.12(0.17)	1.13(0.14)	1.50(0.77)	0.35(0.49)	1.51(2.02)
n=25	0.87(0.39)	0.92(0.42)	0.95(0.23)	0.99(0.21)	0.97(0.46)	0.56(0.41)	0.81(0.68)
n=50	0.63(0.32)	0.71(0.27)	0.81(0.27)	0.89(0.26)	0.76(0.32)	0.56(0.29)	0.64(0.36)

Table 2. Posterior medians and MLE for $\lambda = 4$ and $\gamma = 2$ for 1000 samples of sizes 10, 25 and 50.

$\lambda = 4$	Jeffreys	MDIP	Ref. λ	Ref. γ	Uniform	Gamma	MLE
n=10	4.85(1.21)	4.62(1.50)	5.31(1.55)	4.73(0.99)	5.08(1.25)	5.08(1.25)	4.86(2.03)
n=25	4.56(0.81)	4.02(0.77)	3.82(0.71)	3.90(0.73)	4.77(0.87)	3.50(0.84)	4.41(1.42)
n=50	3.76(0.57)	3.98(0.61)	3.51(0.41)	3.56(0.40)	3.40(0.64)	3.67(0.65)	4.22(0.94)

$\gamma = 2$	Jeffreys	MDIP	Ref. λ	Ref. γ	Uniform	Gamma	MLE
n=10	3.07(0.50)	1.81(0.40)	3.25(0.42)	3.33(0.40)	3.26(0.42)	3.26(0.42)	3.65(3.23)
n=25	2.70(0.49)	1.76(0.34)	1.62(0.25)	1.74(0.27)	2.94(0.53)	1.52(0.56)	2.87(2.06)
n=50	1.71(0.45)	1.90(0.41)	1.53(0.29)	1.60(0.32)	1.95(0.45)	1.70(0.52)	2.43(1.16)

From Tables 1 and 2, it is observed that when the hazard function is decreasing $(0 < \gamma < 1)$ the prior distribution given by product of independent gamma distributions gives the best estimation for the parameters while for the increasing hazard function $(\gamma > 1)$ the MDIP prior distribution provides the best one for all sample sizes considered.

A criterion for comparison of the prior distributions consists on checking the frequentist coverage probabilities of the posterior intervals. We therefore compare the frequency at which the true values of γ and λ are included in their 95% posterior intervals. This frequency should be close to 95% for large numbers of repeated experiments.

Table 3. Coverage probabilities for λ = 2 and γ = ½ for 1000 samples of sizes 10, 25 and 50.

λ = 2	Jeffreys	MDIP	Ref. λ	Ref. γ	Uniform	Gamma	MLE
n=10	98.50%	99.10%	95.70%	96.70%	90.60%	96.50%	95.20%
n=25	95.20%	91.60%	89.70%	90.00%	93.40%	97.20%	95.00%
n=50	97.50%	94.30%	95.00%	91.30%	92.20%	96.60%	94.70%
γ = ½	Jeffreys	MDIP	Ref. λ	Ref. γ	Uniform	Gamma	MLE
n=10	99.20%	98.00%	95.70%	94.80%	90.60%	97.10%	94.60%
n=25	95.50%	97.80%	97.90%	97.70%	95.00%	97.80%	95.50%
n=50	97.60%	97.90%	96.00%	95.50%	93.70%	97.70%	95.40%

Table 4. Coverage probabilities for λ = 4 and γ = 2 for 1000 samples of sizes 10, 25 and 50.

| λ = 4	Jeffreys	MDIP	Ref. λ	Ref. γ	Uniform	Gamma	MLE
n=10	95.50%	96.60%	88.10%	98.00%	91.20%	91.20%	96.60%
n=25	94.60%	98.20%	96.80%	97.00%	94.00%	95.60%	94.30%
n=50	97.30%	98.40%	96.00%	97.60%	98.50%	96.60%	95.00%
γ = 2	Jeffreys	MDIP	Ref. λ	Ref. γ	Uniform	Gamma	MLE
n=10	97.10%	99.50%	93.80%	97.00%	96.60%	96.60%	92.50%
n=25	98.50%	99.10%	99.50%	99.90%	98.50%	97.10%	92.00%
n=50	98.20%	98.90%	97.00%	99.50%	98.80%	96.50%	93.90%

An example with literature data

Now consider a lifetime dataset related to an electrical insulator subjected to constant stress and strain, introduced by Lawless (1982). The dataset does not have censored values and represents the lifetime (in minutes) to failure: 0.96, 4.15, 0.19, 0.78, 8.01, 31.75, 7.35, 6.50, 8.27, 33.91, 32.52, 16.03, 4.85, 2.78, 4.67, 1.31, 12.06, 36.71 and 72.89.

Assume that the EEG distribution is appropriated to analyze this dataset, and then it will be compared with other lifetime distributions such as Weibull, Gamma, and Lognormal. As shown, the efficiency of the different non-informative prior distributions changes with the shapes of the hazard functions, therefore, to get good inferences on parameters of interest it is necessary to have some prior information on how the hazard function behaves for the Lawless data set. In this way, Barlow & Campo (1975) proposed a simple graphical technique that has been widely used to verify the behavior of the risk function called TTT plot (total time for testing). The graph is constructed with the plot of the consecutive quantities [r/n, $G(r/n)$], where $G(r/n)$ is a function given by

$$G(r/n) = \frac{\sum_{i=1}^{r} t_i + (n-r)t_{(r)}}{\sum_{i=1}^{n} t_i},$$

where $r = 1, ..., n$ and $t_{(i)}$, $i = 1, ..., n$ are the order statistics in the sample.

Using the TTT curve in an empirical scale, one can determine the shape of the hazard function for the lifetime data. A diagonal line indicates that the data have a constant hazard function; if the curve is convex the risk is decreasing; if it is concave, there is an indication that the risk is increasing; if first is convex and after this is concave then there is an indication that there is a bathtub shape for the hazard function; if it is first concave and after this convex, there is an indication of inverse form of the bath for the hazard function. The Figure 3 shows how to verify the behavior of the hazard function.
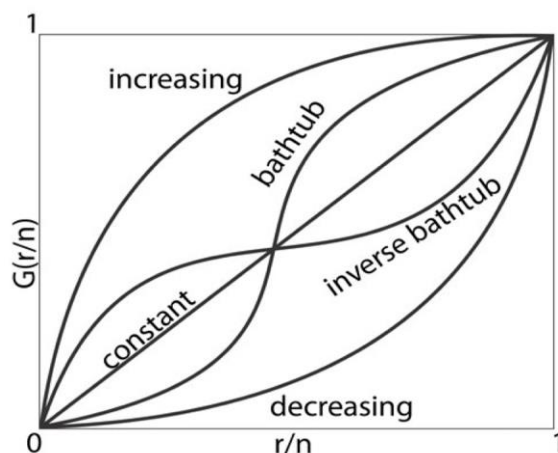
Figure 3. TTT plots for different distributions indicating the shape of the hazard function.

Some TTT transformations can be studied to solve other problems. Nair et al. (2008) show some of these transformations applied in survival analysis. Figure 4 shows the TTT plot for the Lawless data set.
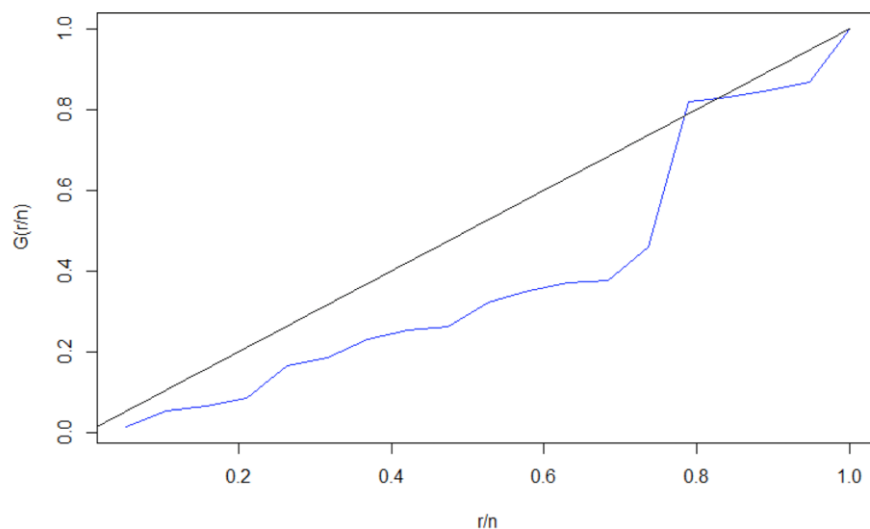
Figure 4. TTT plot for the dataset lifetime related to an electrical insulator subjected to constant stress and strain (Lawless data).

It is observed in Figure 4 that the TTT plot is convex; then it can be concluded that the risk is a decreasing function. When the hazard function is decreasing, it was observed from the results of section 4.1, that non-informative priors obtained through the product of independent gamma distributions is the best prior with little prior information about the parameters of interest. The joint posterior distribution of λ and γ (24) is obtained by replacing $\pi(\gamma, \lambda)$ by (12). It is necessary to use numerical methods to extract information from the marginal posterior distributions λ and γ. MCMC methods are used to simulate samples for the joint posterior distribution; that is, also for the marginal posterior distributions of interest.

It was generated 110,000 iterations with a "burn-in" of 10,000 values and jumps of size 10; so we get chains of the marginal posterior distributions for λ and γ of size 10,000 obtained using MCMC methods. To verify the convergence of the chains, we have used Geweke (1992) diagnosis, which indicated the convergence of the two chains. The convergence and autocorrelations is also observed in the trace-plots of the simulated series given in Figure 5.

To verify the performance of other lifetime distributions we also consider as non-informative prior, the product of independent gamma distributions $\gamma \sim Gamma$ (0.01, 0.01), $\lambda \sim Gamma$ (0.01, 0.01) assuming the following lifetime distributions: EGE, Weibull, Gamma and Lognormal distribution. The results are compiled in Tables 5 and 6, respectively.

239
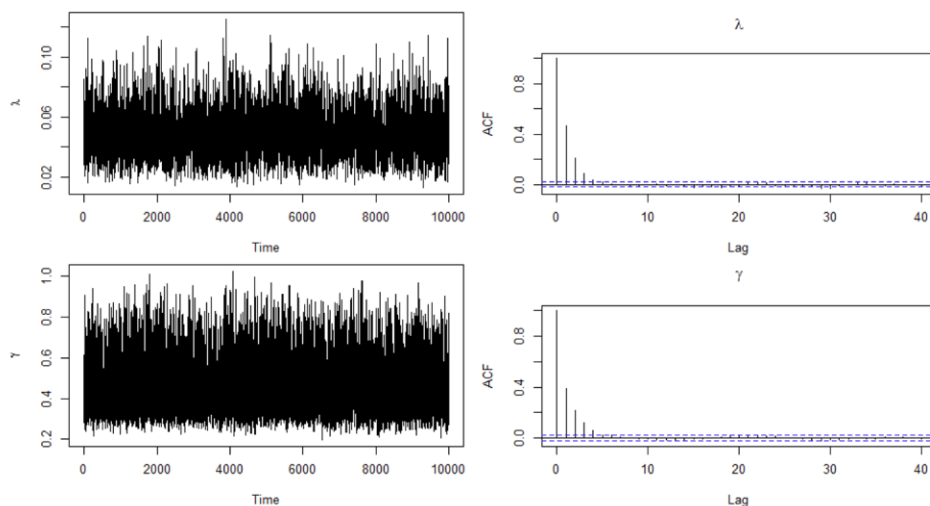
Figure 5. Trace-plots and autocorrelation graphs for the generated values of λ and γ.

Table 5. Posterior estimates (means) for γ and λ considering different probability distributions for the Lawless data set.

Parameters	EGE	Weibull	Gamma	Log-normal
γ	0.0482 (0.0161)	0.7629(0.1356)	0.6725(0.1870)	1.6880(0.3969)
λ	0.4513 (0.1649)	7.8250(4.2250)	0.0474(0.0183)	2.6410(1.0200)

Table 6. Obtained results for the DIC, BIC and AIC criteria for the different probability distributions for the Lawless data set.

Criteria	EGE	Weibull	Gamma	Log-normal
DIC	**138.96**	140.70	141.30	141.30
BIC	**140.55**	148.06	143.13	151.62
AIC	**142.44**	149.95	141.24	149.73

Based on any of the criteria used by the table it can be concluded that EGE was the best fit to the offered data.

Conclusion

The use of extended exponential geometric (EEG) distributions showed a good flexibility of fit for lifetime data applications and could be an alternative

distribution to other usual distributions in Survival analysis. The great number of existing non-informative prior distributions can cause difficulties in the choice of an adequate prior with little information a priori, mainly when these prior distributions do not produce similar posterior summaries. In this way, the development of a general theory for the construction of non-informative prior distributions is an important topic to be investigated by researchers in the Bayesian inference.

The results showed the effects of different non-informative prior distributions related to the changes in the risk function using extended exponential geometric (EEG) distributions. Therefore, we recommend the product of gamma distributions Gamma (0.01, 0.01) when the hazard function is decreasing and the non-informative MDIP prior distribution when the hazard function is increasing. With these choices of prior we surely get better inferences for the parameters.

References

Achcar, J. A, Moala, F. A, & Boleta, J. (March, 2010). *Generalized Exponential Distribution: A Bayesian approach using MCMC methods.* Poster presented at the 10th Bayesian Statistics Brazilian Meeting, Rio de Janeiro, Brazil.

Adamidis, K, Dimitrakopoulou, T, & Loukas, S. (2005). On an Extension of the Exponential Geometric Distribution. *Statistics and Probability Letters*, *73*, 259-269.

Adamidis, K, & Loukas, S. (1998). A Lifetime Distribution with Decreasing Failure Rate. *Statistics and Probability Letters, 39*, 35-42.

Barlow, R. E., & Campo, R. A. (1975). Total Time on Test processes and applications to failure data analysis. In R. E. Barlow, J. B. Fussel, & N. D. Singpurwalla (Eds.), *Reliability and Fault Tree Analysis: Theoretical and Applied Aspects of System Reliability and Safety Assessment* (451-481). Philadelphia, PA: Society for Industrial and Applied Mathematics.

Bayes, T. R. (1958). Essay towards solving a problem in the doctrine of changes. Reprinted in *Biometrika, 45*, 243-315, 1763.

Berger, J. O, & Bernardo, J. M. (1992). *On the Development of the Reference Prior Method*. Fourth Valencia International Meeting on Bayesian Statistics, Spain.

Bernardo, J. M. (1979). Reference Posterior Distributions for Bayesian Inference. *Journal Royal Statistical Society, 41*(2), 113-147.

Bernardo, J. M. (2005). Reference analysis. In D.K. Dey and C. R. Rao, (Eds.), *Handbook of Statistics, 25* (17-90). Amsterdam: Elsevier.

Box, G. E. P., & Tiao, G. C. (1973). *Bayesian inference in statistical analysis.* Reading, MA: Addison-Wesley.

Chib, S., & Greenberg, E. (1995). Understanding The Metropolis-Hasting Algorithm. *The American Statistician, 49*(4), 327-335.

Degroot, M. H., & Schervish, M. J. (2002). *Probability and Statistics*. Reading, MA: Addison-Wesley.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm (with discussion). *Journal of the Royal Statistical Society: Series B (Methodological), 39*(1), 1-38.

Erdelyi, A., Maguns, W., Oberhettinger, F., & Tricomi, F.G. (1953). *Higher Transcendental Functions.* New York: McGraw-Hill.

Gelfand, A. E., & Smith, F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association, 85*, 398-409.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In J. M. Bernado, J. O. Berger, A. P. Dawid, & A. F. M. Smith (Eds.), *Bayesian Statistics 4*. Oxford, UK: Clarendon Press.

Gupta, R. D., & Kundu, D. (1999). Generalized exponential distributions. *Australian and New Zealand Journal of Statistics, 41*, 173-188.

Gupta, R. D., & Kundu, D. (2001). Generalized exponential distributions: different methods of estimation. *Journal of Statistical Computation and Simulation, 69*, 315-338.

Gupta, R. D., & Kundu, D. (2007). Generalized exponential distribution: existing results and some recent developments. *Journal of Statistical Planning and Inference, 137*(11), 3537-3547. doi: 10.1016/j.jspi.2007.03.030.

Gupta, R. D., & Kundu, D. (2008). Generalized exponential distribution; Bayesian Inference. *Computational Statistics and Data Analysis, 52*(4), 1873-1883.

Hamada, M., Wilson, A. G., Reese, C. S., & Martz, H. F. (2008). *Bayesian Reliability*. New York: Springer.

Jeffreys, H. (1967). *Theory of probability* (3rd ed.). London: Oxford University Press.

Kitidamrongsuk, P. (2010). *Discriminating Between the Extended Exponential Geometric Distribution and the Gamma Distribution*. (Unpublished Doctoral Dissertation). Assumption University of Thailand.

Laplace, P. (1774). Mémoire sur la probabilité des causes par les é venemens. Mem. Acad. R. Sci. Presentés par Divers Savans, v. 6, p. 621-656 (translated in *Statistical Science*, 1, 359-378).

Lawless, J. F. (1982). *Statistical models and methods for lifetime data.* New York: John Wiley and Sons.

Marshall, A.W., & Olkin, I. (1997). A New Method for Adding a Parameter to a Family of Distributions with Application to the Exponential and Weibull Families. *Biometrika, 84*, 641-652.

Moala, F. A. (2010). Bayesian analysis for the Weibull parameters by using noninformative prior distributions. *Advances and Applications in Statistics, 14*(2), 117-143.

Nair, N. U., Sankaran, P. G., & Kumar, B. V., (2008). Total time on test transforms of order *n* and its implications in reliability analysis. J*ournal of Applied Probability, 45*(4), 1126–1139.

Raqab, M. Z. (2002). Inferences for generalized exponential distribution based on record statistics. *Journal of Statistical Planning and Inference, 104*, 339-350.

Raqab, M. Z., & Ahsanullah, M. (2001). Estimation of the location and scale parameters of generalized exponential distribution based on order statistics. *Journal of Statistical Computation and Simulation, 69*, 109-124.

Sarhan, A. M. (2007). Analysis of incomplete, censored data in competing risks models with generalized exponential distributions. *IEEE Transactions on Reliability, 56*(1), 132-138.

Shannon, C. E. (1948). A Mathematical theory of communication. *Bell System Technical Journal, 27*, 623-659.

Zellner, A. (1977). Maximal Data Information Prior Distributions. In A. Aykac & C. Brumat (Eds.), *New Methods in the applications of Bayesian Methods*. Amsterdam: North-Holland Publishing Co.

Zellner, A. (1984). *Maximal Data Information Prior Distributions: Basic Issues in Econometrics*. Chicago, IL: University of Chicago Press.

Zheng, G. (2002). Fisher information matrix in type-II censored data from exponentiated exponential family. *Biometrical Journal, 44*, 353-357.