International Journal of Mathematical Archive-11(6), 2020, 37-52
Available online through www.ijma.info ISSN 2229 – 5046

ON A CHARACTERIZATION OF THE EXPONENTIATED EXPONENTIAL DISTRIBUTIONBASED ON THE MINIMUM CHI-SQUARED DIVERGENCE PRINCIPLE

MAHMOUD ABD ALMO'MEN ATALLAH*[1], ABDULAZIZ ABD ULLAH ALQARNI[2]

[1]**Department of Physics and Engineering Mathematics, Faculty of Engineering, AIN Shams University, Cairo, Egypt.**

[2]**Department of General Studies, Jeddah College of Technology, Technical and Vocational Training Corporation, Jeddah, Saudi Arabia.**

(Received On: 27-02-20; Revised & Accepted On: 19-05-20)

ABSTRACT

A probability distribution can be characterized through various methods. Given a prior probability distribution g(x) and some available information on moments of the random variable X, the original probability distribution f(x) is determined such that the χ^2- divergence measure of the distance between f(x) and g(x) is a minimum. The minimum chi-squared density function f(x) is determined. The expressions of the non-central moments as well as the cumulative distribution function F(x), survival function S(x), and hazard function h(x) are also determined under different available information on moments. In this paper we discussed the characterization of the Exponentiated exponential (EExp) distribution. The available information on moments included: first moment, second moment, the first two moments, and the first three moments. Some illustrative examples are included for special values of the parameters. We tabulated the results and the corresponding characteristics of the distribution are graphically, compared.

Keywords: *Characterization, Exponentiated Exponential Distribution, Chi-square Divergence, Hazard Function, Available Information, Optimization Principle.*

1. INTRODUCTION

In mathematics in general, a characterization theorem says that a particular object a function, a space, etc. is the only one that possesses properties specified in the theorem. A characterization of a probability distribution accordingly states that it is the only probability distribution that satisfies specified conditions. More precisely, the model of characterization of probability distribution was described by Zolotarev [1976], in such manner. On the probability space we define the space $x = \{X\}$ of random variables with values in measurable metric space (U, d_u) and the space $y = \{Y\}$ of random variables with values in measurable metric space (V, d_u) by characterizations of probability distributions we understand general problems of description of some set C in the space x by extracting the sets $A \subseteq x$ and $B \subseteq y$ which describe the properties of random variables $X \in A$ and their images $Y = FX \in B$, obtained by means of a specially chosen mapping $F: x \to y$. The description of the properties of the random variables X and of their images $Y = FX$ is equivalent to the indication of the set $A \subseteq x$ from which X must be taken and of theset $B \subseteq y$ into which its image must fall. So, the set which interests us appears therefore in the following form:

$$X \in A, FX \in B \Leftrightarrow X \in C, i.e. C = F^{-1}B,$$

Where $F^{-1}B$ denotes the complete inverse image of B in A.This is the general model of characterization of probability distribution; see Hamedani and Ahsanullah [2011].

The problem of characterizing a distribution is an important problem which has recently attracted the attention of many researchers. Thus, various characterizations have been established in many different directions see Yanushkevichius [2014]. The minimum χ^2- divergence principle states: When a prior probability density function (PDF) of X, g(x), which estimates the underlying PDF f(x) is given in addition to some constraints, then among all the density functions f(x) which satisfy the given constraints we should select that probability density function which minimizes the $\chi 2$- divergence, Kumar [2005].

Corresponding Author: Mahmoud AbdAlMo'men Atallah*[1], Department of Physics and Engineering Mathematics, Faculty of Engineering, Ain Shams University, Cairo, Egypt.

The rest of the article is organized as follows: section 2 presentsminimum chi-squared divergence probability distributions. In section 3, the exponentiated exponential distribution, and The Available information of the moments for the EExp distribution is presented. Finally, section 4 offers, Illustrative Examples, and concluding remarks.

2. MINIMUM CHI-SQUARED DIVERGENCE PROBABILITY DISTRIBUTIONS

The maximum entropy principle Jaynes [(1957] and the minimum discrimination (minimum cross-entropy) principle Kullback(1997) are used in characterization of many univariate and multivariate probability distributions see Kagan *et al.* [1973], Kapur [1983], Kapurand Kesavan [1989]. These two principals were used, Kapur [1982] in characterizing some univariate distributions. The above-mentioned principles are generalized, Kapur and Kesavan. [1989]. Kawamuraand Iwase [2003] applied the maximum entropy principlein characterizing the power inverse Gaussian, and generalized gumbel probability distributions. The equivalence of minimum discrimination information principle and the maximum likelihood principle and Gauss's principle have been discussed in Campbell [1970] and Shoreand Johnson [1980]. In literature, the chi-square divergence principle is well known. Here, we make use of and prove the results of Kumar (2006) and apply them to characterize the EExp distribution, under different available information on moments.

Let the random variable X be a continuous variable with PDFf(x), defined on the interval (0, ∞). According to Kullback [1997], the minimum cross entropy principle states that: When a prior PDF of X, g(x), which estimates the underlying PDFf(x) is given in addition to some constraints, then among all the density functions f(x) which satisfy the given constraints we should select that PDF which minimizes the Kullback and Leibler divergence:

$$K(f,g)=\int_0^\infty f(x)\ln\frac{f(x)}{g(x)}\,dx \tag{2.1}$$

Kumar and Taneja [2004] have considered the minimum χ^2– divergence principle is as follows:
when a prior PDF X, g(x) of the random variable is the prior PDF estimating the underlying PDF f(x) is given with some available information on moments of f(x), then f(x) is called the minimum χ^2 – divergence PDF if it satisfies the given partial available information and minimizes the χ^2 – divergence:

$$\chi^2(f,g)=\int_0^\infty \frac{f^2(x)}{g(x)}dx-1. \tag{2.2}$$

The minimum cross-entropy principle and the minimum χ^2 – divergence principle applies to both the discrete and continuous random variables. Kumar and Taneja [2004] defined the minimum χ^2 –divergence probability distribution for continuous random variable as:

Definition 2.1: f(x) is the probability density of the minimum χ^2 – divergence continuous probability distribution of random variable X if it minimizes the χ^2 – divergence:

$$\chi^2(f,g)=\int_0^\infty \frac{f^2(x)}{g(x)}dx-1$$

That is, f(x) is the minimum χ^2 – divergence PDF of the random variable X if;

1. Probability density function constraints: f(x) ≥ 0, and ∫f(x) d x = 1,
2. a prior probability density function: g(x) ≥ 0, and ∫g(x) d x = 1,
3. It minimizes χ^2 (f, g) defined in (2.2), and

   Partial information in terms of averages: $\int_0^\infty [h(x)]^t f(x)\,dx=\mu'_{h(X),t,f}$

Where $h(x)$ is any real-valued function of x and $t=0,1,2,\ldots,r$.

Lemma 2.1: Let the given prior PDF be g(x) and the constraints are:

$$f(x)\ge 0,\ \int_0^\infty f(x)\,dx=1,\ \int_0^\infty [h(x)]^t f(x)\,dx=\mu'_{h(X),t,f} \tag{2.2}$$

Then the minimum χ^2 – divergence probability distribution of the random variable x has the PDF;

$$f(x)=\frac{g(x)}{2}\left(\alpha_0+\sum_{t=1}^{r}\alpha_t[h(x)]^t\right). \tag{2.3}$$

The coefficients $\alpha_t, t = 0,1,2,\dots,r$ are determined from;

$$\int_0^\infty \frac{g(x)}{2}\left(\alpha_0 + \sum_{t=1}^{r}\alpha_t[h(x)]^t\right)dx = 1, \tag{2.4}$$

$$\int_0^\infty [h(x)]^t \frac{g(x)}{2}\left(\alpha_0 + \sum_{t=1}^{r}\alpha_t[h(x)]^t\right)dx = \mu'_{h(X),t,f}. \tag{2.5}$$

The minimum χ^2 – divergence measure is:

$$\chi^2_{\min}(f,g) = \int_0^\infty \frac{g(x)}{4}\left(\alpha_0 + \sum_{t=1}^{r}\alpha_t[h(x)]^t\right)^2 dx - 1. \tag{2.6}$$

3. EXPONENTIATED EXPONENTIAL DISTRIBUTION

The EExpdistribution is a member of the exponentiatedweibull (EW) distribution, see Mudholkar and Srivastava [1993]. In GuptaandKundu [2001], the two-parameter EExp distribution is quite effectively, used in analyzing lifetime's data, particularly in place of two-parameter gamma or two-parameter weibull distribution. If the shape parameter θ=1, then all the three distributions coincide with the one-parameter exponential distribution. Therefore, all the three distributions are generalizations of the exponential distribution in different ways. The PDF of the EExp distribution has different shapes. It is unimodal for θ>1, and is reversed J-shaped for θ≤1. It is log-convex if θ<1 and log-concave if θ>1. The hazard function is non-decreasing when θ>1, non-increasing when θ<1, and constant when θ=1, see Nadarajah [2011].

3.1 The Available information is the n^{th} moment

For the case, when the prior is the EExp (θ, λ) distribution and the available information is $\mu'_{n,f}$, n^{th} non-central moment of f(x), we have;

Theorem 3.1: Suppose that the observed estimated prior is the exponentiated exponential distribution with parameters θ and λ and the available information be;

$$g(x;\theta,\lambda) = \lambda\theta\exp(-\lambda x)\left(1-\exp(-\lambda x)\right)^{\theta-1}, x>0, \theta>0, \lambda>0 \tag{3.1}$$

$$f(x)\ge 0, \int_0^\infty f(x)dx = 1, E(X^n) = \int_0^\infty x^n f(x)dx = \mu'_{n,f} \tag{3.2}$$

Then the corresponding minimum χ^2 – divergence PDF f(x) is:

$$f(x;\theta,\lambda) = w_1(n)\,g(x;\theta,\lambda) + w_2(n)\,g_n(x;\theta,\lambda), \tag{3.3}$$

$$\mu'_{n,f} = t\,\mu'_{n,g} + (1-t)\frac{\mu'_{2n,g}}{\mu'_{n,g}},\ \ 0\le t\le 1, \tag{3.4}$$

$$w_1(n) = \frac{1}{\sigma^2_{n,g}}\left(\mu'_{2n,g} - \mu'_{n,f}\,\mu'_{n,g}\right), \tag{3.5}$$

$$w_2(n) = \frac{1}{\sigma^2_{n,g}}\left(\mu'_{n,f} - \mu'_{n,g}\right)\mu'_{n,g}, \tag{3.6}$$

$$\chi^2_{\min}(f,g) = \frac{\left(\mu'_{n,f}-\mu'_{n,g}\right)^2}{\mu'_{2n,g}-\mu'^2_{n,g}} = \frac{\left(\mu'_{n,f}-\mu'_{n,g}\right)^2}{\sigma^2_{n,g}}, \tag{3.7}$$

For non-negative integral values of (θ – 1).

$$\mu'_{k,g} = E(X^k) = \frac{\theta}{\lambda^k}\sum_{i=0}^{\theta-1}\frac{(-1)^i\,\Gamma(\theta)\,k!}{\Gamma(i+1)\Gamma(\theta-i)(i+1)^{k+1}}, k=1,2,\dots \tag{3.8}$$

$$g_n(x;\theta,\lambda) = \frac{1}{\mu'_{n,g}}x^n\,\lambda\theta\exp(-\lambda x)\left(1-\exp(-\lambda x)\right)^{\theta-1}, x>0, \theta,\lambda>0. \tag{3.9}$$

Proof: Eq. (3.8) implies:

$$\mu'_{k,g} = \int_0^\infty x^k \lambda\theta e^{-\lambda x}(1-e^{-\lambda x})^{\theta-1} dx$$

$$= \int_0^\infty \lambda\theta x^k e^{-\lambda x} \sum_{i=0}^{\theta-1} {}^{\theta-1}C_i \, (-e^{-\lambda x})^i \, dx$$

$$= \lambda\theta \sum_{i=0}^{\theta-1} (-1)^i \; {}^{\theta-1}C_i \int_0^\infty x^k e^{-\lambda x} e^{-\lambda i x} dx$$

$$= \lambda\theta \sum_{i=0}^{\theta-1} (-1)^i \cdot \frac{(\theta-1)!}{i!(\theta-i-1)!} \int_0^\infty x^k e^{-(i+1)\lambda x} dx$$

$$\text{Let } (i+1)\lambda x = y$$

$$\Rightarrow x = \frac{y}{(i+1)\lambda}, \quad dx = \frac{dy}{(i+1)\lambda}$$

$$\Rightarrow \mu'_{k,g} = \lambda\theta \sum_{i=0}^{\theta-1} (-1)^i \frac{(\theta-1)!}{i!(\theta-i-1)!} \int_0^\infty \frac{y^k}{(i+1)^k \lambda^k} e^{-y} \frac{dy}{\lambda(i+1)}$$

$$= \lambda\theta \sum_{i=0}^{\theta-1} (-1)^i \cdot \frac{(\theta-1)!}{i!(\theta-i-1)!} \cdot \frac{1}{(i+1)^{k+1}\lambda^{k+1}} \cdot \Gamma(k+1)$$

$$= \frac{\theta}{\lambda^k} \sum_{i=0}^{\theta-1} (-1)^i \frac{\Gamma(\theta)\, k!}{\Gamma(i+1)\Gamma(\theta-1)(i+1)^{k+1}}$$

Otherwise, the sum runs to ∞. We have the following corollaries:

Corollary 3.1: When t = 1, equation (3.4) implies $\mu'_{n,f} = \mu'_{n,g}$, $w_1(n) = 1$, $w_2(n) = 0$, and f(x; θ, λ) = g (x; θ, λ) with zero minimum χ^2 – divergence measure, see Atallah (2009).

Corollary 3.2: When t = 0, equation (3.4) implies $\mu'_{n,f} = \mu'_{2n,g} / \mu'_{n,g}$, $w_1(n) = 0$ and $w_2(n) = 1$. Hence, f(x; θ, λ) = g_n(x; θ, λ) defined in equation (3.9), with $\chi^2_{\min}(f,g) = \sigma^2_{n,g} / \mu'^2_{n,g}$.

Corollary 3.3: For 0< t <1, $\mu'_{n,f} = t\,\mu'_{n,g} + (1-t)\,\mu'_{2n,g} / \mu'_{n,g}$, and f(x) is the proper mixture described in equation (3.3), i.e. $w_1(n) \neq 0$, $w_2(n) \neq 0$ and sum to 1.

For example, t = ½, $\mu'_{n,f} = \frac{1}{2}\left[\mu'_{n,g} + \mu'_{2n,g} / \mu'_{n,g}\right]$, $w_1(n) = ½$, and $w_2(n) = ½$.

$$f(x;\theta,\lambda) = \frac{1}{2} g(x;\theta,\lambda) + \frac{1}{2} g_n(x;\theta,\lambda), \tag{3.10}$$

$$\chi^2_{\min}(f,g) = \frac{\sigma^2_{n,g}}{4\mu'^2_{n,g}}.$$

3.1.1 Some properties of the minimum χ^2–divergence:

Property 3.1: The r^{th} non-central moment, mean, and variance of f(x) are;

$$\mu'_{r,f} = A(n)\mu'_{r,g} + B(n)\mu'_{n+r,g}, \tag{3.11}$$

$$\mu'_{1,f} = A(n)\mu'_{1,g} + B(n)\mu'_{n+1,g}, \tag{3.12}$$

$$\sigma^2_f = A(n)\mu'_{2,g} + B(n)\mu'_{n+2,g} - \mu'^2_{1,f}. \tag{3.13}$$

Property 3.2: The cumulative distribution function of f(x) has the form:

$$F(x) = w_1(n)G(x) + w_2(n)G_n(x). \tag{3.14}$$

The functions G(x) and G_n (x) are the CDFs of g(x; θ, λ) and g_n(x; θ, λ) respectively. That is;

$$G(x) = \left(1 - \exp(-\lambda x)\right)^{\theta}, \tag{3.15}$$

$$G_n(x) = \int_0^x \frac{1}{m_{n,g}} u^n \lambda\theta \exp(-\lambda u)\left(1 - \exp(-\lambda u)\right)^{\theta - 1} du, x > 0 \tag{3.16}$$

Property 3.3: The Survival Function of f(x) is;

$$S(x) = w_1(n)[1 - G(x)] + w_2(n)[1 - G_n(x)] \tag{3.17}$$

Property 3.4: The hazard function of f(x) is;

$$h(x) = \frac{w_1(n)g(x;\theta,\lambda) + w_2(n)g_n(x;\theta,\lambda)}{w_1(n)[1 - G(x)] + w_2(n)[1 - G_n(x)]} \tag{3.18}$$

3.2 Available information is the first moment

For the special case n = 1, **Theorem 3.1** reduces to:

Theorem 3.2: Suppose that the observed estimated prior is the EExp distribution with parameters θ and λ, given in Eq. (3.1).Let the available information be:

$$f(x) \geq 0, \int_0^\infty f(x)dx = 1, E(X) = \int_0^\infty x f(x)dx = \mu'_{1,f}, \tag{3.19}$$

The minimum χ^2 –divergence PDF of X is;

$$f(x;\theta,\lambda) = [A(1) + B(1)x]g(x;\theta,\lambda)$$

$$f(x;\theta,\lambda) = w_1(1)g(x;\theta,\lambda) + w_2(1)g_1(x;\theta,\lambda) \tag{3.20}$$

$$\mu'_{1,f} = t\mu'_{1,g} + (1-t)\frac{\mu'_{2,g}}{\mu'_{1,g}}, \quad 0 \leq t \leq 1. \tag{3.21}$$

The minimum χ^2 – divergence measure between f(x) and g(x) is;

$$\chi^2_{\min}(f,g) = \frac{\left(\mu'_{1,f} - \mu'_{1,g}\right)^2}{\sigma^2_{1,g}}. \tag{3.22}$$

Eq. (3.20) expresses the function f(x; θ, λ) as a weighted mixture of two PDFs g(x; θ, λ), and g_1(x; θ, λ) with respective weights w_1 (1) and w_2 (1), see Bairamov and Özkal (2007).

The quantities A (1), B (1), w_1 (1), w_2 (1), and g_1 (x; θ, λ) are determined by substituting n = 1 into equations (3.5), (3.6), (3.9) respectively. From the results of theorem 3.2, we have;

Corollary 3.4: For t = 1, equation (3.21) implies $\mu'_{1,f} = \mu'_{1,g}$, w_1 (1) = 1 and w_2 (1) = 0. Hence, f(x; θ, λ) = g(x; θ, λ) with zero minimum χ^2 – divergence measure between f(x; θ, λ) and g(x; θ, λ), see Ord (1975).

Corollary 3.5: When t = 0, equation (3.21) implies $\mu'_{1,f} = \mu'_{2,g} / \mu'_{1,g}$, w_1 (1) = 0 and w_2 (1) = 1. Hence,

f(x; θ, λ) = g_1 (x; θ, λ) and $\chi^2_{\min}(f,g) = \sigma^2_{1,g} \big/ \mu'^2_{1,g}$.

Corollary 3.6: When 0< t <1, $\mu'_{1,f} = t\mu'_{1,g} + (1-t)\mu'_{2,g} / \mu'_{1,g}$, and f(x) is the proper mixture, given by equation (3.20).

For example, when t = ½, $\mu'_{1,f} = t\mu'_{1,g} + (1-t)\mu'_{2,g} / \mu'_{1,g}$ w_1 (1) = ½ and w_2 (1) = ½.

$$f(x;\theta,\lambda) = \frac{1}{2}g(x;\theta,\lambda) + \frac{1}{2}g_1(x;\theta,\lambda), \tag{3.23}$$

$$\chi^2_{\min}(f,g) = \frac{\sigma^2_{1,g}}{4\mu'^2_{1,g}}.$$

3.2.1 Some properties of the minimum χ^2–divergence in this moment:

Property 3.5: The r^{th} non-central moment, mean, and variance of f(x; θ, λ) are;

$$\mu'_{r,f} = A(1)\mu'_{r,g} + B(1)\mu'_{r+1,g}\ , \tag{3.24}$$

$$\mu'_{1,f} = A(1)\mu'_{1,g} + B(1)\mu'_{2,g}\ , \tag{3.25}$$

$$\sigma^2_f = A(1)\mu'_{2,g} + B(1)\mu'_{3,g} - \mu'^2_{1,f}\ . \tag{3.26}$$

Property 3.6: The cumulative distribution function has the form;

$$F(x) = w_1(1)G(x) + w_2(1)G_1(x). \tag{3.27}$$

Where, G(x) and G_1(x) are the CDFs of the PDFs: g(x; θ, λ) and g_1(x; θ, λ), respectively.

Property 3.7: The Survival Function of f(x; θ, λ) is;

$$S(x) = w_1(1)[1 - G(x)] + w_2(1)[1 - G_1(x)]\cdot \tag{3.28}$$

Property 3.8: The hazard function of f(x; θ, λ) is;

$$h(x) = \frac{w_1(1)g(x;\theta,\lambda) + w_2(1)g_1(x;\theta,\lambda)}{w_1(1)[1-G(x)] + w_2(1)[1-G_1(x)]}\cdot \tag{3.29}$$

3.3 Available information is the second moment

For the special case n = 2, Theorem 3.1 reduces to:

Theorem 3.3: Suppose that the observed estimated prior is the EExp distribution with parameters θ and λ, given in Eq. (3.1). Let the available information be:

$$f(x) \ge 0, \int_0^\infty f(x)dx = 1,\ E(X^2) = \int_0^\infty x^2 f(x)dx = \mu'_{2,f}\ . \tag{3.30}$$

The minimum χ^2 – divergence PDF f(x) is;

$$f(x;\theta,\lambda) = [A(2) + B(2)x^2]g(x;\theta,\lambda),$$

$$f(x;\theta,\lambda) = w_1(2)g(x;\theta,\lambda) + w_2(2)g_2(x;\theta,\lambda), \tag{3.31}$$

$$\mu'_{2,f} = t\ \mu'_{2,g} + (1-t)\frac{\mu'_{4,g}}{\mu'_{2,g}},\ \ 0 \le t \le 1. \tag{3.32}$$

The minimum χ^2 – divergence measure between f(x) and g(x) is;

$$\chi^2_{\min}(f,g) = \frac{\left(\mu'_{2,f} - \mu'_{2,g}\right)^2}{\mu'_{4,g} - \mu'^2_{2,g}} = \frac{\left(\mu'_{2,f} - \mu'_{2,g}\right)^2}{\sigma^2_{2,g}}, \tag{3.33}$$

Eq. (3.31) expresses the PDF f(x; θ, λ) as a weighted mixture of the two PDFs g(x; θ, λ) and g_2 (x; θ, λ) with respective weights w_1 (2) and w_2 (2) see Gupta et al. (1997).

The quantities A (2), B (2), w_1 (2), w_2 (2), and g_2 (x; θ, λ) are determined by putting n=2 into equations (3.5), (3.6), and (3.9) respectively. Some Properties of the minimum χ2–divergence PDF f(x; θ, λ)

Corollary 3.7: For t = 1, equation (3.32) implies $\mu'_{2,f} = \mu'_{2,g}$, w_1 (2) = 1 and w_2 (2) = 0. Hence, f(x; θ, λ) = g(x; θ, λ) with zero minimum χ^2 – divergence measure between f(x; θ, λ) and g(x; θ, λ).

Corollary 3.8: When t = 0, equation (3.32) implies $\mu'_{2,f} = \mu'_{4,g} / \mu'_{2,g}$, w_1 (2) = 0, and w_2 (2) = 1. Hence, f(x) = g_2(x) and $\chi^2_{\min}(f,g) = \sigma^2_{2,g} \big/ \mu'^2_{2,g}$.

Corollary 3.9: When 0< t <1, $\mu'_{2,f} = t\mu'_{2,g} + (1-t)\mu'_{4,g} / \mu'_{2,g}$, and f(x; θ, λ) is the proper mixture, given by Eq. (3.31),see Khan et al.(2009).

For example, when t = ½: $\mu'_{2,f} = \frac{1}{2}\left[\mu'_{2,g} + \mu'_{4,g}/\mu'_{2,g}\right]$, $w_1(2) = ½$ and $w_2(2) = ½$,

$$f(x;\theta,\lambda) = \frac{1}{2}g(x;\theta,\lambda) + \frac{1}{2}g_2(x;\theta,\lambda), \tag{3.34}$$

$$\chi^2_{\min}(f,g) = \frac{\sigma^2_{2,g}}{4\mu'^2_{2,g}}.$$

3.3.1 Some properties of the minimum χ^2–divergence in this moment

Property 3.9: The r^{th} non-central moment, mean, and variance of f(x; θ, λ) are;

$$\mu'_{r,f} = A(2)\mu'_{r,g} + B(2)\mu'_{r+2,g}, \tag{3.35}$$

$$\mu'_{1,f} = A(2)\mu'_{1,g} + B(2)\mu'_{3,g}, \tag{3.36}$$

$$\sigma^2_f = A(2)m_{2,g} + B(2)m_{4,g} - \mu'^2_{1,f}, \tag{3.37}$$

Property 3.10: The cumulative distribution function has the form;

$$F(x) = w_1(2)G(x) + w_2(2)G_2(x), \tag{3.38}$$

The PDFs g(x; θ, λ) and g_2(x; θ, λ) have CDFs G(x) and G_2(x) respectively.

Property 3.11: The Survival Function of f(x; θ, λ) is;

$$S(x) = w_1(2)[1-G(x)] + w_2(2)[1-G_2(x)], \tag{3.39}$$

Property 3.12: The hazard function of f(x; θ, λ) is;

$$h(x) = \frac{w_1(2)g(x;\theta,\lambda) + w_2(2)g_2(x;\theta,\lambda)}{w_1(2)[1-G(x)] + w_2(2)[1-G_2(x)]}, \tag{3.40}$$

3.4 Available information is the first two moments

When the prior distribution is the EExpdistribution with parameters θ and λ, and the available information is $\mu'_{k,f}, k = 1,2$ .We have;

Theorem 3.4: Suppose that the observed estimated prior is the EExpwith parameters θ, λ. Let the available information be:

$$E(X) = \int_0^\infty x f(x)\,dx = \mu'_{1,f},\ E(X^2) = \int_0^\infty x^2 f(x)\,dx = \mu'_{2,f}, \tag{3.41}$$

$$f(x) \geq 0, \text{ and } \int_0^\infty f(x)\,dx = 1$$

The minimum χ^2 – divergence PDF f(x; θ, λ) is;

$$f(x;\ \theta,\ \lambda) = (A + Bx + Cx^2)g(x;\theta,\lambda).$$

$$f(x;\ \theta,\ \lambda) = w_1 g(x;\theta,\lambda) + w_2 g_1(x;\theta,\lambda) + w_3 g_2(x;\theta,\lambda), \tag{3.42}$$

$$A = \frac{\mu'_{1,f}(\mu'_{2,g}\mu'_{3,g} - \mu'_{1,g}\mu'_{4,g}) + \mu'_{2,f}(\mu'_{1,g}\mu'_{3,g} - \mu'^2_{2,g}) + \mu'_{2,g}\mu'_{4,g} - \mu'^2_{3,g}}{\mu'_{2,g}\mu'_{4,g} + 2\mu'_{1,g}\mu'_{2,g}\mu'_{3,g} - \mu'^2_{1,g}\mu'_{4,g} - \mu'^2_{3,g} - \mu'^3_{2,g}}, \tag{3.43}$$

$$B = \frac{\mu'_{1,f}(\mu'_{4,g} - \mu'^2_{2,g}) + \mu'_{2,f}(\mu'_{1,g}\mu'_{2,g} - \mu'_{3,g}) + \mu'_{2,g}\mu'_{3,g} - \mu'_{1,g}\mu'_{4,g}}{\mu'_{2,g}\mu'_{4,g} + 2\mu'_{1,g}\mu'_{2,g}\mu'_{3,g} - \mu'^2_{1,g}\mu'_{4,g} - \mu'^2_{3,g} - \mu'^3_{2,g}}, \tag{3.44}$$

$$C = \frac{\mu'_{1,f}(\mu'_{1,g}\mu'_{2,g} - \mu'_{3,g}) + \mu'_{2,f}(\mu'_{2,g} - \mu'^2_{1,g}) + \mu'_{1,g}\mu'_{3,g} - \mu'^2_{2,g}}{\mu'_{2,g}\mu'_{4,g} + 2\mu'_{1,g}\mu'_{2,g}\mu'_{3,g} - \mu'^2_{1,g}\mu'_{4,g} - \mu'^2_{3,g} - \mu'^3_{2,g}}, \tag{3.45}$$

$$w_1 = A,\ w_2 = B\mu'_{1,g},\ w_3 = C\mu'_{2,g} \tag{3.46}$$

Eq. (3.42) expresses the PDF f(x; θ, λ) as a weighted mixture of the three PDFs g(x; θ, λ), g_1(x; θ, λ), and g_2(x; θ, λ) with respective weights w_1, w_2, and w_3. Notice that, $w_1 + w_2 + w_3 = 1$,

$$g_i(x;\theta,\lambda) = \frac{1}{\mu'_{i,g}} x^i \lambda\theta \exp(-\lambda x)(1-\exp(-\lambda x))^{\theta-1}, x > 0, i = 1,2 \tag{3.47}$$

The minimum χ^2-divergence measure between f(x; θ, λ) and g(x; θ, λ) is;

$$\chi^2_{\min}(f,g) = A^2 + B^2\mu'_{2,g} + C^2\mu'_{4,g} + 2AB\mu'_{1,g} + 2AC\mu'_{2,g} + 2BC\mu'_{3,g} - 1, \qquad (3.48)$$

Where A, B, and C are as given in equations (3.43) – (3.45). We have the following corollary;

Corollary 3.10: When $\mu'_{k,f}, k = 1,2$, and $\mu'_{k,f}, k = 1,2$, $w_1 = 1$, $w_2 = 0$, and $w_3 = 0$. Therefore, f(x; θ, λ) = g(x; θ, λ) and the minimum χ^2– divergence measure between f(x) and g(x; θ, λ) equals zero. Some Properties of the minimum χ^2–divergence PDF f(x; θ, λ), see Gertsbakh and Kagan (1999).

3.4.1 Some properties of the minimum χ^2–divergence in this moment

Property 3.13: The r^{th} non-central moment, mean, and variance of f(x) are;

$$\mu'_{r,f} = A\mu'_{r,g} + B\mu'_{r+1,g} + C\mu'_{r+2,g} \qquad (3.49)$$

$$\mu'_{1,f} = A\mu'_{1,g} + B\mu'_{2,g} + C\mu'_{3,g}, \qquad (3.50)$$

$$\sigma^2_f = A\mu'_{2,g} + B\mu'_{3,g} + C\mu'_{4,g} - \mu'^2_{1,f}. \qquad (3.51)$$

Property 3.14: The cumulative distribution function of f(x; θ, λ) has the form

$$F(x) = w_1G(x) + w_2G_1(x) + w_3G_2(x). \qquad (3.52)$$

Where G(x), G_1(x), and G_3(x) are respectively, the CDFs of g(x; θ, λ), g_1(x; θ, λ), and g_2(x; θ, λ) defined in Eq. (3.47).

Property 3.15: The Survival Function of f(x; θ, λ) is;

$$S(x) = w_1[1-G(x)] + w_2[1-G_1(x)] + w_3[1-G_2(x)]. \qquad (3.53)$$

Property 3.16: The hazard function of f(x; θ, λ) is;

$$h(x) = \frac{w_1\, g(x;\theta,\lambda) + w_2\, g_1(x;\theta,\lambda) + w_3\, g_2(x;\theta,\lambda)}{w_1[1-G(x)] + w_2[1-G_1(x)] + w_3[1-G_2(x)]}. \qquad (3.54)$$

3.5 Available information is the first three moments

When the prior is the EExp distribution with parameters θ and λ and the available information is the first three moments we have;

Theorem 3.5: Suppose that the observed estimated prior is the EExp distribution with a PDFg(x; θ, λ). Let the available information be;

$$f(x) \geq 0, \int_0^\infty f(x)\,dx = 1, \int_0^\infty x^k f(x)\,dx = \mu'_{k,f}, k = 1,2,3, \qquad (3.55)$$

The minimum χ^2 - divergence PDF f(x) is;

$$f(x;\ \theta,\ \lambda) = (A + Bx + Cx^2 + Dx^3)g(x;\theta,\lambda).$$

Equivalently,

$$f(x;\ \theta,\ \lambda) = w_1\, g(x;\theta,\lambda) + w_2\, g_1(x;\theta,\lambda) + w_3\, g_2(x;\theta,\lambda) + w_4\, g_3(x;\theta,\lambda), \qquad (3.56)$$

Where g(x; θ, λ), g_1 (x; θ, λ), g_2 (x; θ, λ), g_3 (x; θ, λ) are given by equations (3.1) – (3.9). The weights w_i, i = 1, 2, 3, 4 can be obtained from equation (2.27), see Tavangar and Asadi (2010).

4. ILLUSTRATIVE EXAMPLES

4.1 The first moment is available

Consider the case when the prior estimated PDF of f(x) is the EExp (θ, λ) = EExp in Eq. (2, 1) and the available information is E(X). In this case:

$$g(x) = 2e^{-x}(1-e^{-x}), x > 0$$

$$f(x) = w_1(1)g(x) + w_2(1)g_1(x)$$

$$\mu'_{1,f} \in [1.5,\ 2.33].$$

We have three possibilities, the first possibility in case (1), the second possibility in case (2), and third possibility in cases (3) – (6).

Table 1, displays the different cases of Prior is EExp distribution, According to the values of the first moment $\mu'_{1,f}$, the parent distribution is determined. The PDF f (x) of the parent distribution is characterized as a single PDF in cases (1) and (2), while in cases (3) – (6) it is a proper weighted binary mixture of two PDFs g(x) and g_1(x). The values in the final column, indicates how close are the different forms of f (x) and the prior estimated PDF g (x). As can be seen from Figure 1, Figure 2.

Table-1: Prior is EExp and the available is $\mu'_{1,f}$

Case	$\mu'_{1,f}$	w_1	w_2	f(x)	$\chi^2_{\min}(f,g)$
1	1.5	1	0	$f_1(x) = g(x)$	0
2	2.33	0	1	$f_2(x) = g(x)$	0.556
3	1.917	0.5	0.5	$f_3(x) = 0.5\,g(x) + 0.5\,g_1(x)$	0.139
4	2.125	0.25	0.75	$f_4(x) = 0.25\,g(x) + 0.75\,g_1(x)$	0.313
5	1.738	0.714	0.286	$f_5(x) = 0.714\,g(x) + 0.286\,g_1(x)$	0.045
6	2.056	0.333	0.667	$f_6(x) = 0.333\,g(x) + 0.667\,g_1(x)$	0.247

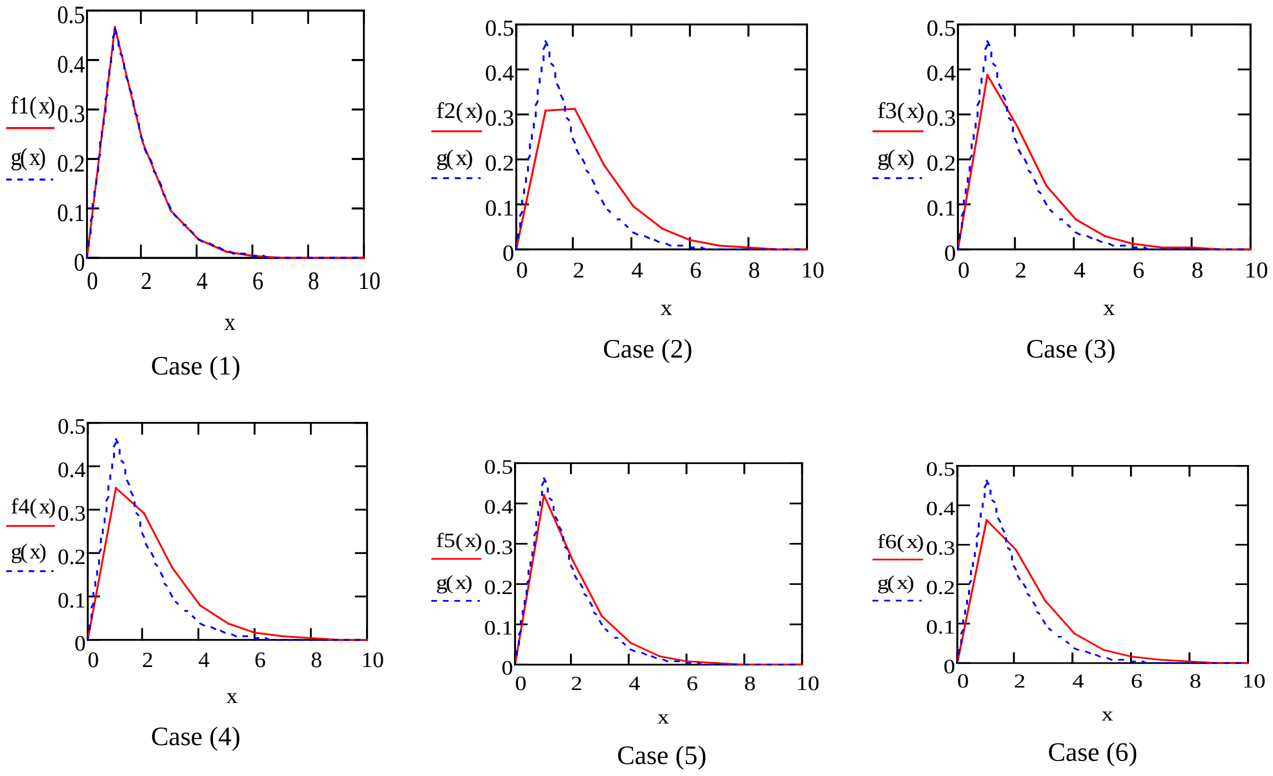
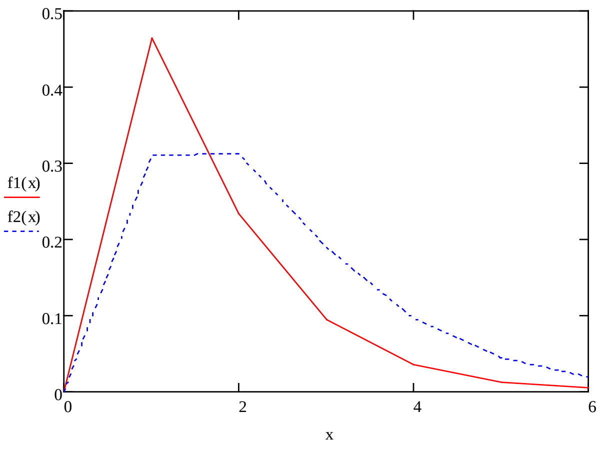
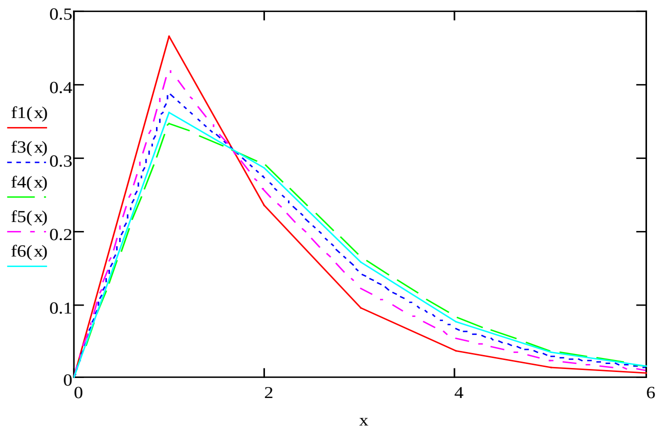


Figure-1: RepresentPDFs with prior EExp in cases (1) – (6).



Cases (1) – (2)



Cases (3) – (6)

Figure-2: PDFs with prior EExp when $\mu'_{1,f}$ is available in all cases, collectively.

4.2 The second moment is available

In this case, we have

$$f(x) = w_1(2)g(x) + w_2(2)g_2(x)$$

$$\mu'_{2,f} \in [3.5, 13.286]$$

We have three possibilities, the first possibility in case (1), the second possibility in case (2), and third possibility in cases (3) – (6).

Table 2, displays the different cases of Prior is Exp distribution, According to the values of the second moment $\mu'_{2,f}$, the parent distribution is determined. The PDF f(x) of the parent distribution is characterized as a single PDF in cases (1) and (2), while in cases (3) – (6) it is a proper weighted binary mixture of two PDFs g(x) and $g_2(x)$. The values in the final column, indicates how close are the different forms of f(x) and the prior estimated PDF g(x). As can be seen from Figure 3, Figure 4.

Table-2: Prior is EExp and the available is $\mu'_{2,f}$

Case	$\mu'_{2,f}$	w_1	w_2	f(x)	$\chi^2_{\min}(f,g)$
1	3.5	1	0	$f_1(x) = g(x)$	0
2	13.286	0	1	$f_2(x) = g(x)$	2.796
3	8.393	0.5	0.5	$f_3(x) = 0.5\,g(x) + 0.5\,g_1(x)$	0.699
4	10.839	0.25	0.75	$f_4(x) = 0.25\,g(x) + 0.75\,g_1(x)$	1.573
5	6.296	0.714	0.286	$f_5(x) = 0.714\,g(x) + 0.286\,g_1(x)$	0.228
6	10.024	0.333	0.667	$f_6(x) = 0.333\,g(x) + 0.667\,g_1(x)$	1.243

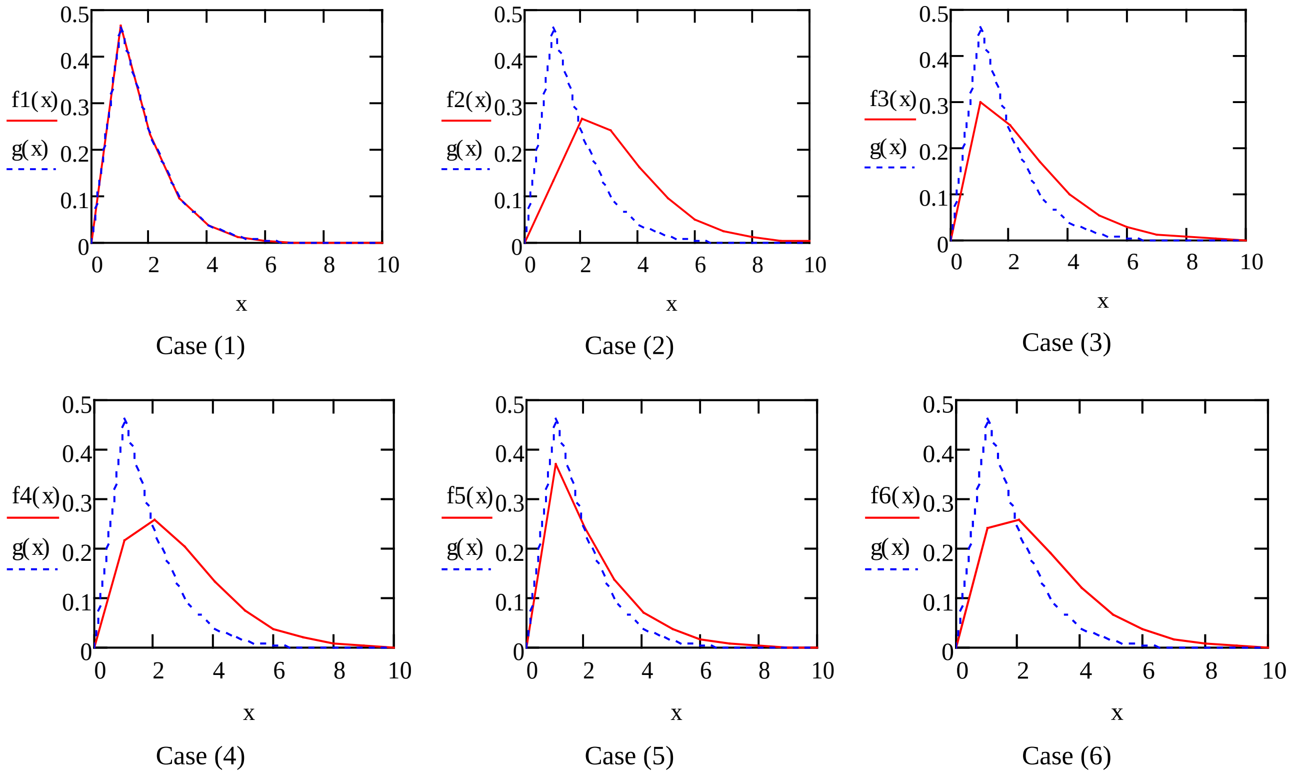
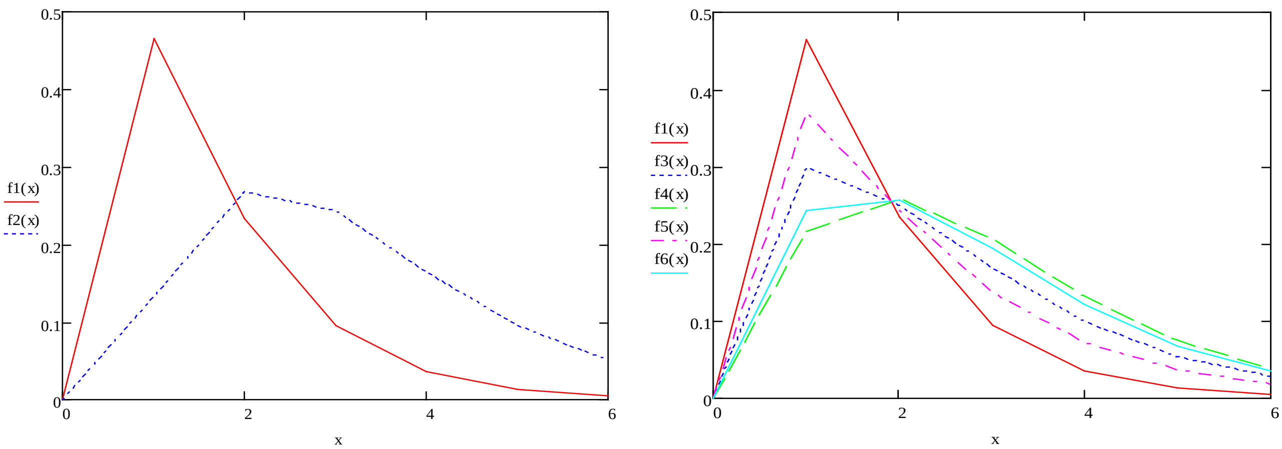


Figure-3: Represent PDFs with prior EExp in cases (1) – (6).



Cases (1) – (2) Cases (3) – (6)

Figure-4: PDFs with prior EExp when $\mu'_{2,f}$ is available in all cases, collectively.

4.3 The first two moments are available

In this case

$$f(x) = A\, g(x) + B\, g_1(x) + C\, g_2(x)$$

Or, equivalently

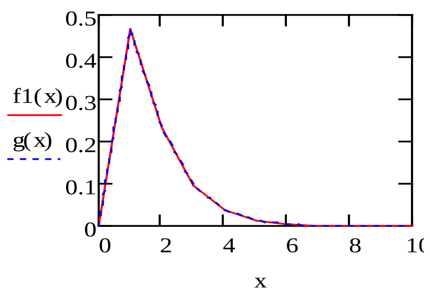
$$f(x) = w_1\, g(x) + w_2\, g_1(x) + w_3\, g_2(x),$$
$$\mu'_{1,f} \in [1.5, 2.33],$$
$$\mu'_{2,f} \in [3.5, 13.286],$$

where A, B, C, w_1, w_2, and w_3 are expressed in equations (3.43) – (3.46).

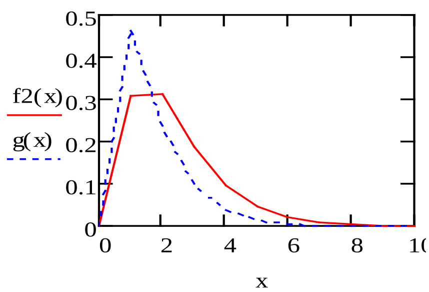
Table-3: Prior is EExp distribution and available is $\mu'_{k,f}, k=1,2$

Case	$\mu'_{1,f}$	$\mu'_{2,f}$	w_1	w_2	w_3	f(x)	$\chi^2_{\min}(f,g)$
1	1.5	3.5	1	0	0	$f_1(x) = g(x)$	0
2	2.333	7.5	0	1	0	$f_2(x) = g_1(x)$	0.556
3	1.917	5.5	0.5	0.5	0	$f_3(x) = 0.5g(x)+0.5g_1(x)$	0.139
4	2.056	6.167	0.333	0.667	0	$f_4(x) = 0.333g(x)+0.667g_1(x)$	0.247
5	2.071	6.762	0.667	0	0.333	$f_5(x) = 0.667g(x)+0.333g_2(x)$	0.311
6	2.01	6.257	0.333	0.333	0.333	$f_6(x) = 0.333g(x)+0.333g_1(x)+0.333g_2(x)$	0.225
7	1.711	4.666	0.851	0.051	0.098	$f_7(x) = 0.851g(x)+0.051g_1(x)+0.098g_2(x)$	0.04
8	1.7	4.605	0.858	0.048	0.093	$f_8(x) = 0.858g(x)+0.048g_1(x)+0.093g_2(x)$	0.034

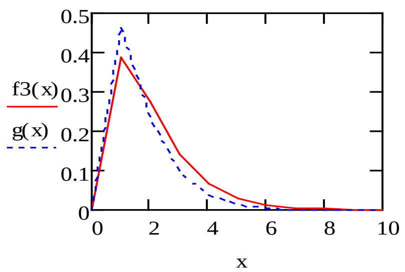
Table 3. Previous, displays the different cases of Prior is EExp distribution, According to the values of the first two moments $\mu'_{1,f}, \mu'_{2,f}$, the parent distribution is determined. The PDF f(x) of the parent distribution is characterized as a single PDF in cases (1) and (2). In cases (3) – (5), it is determined as a proper weighted binary mixture of two PDFs, while in cases (6) – (8), f(x) is determined as a proper weighted mixture of three PDFs. The values in the final column, indicates how close are the different forms of f(x) and the prior estimated PDF g(x). As can be seen from Figure 5, Figure 6.



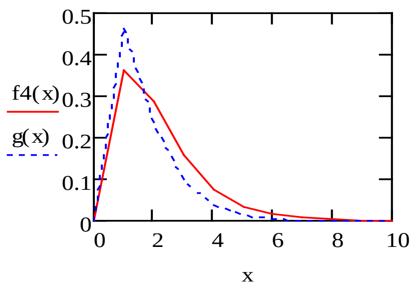
Case (1)



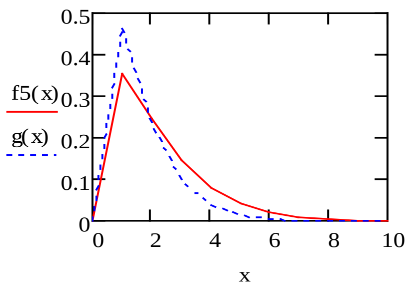
Case (2)



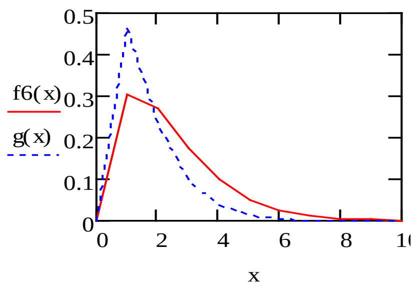
Case (3)



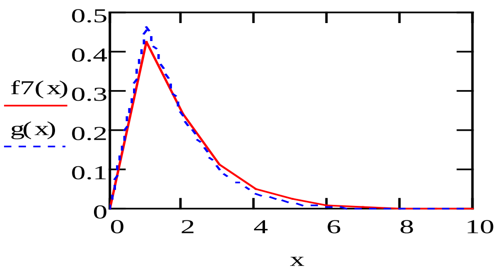
Case (4)



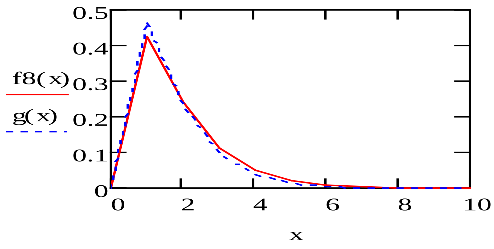
Case (5)



Case (6)

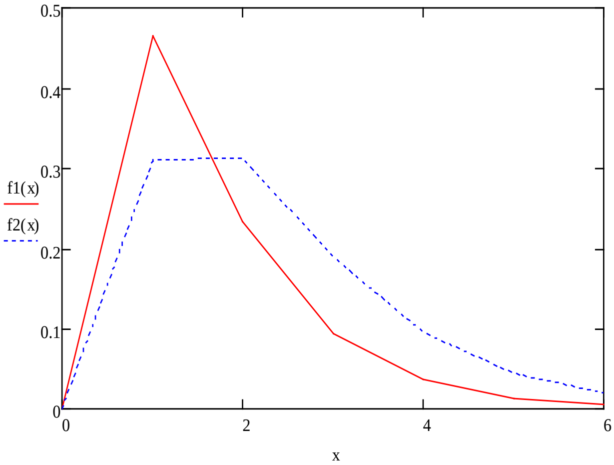


Case (7)

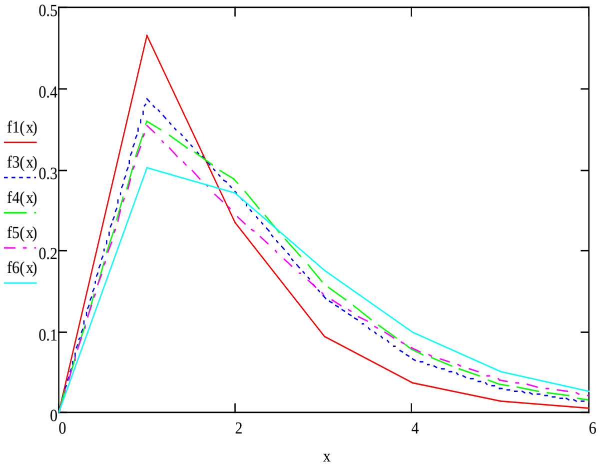


Case(8)

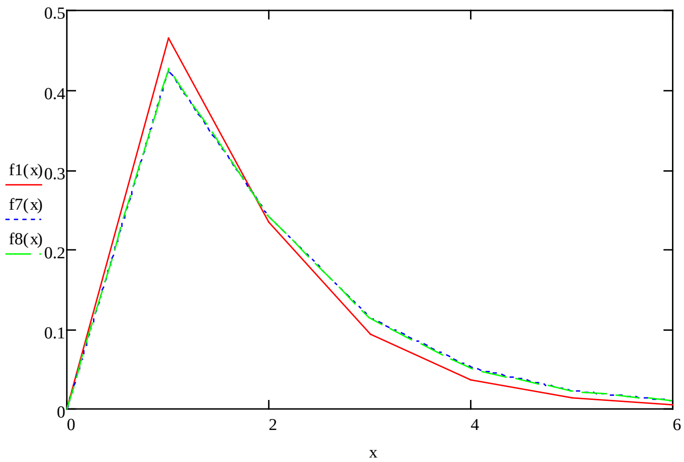
Figure-5: Represent PDFs with prior EExp in cases (1) – (8).



Cases (1) – (2)



Cases (3) – (6)



Cases (7) – (8)

Figure-6: PDFs with prior EExp, μ'_k, f, k $=1,2.$ is available in all cases, collectively.

4.4 The first three moments are available

In this case:

$$f(x) = A\,g(x) + B\,g_1(x) + C\,g_2(x) + D\,g_3(x)$$

Or, equivalently:

$$f(x) = w_1\,g(x) + w_2\,g_1(x) + w_3\,g_2(x) + w_4\,g_3(x)$$

where A, B, C, D, w_1, w_2, w_3, and w_4 are expressed in Eq. (3.56).

Table 4, displays the different cases of Prior is EExp distribution, According to the values of the first three moments $\mu'_{1,f}$, $\mu'_{2,f}$, $\mu'_{3,f}$, the parent distribution is determined. The PDF f(x) of the parent distribution is characterized as a single PDF in cases (1) and (2). In case (3), it is determined as a proper weighted binary mixture of two PDFs, while in case (4), f(x) is determined as a proper weighted mixture of three PDFs. The values in the final column, indicates how close are the different forms of f(x) and the prior estimated PDF g(x). As can be seen from Figure 7, Figure 8.

Table-4: Prior is EExp distribution and available is $\mu'_{k,f}; k = 1,2,3$

Case	$\mu'_{1,f}$	$\mu'_{2,f}$	$\mu'_{3,f}$	w_1	w_2	w_3	w_4	f(x)	$\chi^2_{\min}(f,g)$
1	1.5	3.5	11.25	0	0	0	1	$f_1(x) = g(x)$	0
2	2.333	7.5	31	0	0	1	0	$f_2(x) = g_1(x)$	0.556
3	1.917	5.5	21.125	0	0	0.5	0.5	$f_3(x) = 0.5g(x) + 0.5g_1(x)$	0.139
4	2.01	6.257	26.45	0	0.2	0.2	0.6	$f_4(x) = 0.6g(x) + 0.2g_1(x) + 0.2g_2(x)$	0.247

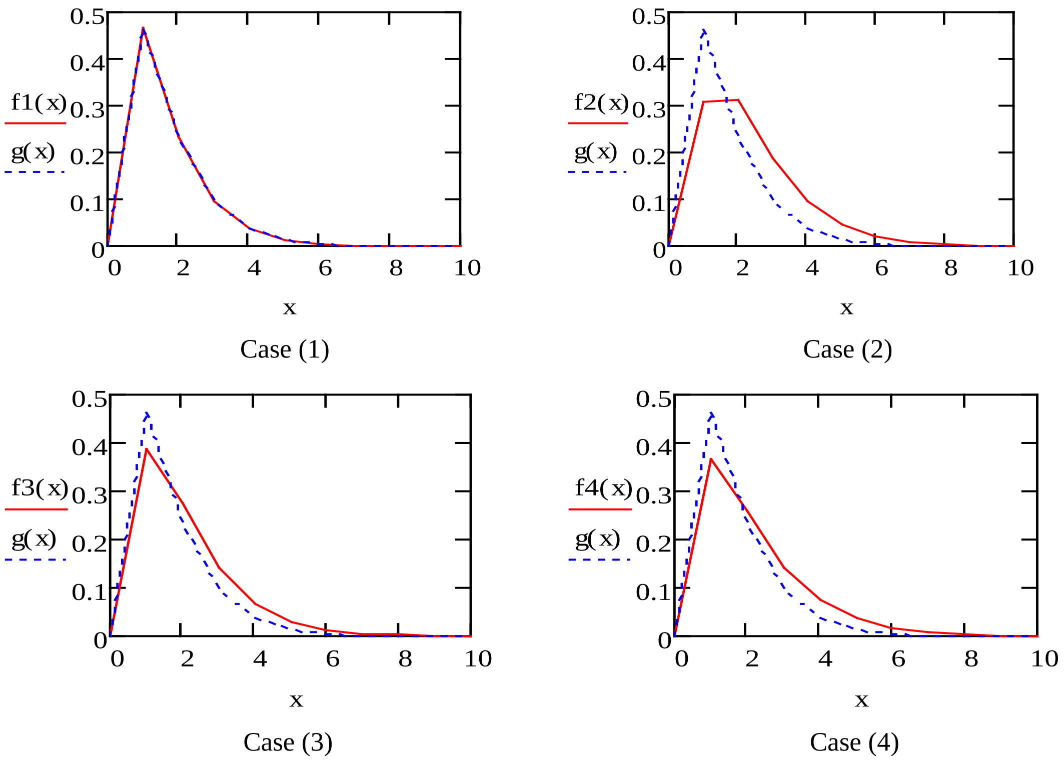
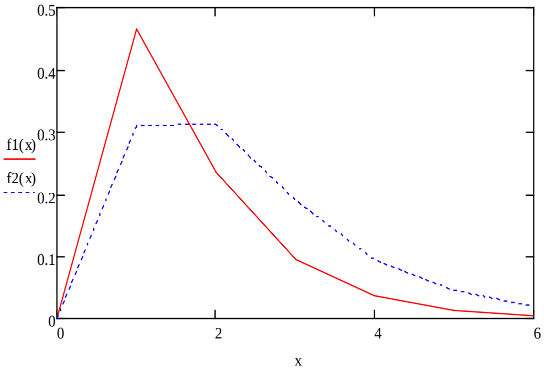
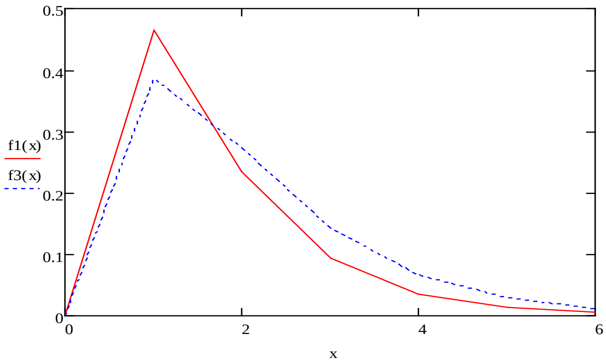


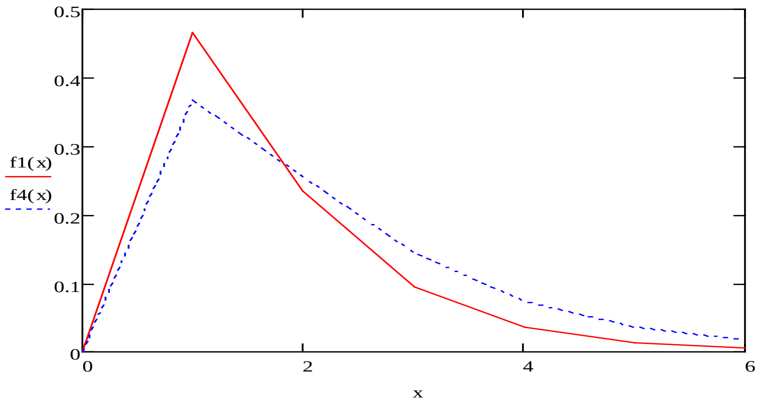
Figure-7: Represent PDFs with prior EExp in cases (1) – (4).



Cases (1) – (2)



Case (3)



Case (4)

Figure-8: PDFs with prior EExp μ'_k, f, $k=1,2,3$ is available inall cases, collectively.

5. CONCLUSION

This paper discusses the characterization of probability distributions conditioned on available a certain prior probability distribution and available information on moments. More precisely, we obtained characterizations when the prior is the EExp (θ, λ), and different available information on moments. Namely, the following available information on moments is considered for different cases, the first moment is available, the second moment is available, the first two moments are available, and the first three moments are available. As the special case EExp (1, λ). The properties of the EExp (θ, λ) are determined, given a prior distribution as EExp and the new (current) information on moments, and from these tables and figures, there is evidence that the new (minimum chi-square divergence) distribution, is not the EExp distribution. However according to the minimum cross-entropy principle the new distribution remains the same as the given prior distribution. Thus, given a prior information about the underlying distribution as EExp distribution, in addition to the partial information in terms of the moments, the minimum chi-square divergence principle provides a useful methodology for characterizing EExp probability distributions.

REFERENCES

1. Atallah, M. A. (2009). Erratum to:"Probability distributions conditioned by the available information: Gamma Distribution and Moments"[Comput. Math. Appl. 52 (2006) 289–304]. *Computers & Mathematics with Applications*, *58*(9), 1878-1886.
2. Ord, J. K. (1975). Characterization Problems in Mathematical Statistics. *Journal of the Royal Statistical Society: Series A (General)*, *138*(4), 576-577.
3. Bairamov, I., & Özkal, T. (2007). On characterization of distributions through the properties of conditional expectations of order statistics. *Communications in Statistics—Theory and Methods*, *36*(7), 1319-1326.
4. Campbell, L. L. (1970). Equivalence of Gauss's principle and minimum discrimination information estimation of probabilities. *The Annals of Mathematical Statistics*, *41*(3), 1011-1015.
5. Gertsbakh, I., & Kagan, A. (1999). Characterization of the Weibull distribution by properties of the Fisher information under type-I censoring. *Statistics & probability letters*, *42*(1), 99-105.
6. Gupta, A. K., Nguyen, T. T., & Zeng, W. B. (1997). Characterization of multivariate distributions through a functional equation of their characteristic functions. *Journal of statistical planning and inference*, *63*(2), 187-201.

7. Gupta, R. D., & Kundu, D. (2001). Exponentiated exponential family: an alternative to gamma and Weibull distributions. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, *43*(1), 117-130.
8. Hamedani, G., & Ahsanullah, M. (2011). Characterizations of Weibull geometric distribution. *Journal of Statistical Theory and Applications*, *10*(4), 581-590.
9. Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical review*, *106*(4), 620.
10. Kapur, J. N. (1982). Maximum-entropy probability distribution for a continuous random variate over a finite interval'. *J. Math. Phys. Sci*, *16*(1), 97-109.
11. Kapur, J. N. (1983). Twenty-five years of maximum-entropy principle. *Journal of Mathematical and Physical Sciences*, *17*, 103-156.
12. Kesavan, H., & Kapur, J. (1989). The generalized maximum entropy principle. *IEEE Transactions on systems, Man, and Cybernetics*, *19*(5), 1042-1052.
13. Kawamura, T., & Iwase, K. (2003). Characterizations of the distributions of power inverse Gaussian and others based on the entropy maximization principle. *Journal of the Japan Statistical Society*, *33*(1), 95-104.
14. Khan, A., Faizan, M., & Haque, Z. (2009). Characterization of probability distributions through order statistics. In *ProbStat Forum* (Vol. 2, pp. 132-136).
15. Kullback, S. (1997). *Information theory and statistics*. Courier Corporation. New York, John Wiley.
16. Kumar, P. (2006). Probability distributions conditioned by the available information: Gamma distribution and moments. *Computers & Mathematics with Applications*, *52*(3-4), 289-304.
17. Kumar, P. (2005). Characterizations of Beta Probability Distributions Based on the Minimum χ^2–Divergence Principle.
18. Mudholkar, G., &Srivastava, D. (1993). ExponentiatedWeibull family for analyzing bathtub failure-rate data. *IEEE transactions on reliability*, *42*(2), 299-302.
19. Shore, J., & Johnson, R. (1980). Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Transactions on information theory*, *26*(1), 26-37.
20. Tavangar, M., & Asadi, M. (2010). Some new characterization results on exponential and related distributions.
21. Zolotarev, V. M. (1976). Metric distances in spaces of random variables and their distributions. *Mathematics of the USSR-Sbornik*, *30*(3), 373.
22. Yanushkevichius, R. (2014). *Stability characterizations of some probability distributions*. LAP LAMBERT Academic Publishing.
23. Kumar, P., &Taneja, I. (2004). Chi square divergence and minimization problem. *J. Comb. Inf. Syst. Sci*, *28*, 181-207.
24. Nadarajah, S. (2011). The exponentiated exponential distribution: a survey, *AStA Advances in Statistical Analysis*, 95, 219 – 251.

Source of support: Nil, Conflict of interest: None Declared.