Characterizations of Beta Probability Distributions Based on the Minimum χ^2-Divergence Principle*

March 8, 2005

Pranesh Kumar
(*kumarp@unbc.ca*)
Mathematics Department
College of Science and Management
University of Northern British Columbia
Prince George BC V2N4Z9, Canada

Abstract

The minimum χ^2-divergence principle states: *When a prior probability density function of X , $g(x)$, which estimates the underlying probability density function $f(x)$ 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 χ^2-divergence* . We study minimun χ^2-divergence probability distributions based on this principle given a *prior* beta distribution and the partial information on moments.

Keywords: Optimization principle, Chi-square divergence, Characterization, Beta probability distributions.

MSC [2000]: 60E05; 62H05

1 Introduction

In this paper, we consider characterizations of beta probability distributions based on the minimum χ^2-divergence principle when given is: (i) a *prior* beta distribution and (ii) partial information in the form of averages.

2 Minimum χ^2-Divergence Probability Distributions

Let the random variable X be a continuous variable with probability density function $f(x)$ defined over the open interval $(-\infty, +\infty)$ or finite closed interval $[a, b]$. The minimum cross-entropy principle due to Kullback(1959) is:

When a prior probability density function of X , $g(x)$, which estimates the underlying probability density function $f(x)$ 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 Kullback - Leibler divergence

$$K(f,g) = \int f(x) \ln \frac{f(x)}{g(x)} dx. \tag{2.1}$$

*This research is supported by the author's Natural Sciences and Engineering Research Council Discovery Grant.

Kumar and Taneja (2004) has considered the minimum χ^2-divergence principle as:

When a prior probability density function of X , $g(x)$, which estimates the underlying probability density function $f(x)$ 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 χ^2-divergence

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

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*

$$\chi^2(f,g) = \int \frac{f^2(x)}{g(x)}dx - 1, \tag{2.3}$$

given:

(i) a prior probability density function: $g(x) \geq 0, \int g(x)dx$=1,
(ii) probability density function constraints: $f(x) \geq 0$, $\int f(x)dx$=1,and
(iii) partial information in terms of averages: $\int [h(x)]^t f(x)dx = m_{t,f}$, $t = 1, 2, 3, .., r$,
where $h(x)$ is any real-valued function of x.

Further, they proved:

Lemma 2.1 *Given a prior probability density function $g(x)$ of the continuous random variable X, and the constraints*

$$f(x) \geq 0, \int f(x)dx = 1, \int [h(x)]^t f(x)dx = m_{t,f}, \ t = 1, 2, 3, ..., r, \tag{2.4}$$

the minimum χ^2-divergence probability distribution of X has the probability density function

$$f(x) = \frac{g(x)}{2}\left(\alpha_0 + \sum_{t=1}^{r}[h(x)]^t\alpha_t\right), \tag{2.5}$$

and the $(r+1)$ constants, α_0 and α_t, $t = 1, 2, 3, ..., r$, are determined from

$$\int \frac{g(x)}{2}\left(\alpha_0 + \sum_{t=1}^{r}[h(x)]^t\alpha_k\right)dx = 1, \tag{2.6}$$

and

$$\int \frac{[h(x)]^t g(x)}{2}\left(\alpha_0 + \sum_{t=1}^{rt}[h(x)]^t\alpha_t\right)dx = m_{t,f}. \tag{2.7}$$

The minimum χ^2-divergence measure is:

$$\chi^2_{\min}(f,g) = \int \frac{g(x)}{4}\left(\alpha_0 + \sum_{t=1}^{r}[h(x)]^t\alpha_t\right)^2 dx - 1. \tag{2.8}$$

Given a prior *beta* probability distribution and the partial information on averages, we obtain the minimum χ^2-divergence probability distribution from Lemma 2.1 in the following theorem. In what follows henceforth, integral $\int$ is considered over $[0,1]$.

Theorem 2.1 *Let given be a prior as the beta probability distribution of X with the density function*

$$g(x) = \frac{x^{p-1}(1-x)^{q-1}}{B(p,q)} \ , p, q > 0, 0 < x < 1, \tag{2.9}$$

where

$$B(p,q) = \int_0^1 u^{p-1}(1-u)^{q-1} du, \tag{2.10}$$

and the constraints

$$f(x) \geq 0, \int f(x)dx = 1, \ \int [h(x)]^t f(x)dx = m_{t,f}, \ t = 1, 2, 3, ..., r. \tag{2.11}$$

Then, the minimum χ^2-divergence probability distribution of X has the density function

$$f(x) = \frac{x^{p-1}(1-x)^{q-1}}{2B(p,q)} \left(\alpha_0 + \sum_{t=1}^{r} [h(x)]^t \alpha_t \right), \tag{2.12}$$

and the $(r+1)$ constants, α_0 and α_t, $t = 1, 2, 3, ..., r$, are determined from

$$\int \frac{x^{p-1}(1-x)^{q-1}}{2B(p,q)} \left(\sum_{t=1}^{r} [h(x)]^t \alpha_k \right) dx = 1 - \frac{\alpha_0}{2}, \tag{2.13}$$

and

$$\int \frac{x^{p-1}(1-x)^{q-1}}{2B(p,q)} [h(x)]^t \left(\alpha_0 + \sum_{t=1}^{rt} [h(x)]^t \alpha_t \right) dx = m_{t,f}. \tag{2.14}$$

The minimum χ^2-divergence:

$$\chi^2_{\min}(f,g) = \int \frac{x^{p-1}(1-x)^{q-1}}{4B(p,q)} \left(\alpha_0 + \sum_{t=1}^{r} [h(x)]^t \alpha_t \right)^2 dx - 1. \tag{2.15}$$

Now we consider the applications of Theorem 2.1 when given is a prior *beta* probability distribution and the partial information on different types of averages.

3 Given a Prior Beta Distribution and Partial Information $E(X^n)$

Suppose that *a prior* beta density function $g(x)$ is given and the *partial* information $E(X^n)$ is available. Then from Theorem 2.1:

Theorem 3.1 *Let given be a prior beta probability distribution of X with the density function*

$$g(x) = \frac{x^{p-1}(1-x)^{q-1}}{B(p,q)}, p, q > 0, 0 < x < 1, \tag{3.1}$$

and the constraints

$$f(x) \geq 0, \int f(x)dx = 1, \ E(X) = \int x^n f(x)dx = m. \tag{3.2}$$

Then, the minimum χ^2-divergence probability distribution of X has the density function f(x)= $\frac{x^{p-1}(1-x)^{q-1}}{B(p,q)}$ $\left[\frac{B(p+2n,q)-B(p+n,q)m+\{B(p,q)m-B(p+n,q)\}x^{n}}{B(p,q)B(p+2n,q)-B^{2}(p+n,q)}B(p,q)\right]$

(3.3)

for $m \in [\frac{B(p+n,q)}{B(p,q)}, \frac{B(p+2n,q)}{B(p+n,q)}]$, *and* $\sigma^2_{x^n,g}$ *is*

$$\sigma^2_{x^n,g} = \frac{B(p,q)B(p+2n,q) - B^2(p+n,q)}{B^2(p,q)}. \quad (3.4)$$

The minimum χ^2-*divergence measure is*

$$\chi^2_{\min}(f,g) = \frac{[B(p,q)m - B(p+n,q)]^2}{B(p,q)B(p+2n,q) - B^2(p+n,q)} \quad , \quad (3.5)$$

and

$$0 \le \chi^2_{\min}(f,g) \le \frac{B(p,q)B(p+2n,q) - B^2(p+n,q)}{B^2(p+n,q)}. \quad (3.6)$$

This may be noted from (3.3) that the density function $f(x)$ is the weighted mixture of the $Beta(p,q)$ and $Beta(p+n,q)$ density functions with respective weights

$$\frac{B(p+2n,q) - B(p+n,q)m}{B(p,q)B(p+2n,q) - B^2(p+n,q)}, \quad (3.7)$$

and

$$\frac{B(p,q)m - B(p+n,q)}{B(p,q)B(p+2n,q) - B^2(p+n,q)}, \quad (3.8)$$

, i.e.,

$$X\tilde{\ } \frac{B(p+2n,q)-B(p+n,q)m}{B(p,q)B(p+2n,q)-B^2(p+n,q)}Beta(p,q) + \frac{B(p,q)m-B(p+n,q)}{B(p,q)B(p+2n,q)-B^2(p+n,q)}Beta(p+n,q). \quad (3.9)$$

Properties of the *Minimum* χ^2-*Divergence Probability Distribution* $f(x)$:
i) t^{th} *Moment* of X about the origin:

$$m_{t,f} = \frac{[B(p+2n,q) - B(p+n,q)m]B(p+t,q) + [B(p,q)m - B(p+n,q)B(p+n+t,q)]}{B(p,q)B(p+2n,q) - B^2(p+n,q)}. \quad (3.10)$$

Mean (μ_f) and *Variance* (σ^2_f):

$$\mu_f = \frac{[B(p+2n,q) - B(p+n,q)m]B(p+1,q) + [B(p,q)m - B(p+n,q)B(p+n+1,q)]}{B(p,q)B(p+2n,q) - B^2(p+n,q)}, \quad (3.11)$$

$$\sigma^2_f = \frac{[B(p+2n,q) - B(p+n,q)m]B(p+2,q) + [B(p,q)m - B(p+n,q)B(p+n+2,q)]}{B(p,q)B(p+2n,q) - B^2(p+n,q)} - \mu^2_f. \quad (3.12)$$

ii) *Distribution Function*:

$$F(x) = \frac{[B(p+2n,q) - B(p+n,q)m]\,G(x:p,q) + [B(p,q)m - B(p+n,q)]\,G(x:p+n,q)}{B(p,q)B(p+2n,q) - B^2(p+n,q)}B(p,q)\ , \quad (3.13)$$

where

$$G(x:u,v) = \int_0^x \frac{w^{u-1}(1-w)^{v-1}}{B(u,v)}dw, u,v > 0. \quad (3.14)$$

iii) *Survival Function:*

$$S(x) \ = \frac{[B(p+2n,q)-B(p+n,q)m]\,[1-G(x:p,q)]+[B(p,q)m-B(p+n,q)]\,[1-G(x:p+n,q)]}{B(p,q)B(p+2n,q)-B^2(p+n,q)}B(p,q). \tag{3.15}$$

iv) *Hazard Function*:

$$h(x) = \frac{[B(p+2n,q)-B(p+n,q)m+\{B(p,q)m-B(p+n,q)\}x^n] \quad g(x)}{[B(p+2n,q)-B(p+n,q)m]\,[1-G(x:p,q)]+[B(p,q)m-B(p+n,q)]\,[1-G(x:p+n,q)]}. \tag{3.16}$$

Following are some interesting corollaries:

Corollary 3.1 *For $m = \frac{B(p+n,q)}{B(p,q)}$, the density function $f(x) = g(x)$, that is, the minimum χ^2-divergence probability distribution of X is the beta distribution with parameters p and q.*

Corollary 3.2 *For $m = \frac{B(p+2n,q)}{B(p+n,q)}$, the probability distribution which minimizes the χ^2-divergence between $f(x)$ and $g(x)$ is*

$$f(x) = \frac{x^{p+n-1}(1-x)^{q-1}}{B(p+n,q)}. \tag{3.17}$$

that is, X is the beta distribution with parameters $p+n$ and q.

Corollary 3.3 *For $m \in [m_{\ln x,1,g}, \frac{m_{\ln x,2,g}}{m_{\ln x,1,g}}]$, the probability distribution which minimizes the χ^2-divergence between $f(x)$ and $g(x)$ is not the beta distribution. This distribution is as given in (3.3).*

For example, if $m = \frac{1}{2}[\frac{B(p+n,q)}{B(p,q)} + \frac{B(p+2n,q)}{B(p+n,q)}]$, then the probability distribution has the density function

$$f(x) = \ \frac{1}{2}\left[\frac{x^{p-1}(1-x)^{q-1}}{B(p,q)} + \frac{x^{p+n-1}(1-x)^{q-1}}{B(p+n,q)}\right]. \tag{3.18}$$

The minimum χ^2-divergence is

$$\chi^2_{\min}(f,g) = \quad \frac{B(p,q)B(p+2n,q)-B^2(p+n,q)}{4B^2(p+n,q)} \ . \tag{3.19}$$

Now using the results of this section, we consider some specific cases of $E(X^n)$ for $n = 1$ and $n = 2$ in the following:

3.1 Case $n = 1$,i.e., Given Arithmetic Mean $E(X)$

Given *a prior* beta density function $g(x)$ and the *partial* information $E(X)$, we have from Theorem 2.1:

Theorem 3.2 *Let given be a prior beta probability distribution of X with the density function*

$$g(x) = \frac{x^{p-1}(1-x)^{q-1}}{B(p,q)}, p,q > 0, 0 \le x < 1, \tag{3.20}$$

and the constraints

$$f(x) \ge 0, \int f(x)dx = 1, \ E(X) = \int xf(x)dx = m. \tag{3.21}$$

Then, the minimum χ^2-divergence probability distribution of X has the density function

$$f(x) = \ \frac{p+q}{pq} \ \ [p\{(p+1)-(p+q+1)m\}+(p+q+1)\{(p+q)m-p\}x\ ]\, \frac{x^{p-1}(1-x)^{q-1}}{B(p,q)} \ , \tag{3.22}$$

for $m \in [\frac{p}{p+q}, \frac{p+1}{p+q+1}]$.

The minimum χ^2-*divergence measure is*

$$\chi^2_{\min}(f,g) = \frac{(p+q+1)[m(p+q)-p]^2}{pq}, \tag{3.23}$$

and

$$0 \leq \chi^2_{\min}(f,g) \leq \frac{q}{p(p+q+1)}. \tag{3.24}$$

This may be noted from (3.22) that the density function $f(x)$ is the weighted mixture of the $Beta(p,q)$ and $Beta(p+1,q)$ density functions with respective weights $\frac{(p+q)[(p+1)-(p+q+1)m]}{q}$ and $\frac{(p+q+1)[(p+q)m-p]}{q}$, i.e.,

$$X \sim \frac{(p+q)[(p+1)-(p+q+1)m]}{q} Beta(p,q) + \frac{(p+q+1)[(p+q)m-p]}{q} Beta(p+1,q). \tag{3.25}$$

Properties of the *Minimum* χ^2-*Divergence Probability Distribution* $f(x)$:

i) t^{th} *Moment* of X about the origin $(t = 1,2,3,...)$:

$$m_{t,f} = \frac{(p+q)[(p+1)-(p+q+1)m]}{q} \frac{B(p+t,q)}{B(p,q)} + \frac{(p+q+1)[(p+q)m-p]}{q} \frac{B(p+t+1,q)}{B(p+1,q)} . \tag{3.26}$$

Mean (μ_f) and *Variance* (σ_f^2):

$$\mu_f = m, \tag{3.27}$$

$$\sigma_f^2 = p\,(p+1)\,\frac{(p+q+1)\,[(p+1)-m\,(p+q)]+1}{(p+q)\,(p+q+1)\,(p+q+2)} - m^2. \tag{3.28}$$

ii) *Distribution Function*:

$$F(x) = \frac{p+q}{pq}\,[p\{(p+1)-(p+q+1)m\}G(x:p,q) + (p+q+1)\{(p+q)m-p\}G(x:p+1,q)\,], \tag{3.29}$$

iii) *Survival Function:*

$$S(x) = \frac{p+q}{pq}\,[p\{(p+1)-(p+q+1)m\}(1-G(x:p,q)) + (p+q+1)\{(p+q)m-p\}(1-G(x:p+1,q))\,] . \tag{3.30}$$

iv) *Hazard Function*:

$$h(x) = \frac{[p\{(p+1)-(p+q+1)m\} + (p+q+1)\{(p+q)m-p\}x\,]\,g(x)}{p\{(p+1)-(p+q+1)m\}[1-G(x:p,q)] + (p+q+1)\{(p+q)m-p\}[1-G(x:p+1,q)]}. \tag{3.31}$$

Following are some interesting corollaries:

Corollary 3.4 *For* $m = \frac{p}{p+q}$*, the density function* $f(x) = g(x)$*, that is, the minimum* χ^2-*divergence probability distribution of* X *is the beta distribution with parameters* p *and* q*.*

Corollary 3.5 *For* $m = \frac{p+1}{p+q+1}$*, the probability distribution which minimizes the* χ^2-*divergence is*

$$f(x) = \frac{x^p(1-x)^{q-1}}{B(p+1,q)}. \tag{3.32}$$

that is, X *has a beta distribution with parameters* $(p+1,q)$*.*

Corollary 3.6 *For $m \in [\frac{p}{p+q}, \frac{p+1}{p+q+1}]$, the probability distribution which minimizes the χ^2-divergence between $f(x)$ and $g(x)$ is not the beta. This distribution is as given in (3.22).*

For example, if $m = \frac{1}{2}[\frac{p}{p+q} + \frac{p+1}{p+q+1}] = \frac{2p^2+2pq+2p+q}{2(p+q)(p+q+1)}$, then the probability distribution has the density function

$$f(x) = \frac{1}{2}[\frac{x^{p-1}(1-x)^{q-1}}{B(p,q)} + \frac{x^p(1-x)^{q-1}}{B(p+1,q)}]. \tag{3.33}$$

The mean (μ_f) and variance (σ_f^2) :

$$\mu_f = \frac{2p(p+q+1)+q}{2(p+q)(p+q+1)}, \tag{3.34}$$

$$\sigma_f^2 = \frac{q(4p+8p^2+11pq+2q+4pq^2+4p^3+3q^2+8p^2q)}{4(p+q)^2(p+q+1)^2(p+q+2)}, \tag{3.35}$$

The minimum χ^2-divergence

$$\chi^2_{\min}(f,g) = \frac{q}{4p(p+q+1)}. \tag{3.36}$$

3.2 Case $n = 2$,i.e., Given Second Moment $E(X^2)$

When *a prior* beta density function $g(x)$ and the *partial* information $E(X^2)$ are given, we have from Theorem 2.1:

Theorem 3.3 *Let given be a prior beta probability distribution of X with the density function*

$$g(x) = \frac{x^{p-1}(x^n)^{q-1}}{B(p,q)}, p,q > 0, 0 \le x < 1, \tag{3.37}$$

and the constraints

$$f(x) \ge 0, \int f(x)dx = 1,\ E(X^2) = \int x^2 f(x)dx = v > 0. \tag{3.38}$$

Then, the minimum χ^2-divergence probability distribution of X has the density function

$$f(x) = \left[\frac{B(p+4,q)-vB(p+2,q) + \{vB(p,q) - B(p+2,q)\}x^2}{B(p,q)B(p+4,q)-B^2(p+2,q)} B(p,q)\right] \frac{x^{p-1}(1-x)^{q-1}}{B(p,q)}, \tag{3.39}$$

for $v \in \left[\frac{B(p+2,q)}{B(p,q)}, \frac{B(p+4,q)}{B(p+2,q)}\right]$ or $\left[\frac{p(p+1)}{(p+q)(p+q+1)}, \frac{(p+2)(p+3)}{(p+q+2)(p+q+3)}\right]$.

The minimum χ^2-divergence measure is

$$\chi^2_{\min}(f,g) = \frac{[vB(p,q) - B(p+2,q)]^2}{B(p,q)B(p+4,q) - B^2(p+2,q)}, \tag{3.40}$$

and

$$0 \le \chi^2_{\min}(f,g) \le \frac{B(p,q)B(p+4,q) - B^2(p+2,q)}{B^2(p+2,q)}. \tag{3.41}$$

This may be noted from (3.39) that the density function $f(x)$ is the weighted mixture of the $Beta(p,q)$ and $Beta(p+2,q)$ density functions with respective weights

$$\frac{B(p+4,q) - vB(p+2,q)}{B(p,q)B(p+4,q) - B^2(p+2,q)}, \tag{3.42}$$

and

$$\frac{vB(p,q)\ -B(p+2,q)}{B(p,q)B(p+4,q)-B^2(p+2,q)}, \tag{3.43}$$

, i.e.,

$$X\tilde{\ }\frac{B(p+4,q)-vB(p+2,q)}{B(p,q)B(p+4,q)-B^2(p+2,q)}Beta(p,q)+\frac{vB(p,q)\ -B(p+2,q)}{B(p,q)B(p+4,q)-B^2(p+2,q)}Beta(p+2,q). \tag{3.44}$$

Properties of the *Minimum* χ^2-*Divergence Probability Distribution* $f(x)$:
i) t^{th} *Moment* of X about the origin:

$$m_{t,f}=\frac{[B(p+4,q)-vB(p+2,q)\ ]B(p+t,q)+[vB(p,q)\ -B(p+2,q)B(p+t+2,q)]}{B(p,q)B(p+4,q)-B^2(p+2,q)}. \tag{3.45}$$

Mean (μ_f) and *Variance* (σ_f^2):

$$\mu_f = \frac{[B(p+4,q)-vB(p+2,q)]B(p+1,q)+[vB(p,q)\ -B(p+2,q)B(p+3,q)]}{B(p,q)B(p+4,q)-B^2(p+2,q)}, \tag{3.46}$$

$$\sigma_f^2 = \frac{[B(p+4,q)-vB(p+2,q)\ ]B(p+2,q)+[vB(p,q)\ -B(p+2,q)B(p+4,q)]}{B(p,q)B(p+4,q)-B^2(p+2,q)}-\mu_f^2. \tag{3.47}$$

ii) *Distribution Function*:

$$F(x) = \frac{[B(p+4,q)-vB(p+2,q)]\,G(x:p,q)+[vB(p,q)\ -B(p+2,q)]\,G(x:p+2,q)}{B(p,q)B(p+4,q)-B^2(p+2,q)}B(p,q)\ . \tag{3.48}$$

iii) *Survival Function:*

$$S(x) = \frac{[B(p+4,q)-vB(p+2,q)\ ]\,[1-G(x:p,q)]+[vB(p,q)\ -B(p+2,q)]\,[1-G(x:p+2,q)]}{B(p,q)B(p+4,q)-B^2(p+2,q)}B(p,q). \tag{3.49}$$

iv) *Hazard Function*:

$$h(x)=\frac{[B(p+4,q)-vB(p+n,q)v+\{vB(p,q)v-B(p+2,q)\}x^n]\quad g(x)}{[B(p+4,q)-vB(p+2,q)\ ]\,[1-G(x:p,q)]+[vB(p,q)\ -B(p+2,q)]\,[1-G(x:p+2,q)]}. \tag{3.50}$$

Following are some interesting corollaries:

Corollary 3.7 *For* $v=\frac{B(p+2,q)}{B(p,q)}$*, the density function* $f(x)=g(x)$*, that is, the minimum* χ^2-*divergence probability distribution of* X *is the beta distribution with parameters* p *and* q*.*

Corollary 3.8 *For* $v=\frac{B(p+4,q)}{B(p+2,q)}$*, the probability distribution which minimizes the* χ^2-*divergence between* $f(x)$ *and* $g(x)$ *is*

$$f(x)=\frac{x^{p+1}(1-x)^{q-1}}{B(p+2,q)}. \tag{3.51}$$

that is, X *is the beta distribution with parameters* $p+2$ *and* q*.*

Corollary 3.9 *For* $v \in [\frac{B(p+2,q)}{B(p,q)}, \frac{B(p+4,q)}{B(p+2,q)}]$, *the probability distribution which minimizes the* χ^2-*divergence between* $f(x)$ *and* $g(x)$ *is not the beta distribution. This distribution is as given in (3.39).*

For example, if $v = \frac{1}{2}[\frac{B(p+2,q)}{B(p,q)} + \frac{B(p+4,q)}{B(p+2,q)}]$, *then the probability distribution has the density function*

$$f(x) = \frac{1}{2}\left[\frac{x^{p-1}(1-x)^{q-1}}{B(p,q)} + \frac{x^{p+1}(1-x)^{q-1}}{B(p+2,q)}\right]. \tag{3.52}$$

The minimum χ^2-*divergence is*

$$\begin{aligned} \chi^2_{\min}(f,g) &= \frac{B(p,q)B(p+4,q) - B^2(p+2,q)}{4B^2(p+2,q)} . & (3.53)\\ &= \frac{2q[\ (p+\ q+1)(3+2p)+p]}{p\,(p+1)\,(p+q+2)\,(p+q+3)}. & (3.54) \end{aligned}$$

4 Given a Prior Beta Distribution and Partial Information $E(\ln X)$

Suppose given is *a prior* beta distribution and the *partial* information $E(\ln X)$. Then from Theorem 2.1:

Theorem 4.1 *Let given be a prior beta probability distribution of* X *with the density function*

$$g(x) = \frac{x^{p-1}(1-x)^{q-1}}{B(p,q)}, p, q > 0, 0 \le x < 1, \tag{4.1}$$

and the constraints

$$f(x) \ge 0, \int f(x)dx = 1,\ E(X) = \int \ln x f(x)dx = m_{\ln x,1,f}. \tag{4.2}$$

Then, the minimum χ^2-*divergence probability distribution of* X *has the density function*

$$f(x) = \left[\frac{(m_{\ln x,2,g} - m_{\ln x,1,g}m_{\ln x,1,f}) + (m_{\ln x,1,f} - m_{\ln x,1,g})\ln x}{\sigma^2_{\ln x,g}}\right] \frac{x^{p-1}(1-x)^{q-1}}{B(p,q)}\ , \tag{4.3}$$

for $m_{\ln x,1,f} \ \in [\ m_{\ln x,1,g}, \frac{m_{\ln x,2,g}}{m_{\ln x,1,g}}]$,

where for $t = 1, 2, 3, ...,$

$$m_{\ln x,t,g} = \int \frac{x^{p-1}(1-x)^{q-1}(\ln x)^t}{B(p,q)}dx, \tag{4.4}$$

and $\sigma^2_{\ln x,g}$ *is the variance of* $\ln X$.

The minimum χ^2-*divergence measure is*

$$\chi^2_{\min}(f,g) = \frac{(m_{\ln x,1,f} - m_{\ln x,1,g})^2}{\sigma^2_{\ln x,g}}, \tag{4.5}$$

and

$$0 \le \chi^2_{\min}(f,g) \le \frac{\sigma^2_{\ln x,g}}{m^2_{\ln x,1,g}}. \tag{4.6}$$

Properties of the *Minimum* χ^2-*Divergence Probability Distribution* $f(x)$:

i) t^{th} *Moment* of X about the origin:

$$m_{t,f} = \frac{(m_{\ln x,2,g} - m_{\ln x,1,g}m_{\ln x,1,f})}{\sigma^2_{\ln x,g}}\frac{B(p+t,q)}{B(p,q)} + \frac{(m_{\ln x,1,f} - m_{\ln x,1,g})}{\sigma^2_{\ln x,g}}m_{x\ln x,t,g}\ , \tag{4.7}$$

where

$$m_{x\ln x,t,g} = \int \frac{x^{t+p-1}(1-x)^{q-1}(\ln x)}{B(p,q)}dx. \tag{4.8}$$

Mean (μ_f) and *Variance* (σ_f^2):

$$\mu_f = \frac{p(m_{\ln x,2,g} - m_{\ln x,1,g}m_{\ln x,1,f})}{(p+q)\sigma^2_{\ln x,g}} + \frac{(m_{\ln x,1,f} - m_{\ln x,1,g})}{\sigma^2_{\ln x,g}}m_{x\ln x,1,g}, \tag{4.9}$$

$$\sigma_f^2 = \frac{p(p+1)(m_{\ln x,2,g} - m_{\ln x,1,g}m_{\ln x,1,f})}{(p+q)(P+q+1)\sigma^2_{\ln x,g}} + \frac{(m_{\ln x,1,f} - m_{\ln x,1,g})}{\sigma^2_{\ln x,g}}m_{x\ln x,2,g} - \mu_f^2. \tag{4.10}$$

ii) *Distribution Function*:

$$F(x) = \frac{(m_{\ln x,2,g} - m_{\ln x,1,g}m_{\ln x,1,f})}{\sigma^2_{\ln x,g}}G(x:p,q) + \frac{(m_{\ln x,1,f} - m_{\ln x,1,g})}{\sigma^2_{\ln x,g}}G_{\ln x}(x:p,q)\ , \tag{4.11}$$

where $G(x:u,v)$ is defined in ( ) and

$$G_{\ln x}(x:u,v) = \int_0^x \frac{w^{u-1}(1-w)^{v-1}\ln w}{B(u,v)}dw, u, v > 0. \tag{4.12}$$

iii) *Survival Function:*

$$S(x) = \frac{(m_{\ln x,2,g} - m_{\ln x,1,g}m_{\ln x,1,f})}{\sigma^2_{\ln x,g}}[1 - G(x:p,q)] + \frac{(m_{\ln x,1,f} - m_{\ln x,1,g})}{\sigma^2_{\ln x,g}}[1 - G_{\ln x}(x:p,q)]\ . \tag{4.13}$$

iv) *Hazard Function*:

$$h(x) = \frac{[(m_{\ln x,2,g} - m_{\ln x,1,g}m_{\ln x,1,f}) + (m_{\ln x,1,f} - m_{\ln x,1,g})\ln x]g(x)}{(m_{\ln x,2,g} - m_{\ln x,1,g}m_{\ln x,1,f})\ [1 - G(x:p,q)] + (m_{\ln x,1,f} - m_{\ln x,1,g})[1 - G_{\ln x}(x:p,q)]}. \tag{4.14}$$

Following are some interesting corollaries:

Corollary 4.1 *For $m_{\ln x,1,f} = m_{\ln x,1,g}$, the density function $f(x) = g(x)$, that is, the minimum χ^2-divergence probability distribution of X is the beta distribution with parameters p and q.*

Corollary 4.2 *For $m_{\ln x,1,f} = \frac{m_{\ln x,2,g}}{m_{\ln x,1,g}}$, the probability distribution which minimizes the χ^2-divergence between $f(x)$ and $g(x)$ is*

$$f(x) = \frac{x^{p-1}(1-x)^{q-1}\ln x}{B(p+1,q)}. \tag{4.15}$$

that is, X does not have a beta distribution.

Corollary 4.3 *For $m_{\ln x,1,f} \in [m_{\ln x,1,g}, \frac{m_{\ln x,2,g}}{m_{\ln x,1,g}}]$, the probability distribution which minimizes the χ^2-divergence between $f(x)$ and $g(x)$ is not the beta distribution. This distribution is as given in (4.3).*

For example, if $m_{\ln x,1,f} = \frac{1}{2}[m_{\ln x,1,g} + \frac{m_{\ln x,2,g}}{m_{\ln x,1,g}}]$, then the probability distribution has the density function

$$f(x) = \frac{1}{2}[\frac{x^{p-1}(1-x)^{q-1}}{B(p,q)} + \frac{x^p(1-x)^{q-1}\ln x}{B(p+1,q)m_{\ln x,1,g}}]. \tag{4.16}$$

The minimum χ^2-divergence is

$$\chi^2_{\min}(f,g) = \frac{\sigma^2_{\ln x,g}}{4m^2_{\ln x,1,g}}\ . \tag{4.17}$$

5 Given a Prior Beta Distribution and Partial Information $E[\ln(1+X)]$

When *a prior* beta distribution and the *partial* information $E[\ln(1+x)]$ are available, then from Theorem 2.1:

Theorem 5.1 *Let given be a prior beta probability distribution of X with the density function*

$$g(x) = \frac{x^{p-1}(1-x)^{q-1}}{B(p,q)}, p, q > 0, 0 \leq x < 1, \tag{5.1}$$

and the constraints

$$f(x) \geq 0, \int f(x)dx = 1,\ E(X) = \int \ln(1+x)f(x)dx = m_{\ln 1+x,1,f}. \tag{5.2}$$

Then, the minimum χ^2-divergence probability distribution of X has the density function

$$f(x)= \left[\frac{B(p,q)m_{\ln 1+x,2,g}-(1+2\ln 2-p-q)m_{\ln 1+x,1,f}+[B(p,q)m_{\ln 1+x,1,f}-(1+2\ln 2\ -p-q)]\ln(1+x)}{B(p,q)\sigma^2_{\ln 1+x,g}}\right] \times \frac{x^{p-1}(1-x)^{q-1}}{B(p,q)}\ , \tag{5.3}$$

for $m_{\ln 1+x,1,f} \ \in [\ \frac{1+2\ln 2-p-q}{B(p,q)}, \frac{B(p,q)m_{\ln 1+x,2,g}}{1+2\ln 2-p-q}]$,
where

$$m_{\ln 1+x,2,g} = \int \frac{x^{p-1}(1-x)^{q-1}[\ln(1+x)]^2}{B(p,q)}dx, \tag{5.4}$$

and $\sigma^2_{\ln 1+x,g}$ is the variance of $\ln(1+X)$.
The minimum χ^2-divergence measure is

$$\chi^2_{\min}(f,g) = \frac{[B(p,q)m_{\ln 1+x,1,f} - 1 - 2\ln 2 + p + q]^2}{B^2(p,q)\sigma^2_{\ln 1+x,g}}, \tag{5.5}$$

and

$$0 \leq \chi^2_{\min}(f,g) \leq \frac{B^2(p,q)\ \sigma^2_{\ln 1+x,g}}{(1+2\ln 2-p-q)^2}. \tag{5.6}$$

Properties of the *Minimum χ^2-Divergence Probability Distribution* $f(x)$:
i) t^{th} *Moment* of X about the origin:

$$\mathrm{m}_{t,f} = \frac{\{B(p,q)m_{\ln 1+x,2,g}-(1+2\ln 2+p+q)m_{\ln 1+x,1,f}\}B(p+t,q)+(2\ln 2+1-t-p-q)\{B(p,q)m_{\ln 1+x,1,f}-1-2\ln 2+p+q\}}{B^2(p,q)\sigma^2_{\ln 1+x,g}}. \tag{5.7}$$

Mean (μ_f) and *Variance* (σ^2_f):

$$\mu_f = \frac{\frac{p}{p+q}\{B(p,q)m_{\ln 1+x,2,g}-\ (1+2\ln 2-p-q)m_{\ln 1+x,1,f}\}+(2\ln 2-p-q)\{B(p,q)m_{\ln 1+x,1,f}-\ 1-2\ln 2+p+q\}}{B(p,q)\sigma^2_{\ln 1+x,g}},$$

$$\sigma^2_f = \frac{\frac{p(p+1)}{(p+q)(p+q+1)}\{B(p,q)m_{\ln 1+x,2,g}-\ (1+2\ln 2-p-q)m_{\ln 1+x,1,f}\}+(2\ln 2-1-p-q)\{B(p,q)m_{\ln 1+x,1,f}-1-2\ln 2+p+q\}}{B(p,q)\sigma^2_{\ln 1+x,g}}\ - \mu^2_f. \tag{5.8}$$

(5.9)

ii) *Distribution Function*:

$$F(x)=\frac{\{B(p,q)m_{\ln 1+x,2,g}-\ (1+2\ln 2-p-q)m_{\ln 1+x,1,f}\}G(x{:}p,q)+\{B(p,q)m_{\ln 1+x,1,f}-\ 1-2\ln 2+p+q\}G_{\ln(1+x)}(x{:}p,q)}{B(p,q)\sigma^2_{\ln 1+x,g}}. \tag{5.10}$$

where $G(x:u,v)$ is defined in ( ) and

$$G_{\ln(1+x)}(x:u,v)=\int_0^x\frac{w^{u-1}(1-w)^{v-1}\ln(1+w)}{B(u,v)}dw, u,v>0. \tag{5.11}$$

iii) *Survival Function:*

$$\begin{aligned}S(x)\ =\ &[B(p,q)m_{\ln 1+x,2,g}-\ (1+2\ln 2-p-q)m_{\ln 1+x,1,f}]\ [1-G(x:p,q)]\\ &+[\ B(p,q)m_{\ln 1+x,1,f}-\ 1-2\ln 2+p+q)][1-G_{\ln(1+x)}(x:p,q)].\end{aligned} \tag{5.12}$$

iv) *Hazard Function*:

$$h(x)=\frac{\begin{array}{c}[B(p,q)(m_{\ln 1+x,2,g}+m_{\ln 1+x,1,f})-(1+2\ln 2)\{m_{\ln 1+x,1,f}+\ln(1+x)\}\\ +(p+q)\{m_{\ln 1+x,1,f}+\ln(1+x)\}]g(x)\end{array}}{\begin{array}{c}[B(p,q)m_{\ln 1+x,2,g}-\ (1+2\ln 2-p-q)m_{\ln 1+x,1,f}]\ [1-G(x:p,q)]\\ +[\ B(p,q)m_{\ln 1+x,1,f}-\ 1-2\ln 2+p+q)][1-G_{\ln(1+x)}(x:p,q)]\end{array}}. \tag{5.13}$$

Following are some interesting corollaries:

Corollary 5.1 *For* $m_{\ln 1+x,1,f}=\frac{1+2\ln 2-p-q}{B(p,q)}$, *the density function* $f(x)=g(x)$, *that is, the minimum* χ^2-*divergence probability distribution of* X *is the beta distribution with parameters* p *and* q.

Corollary 5.2 *For* $m_{\ln 1+x,1,f}=\frac{B(p,q)m_{\ln 1+x,2,g}}{1+2\ln 2-p-q}$, *the probability distribution which minimizes the* χ^2-*divergence between* $f(x)$ *and* $g(x)$ *is*

$$f(x)=\frac{x^{p-1}(1-x)^{q-1}\ln(1+x)}{B(p+1,q)}. \tag{5.14}$$

that is, X *does not have a beta distribution.*

Corollary 5.3 *5.4. For* $m_{\ln 1+x,1,f}\in[\ \frac{1+2\ln 2-p-q}{B(p,q)},\frac{B(p,q)m_{\ln 1+x,2,g}}{1+2\ln 2-p-q}]$, *the probability distribution which minimizes the* χ^2-*divergence between* $f(x)$ *and* $g(x)$ *is not the beta distribution. This distribution is as given in (5.3).*

For example, if $m_{\ln 1+x,1,f}=\frac{1}{2}[\ \frac{1+2\ln 2-p-q}{B(p,q)}+\frac{B(p,q)m_{\ln 1+x,2,g}}{1+2\ln 2-p-q}]$, *then the probability distribution has the density function*

$$f(x)=\ \frac{1}{2}[\frac{x^{p-1}(1-x)^{q-1}}{B(p,q)}+\frac{x^p(1-x)^{q-1}\ln(1+x)B(p,q)}{(1+2\ln 2-p-q)B(p+1,q)}]. \tag{5.15}$$

The minimum χ^2-*divergence is*

$$\chi^2_{\min}(f,g)=\frac{B^2(p,q)\sigma^2_{\ln 1+x,g}}{4(1+2\ln 2-p-q)^2}\ . \tag{5.16}$$

6 Given a Prior Beta Distribution and Partial Information $E[\ln(1-X)]$

Suppose that *a prior* beta density function $g(x)$ and the *partial* information $E[\ln(1-x)]$ are given. Then from Theorem 2.1:

Theorem 6.1 *Let given be a prior beta probability distribution of X with the density function*

$$g(x) = \frac{x^{p-1}(1-x)^{q-1}}{B(p,q)}, p, q > 0, 0 \leq x < 1, \tag{6.1}$$

and the constraints

$$f(x) \geq 0, \int f(x)dx = 1, \; E(X) = \int \ln(1-x) f(x)dx = m_{\ln 1-x,1,f}. \tag{6.2}$$

Then, the minimum χ^2-divergence probability distribution of X has the density function

$$f(x) = \frac{x^{p-1}(1-x)^{q-1}\{B(p,q)m_{\ln 1-x,2,g} - (1-p-q)m_{\ln 1-x,1,f} + [B(p,q)m_{\ln 1-x,1,f} - (1-p-q)]\ln(1-x)\}}{B^2(p,q)\sigma^2_{\ln 1-x,g}}, \tag{6.3}$$

for $m_{\ln 1-x,1,f} \in [\frac{1-p-q}{B(p,q)}, \frac{B(p,q)m_{\ln 1-x,2,g}}{1-p-q}]$,
where

$$m_{\ln 1-x,2,g} = \int \frac{x^{p-1}(1-x)^{q-1}[\ln(1-x)]^2}{B(p,q)}dx, \tag{6.4}$$

and $\sigma^2_{\ln 1-x,g}$ is the variance of $\ln(1-x)$.
The minimum χ^2-divergence measure is

$$\chi^2_{\min}(f,g) = \frac{[B(p,q)m_{\ln 1-x,1,f} - 1 + p + q]^2}{B^2(p,q)\sigma^2_{\ln 1-x,g}}, \tag{6.5}$$

and

$$0 \leq \chi^2_{\min}(f,g) \leq \frac{B^2(p,q)\ \sigma^2_{\ln 1-x,g}}{(1-p-q)^2}. \tag{6.6}$$

Properties of the *Minimum χ^2-Divergence Probability Distribution* $f(x)$:
i) t^{th} *Moment* of X about the origin:

$$m_{t,f} = \frac{\{B(p,q)m_{\ln 1-x,2,g} - (1-p-q)m_{\ln 1-x,1,f}\}B(p+t,q) + (1-t-p-q)\{B(p,q)m_{\ln 1-x,1,f} - 1+p+q\}}{B^2(p,q)\sigma^2_{\ln 1-x,g}}. \tag{6.7}$$

Mean (μ_f) and *Variance* (σ^2_f):

$$\mu_f = \frac{\frac{p}{p+q}\{B(p,q)m_{\ln 1-x,2,g} - (1-p-q)m_{\ln 1-x,1,f}\} - (p+q)\{B(p,q)m_{\ln 1-x,1,f} - 1+p+q\}}{B(p,q)\sigma^2_{\ln 1-x,g}}, \tag{6.8}$$

$$\sigma^2_f = \frac{\frac{p(p+1)}{(p+q)(p+q+1)}\{B(p,q)m_{\ln 1-x,2,g} - (1-p-q)m_{\ln 1-x,1,f}\} - (1+p+q)\{B(p,q)m_{\ln 1-x,1,f} - 1+p+q\}}{B(p,q)\sigma^2_{\ln 1-x,g}} - \mu^2_f. \tag{6.9}$$

ii) *Distribution Function*:

$$F(x) = \frac{\{B(p,q)m_{\ln 1-x,2,g} - (1-p-q)m_{\ln 1-x,1,f}\}G(x:p,q) + \{B(p,q)m_{\ln 1-x,1,f} - 1+p+q\}G_{\ln(1-x)}(x:p,q)}{B(p,q)\sigma^2_{\ln 1-x,g}}, \tag{6.10}$$

where $G(x:u,v)$ is defined in ( ) and

$$G_{\ln(1-x)}(x:u,v) = \int_0^x \frac{w^{u-1}(1-w)^{v-1}\ln(1-w)}{B(u,v)}dw, u,v>0. \tag{6.11}$$

iii) *Survival Function:*

$$\begin{aligned} S(x) \;=\; & [B(p,q)m_{\ln 1-x,2,g} - (1-p-q)m_{\ln 1-x,1,f}]\,[1-G(x:p,q)] \\ & +[\,B(p,q)m_{\ln 1-x,1,f} - 1+p+q)][1-G_{\ln(1-x)}(x:p,q)]. \end{aligned} \tag{6.12}$$

iv) *Hazard Function*:

$$h(x) = \frac{\begin{array}{c}[B(p,q)(m_{\ln 1-x,2,g} + m_{\ln 1-x,1,f}) - m_{\ln 1-x,1,f} + \ln(1-x) \\ +(p+q)\{m_{\ln 1-x,1,f} + \ln(1-x)\}]g(x)\end{array}}{\begin{array}{c}[B(p,q)m_{\ln 1-x,2,g} - (1-p-q)m_{\ln 1-x,1,f}]\,[1-G(x:p,q)] \\ +[\,B(p,q)m_{\ln 1-x,1,f} - 1+p+q)][1-G_{\ln(1-x)}(x:p,q)]\end{array}}. \tag{6.13}$$

Following are some interesting corollaries:

Corollary 6.1 *For $m_{\ln 1-x,1,f} = \frac{1-p-q}{B(p,q)}$, the density function $f(x)=g(x)$, that is, the minimum χ^2-divergence probability distribution of X is the beta distribution with parameters p and q.*

Corollary 6.2 *For $m_{\ln 1-x,1,f} = \frac{B(p,q)m_{\ln 1-x,2,g}}{1-p-q}$, the probability distribution which minimizes the χ^2-divergence between $f(x)$ and $g(x)$ is*

$$f(x) = \frac{x^{p-1}(1-x)^{q-1}\ln(1-x)}{B(p+1,q)}. \tag{6.14}$$

that is, X does not have a beta distribution.

Corollary 6.3 *For $m_{\ln 1-x,1,f} \in [\frac{1-p-q}{B(p,q)}, \frac{B(p,q)m_{\ln 1-x,2,g}}{1-p-q}]$, the probability distribution which minimizes the χ^2-divergence between $f(x)$ and $g(x)$ is not the beta distribution. This distribution is as given in (6.3).*

For example, if $m_{\ln 1-x,1,f} = \frac{1}{2}[\frac{1-p-q}{B(p,q)} + \frac{B(p,q)m_{\ln 1-x,2,g}}{1-p-q}]$, then the probability distribution has the density function

$$f(x) = \frac{1}{2}[\frac{x^{p-1}(1-x)^{q-1}}{B(p,q)} + \frac{x^p(1-x)^{q-1}\ln(1-x)B(p,q)}{(1-p-q)B(p+1,q)}]. \tag{6.15}$$

The minimum χ^2-divergence is

$$\chi^2_{\min}(f,g) = \frac{B^2(p,q)\sigma^2_{\ln 1-x,g}}{4(1-p-q)^2}. \tag{6.16}$$

7 Given a Prior Beta Distribution and Partial Information $E(X)$ and $V(X)$

Suppose given are *a prior* beta probability density function and the partial information average $E(X)=a$ and variance $V(X)=v$. Then, we have from Theorem 2.1:

Theorem 7.1 *Let given be a prior beta density function of the continuous random variable X,*

$$g(x:p,q)=\frac{x^{p-1}(1-x)^{q-1}}{B(p,q)},p,q>0,0\leq x<1, \tag{7.1}$$

and

$$f(x)\geq 0,\int f(x)dx=1,\ \int xf(x)dx=a\ ,\int x^2f(x)dx=b=v+a^2. \tag{7.2}$$

Then the minimum χ^2-divergence probability distribution of X has the density function

$$\begin{aligned} f(x) &= \left(\alpha_0+\alpha_1x+\alpha_2x^2\right)\frac{x^{p-1}(1-x)^{q-1}}{B(p,q)}, && (7.3)\\ &= \alpha_0 g(x:p,q)+\frac{p}{p+q}\left(\alpha_1 g(x:p+1,q)+\frac{p+1}{p+q+1}\alpha_2 g(x:p+2,q)\right), && (7.4)\end{aligned}$$

where

$$\begin{aligned} \alpha_0 &= \frac{a(m_{1,g}m_{4,g}-m_{2,g}m_{3,g})-b(m_{1,g}m_{3,g}-m_{2,g}^2\ )+m_{3,g}^2-m_{4,g}m_{2,g}}{m_{2,g}^3+m_{3,g}^2+m_{4,g}m_{1,g}^2-2m_{1,g}m_{2,g}m_{3,g}-m_{4,g}m_{2,g}}, && (7.5)\\ \alpha_1 &= \frac{a(m_{2,g}^2-m_{4,g})+b(m_{3,g}-m_{1,g}m_{2,g})+\ m_{1,g}m_{4,g}-m_{2,g}m_{3,g}}{m_{2,g}^3+m_{3,g}^2+m_{4,g}m_{1,g}^2-2m_{1,g}m_{2,g}m_{3,g}-m_{4,g}m_{2,g}}, && (7.6)\\ \alpha_2 &= \frac{a(m_{3,g}-m_{1,g}m_{2,g}\ )+b(m_{1,g}^2-m_{2,g})+m_{2,g}^2-m_{3,g}m_{1,g}}{m_{2,g}^3+m_{3,g}^2+m_{4,g}m_{1,g}^2-2m_{1,g}m_{2,g}m_{3,g}-m_{4,g}m_{2,g}}, && (7.7)\end{aligned}$$

and

$$m_{t,g}=\frac{B(p+t,q)}{B(p,q)}=\frac{p(p+1)(p+2)....(p+t-1)}{(p+q)(p+q+1)(p+q+2)....(p+q+t-1)}. \tag{7.8}$$

The minimum χ^2-divergence measure is given by

$$\chi^2_{\min}(f,g)=\int\ \left(\alpha_0+\alpha_1x+\alpha_2x^2\right)^2\frac{x^{p-1}(1-x)^{q-1}}{B(p,q)}dx-1. \tag{7.9}$$

Properties of the *Minimum χ^2-Divergence Probability Distribution $f(x)$*:

(i) *t^{th} moment $(t=1,2,3,...)$ about origin of X :*

$$m_{t,f}=\frac{B(p+t,q)}{B(p,q)}\left[\alpha_0+\frac{(p+t\ )}{(p+q+t)}\left(\alpha_1+\frac{(p+t\ +1)}{(p+q+t+1)}\alpha_2\right)\ \right]. \tag{7.10}$$

In particular,

$$m_{1,f}=\frac{p}{p+q}\left[\alpha_0+\frac{(p+1\ )}{(p+q+1)}\left(\alpha_1+\frac{(p+2)}{(p+q+2)}\alpha_2\right)\ \right], \tag{7.11}$$

and

$$m_{2,f}=\frac{p(p+1)}{(p+q)(p+q+1)}\left[\alpha_0+\frac{(p+2\ )}{(p+q+2)}\left(\alpha_1+\frac{(p+3)}{(p+q+3)}\alpha_2\right)\ \right]. \tag{7.12}$$

(ii) *Distribution Function*:

$$F(x)\ =\ \alpha_0G(x:p,q)+\frac{p}{p+q}\left(\alpha_1G(x:p+1,q)+\frac{p+1}{p+q\ +1}\alpha_2G(x:p+2,q)\right). \tag{7.13}$$

(iii) *Survival Function:*

$$S(x) \;=\; \alpha_0\left[1-G(x:p,q)\right]+\frac{p}{p+q}\left[\alpha_1\{1-G(x:p+1,q)\}+\frac{p+1}{p+q\;+1}\alpha_2\{1-G(x:p+2,q)\}\right]\;. \quad (7.14)$$

(iv) *Hazard Function:*

$$h(x)=\frac{\alpha_0(p+q)g(x:p,q)+p\left(\alpha_1 g(x:p+1,q)+\frac{p+1}{p+q+1}\alpha_2 g(x:p+2,q)\right)}{\alpha_0(p+q)\left[1-G(x:p,q)\right]+p\left[\alpha_1\{1-G(x:p+1,q)\}+\frac{p+1}{p+q\;+1}\alpha_2\{1-G(x:p+2,q)\}\right]}. \quad (7.15)$$

8 Numerical Illustration

The minimum χ^2-divergence probability distributions given a prior beta distributions for $(p,q)=(0.2,1)$, $(0.5,0.5)$, $(1,1)$, $(2,4)$ and the information about the current mean $E(X)$, $E(X^2)$ and $V(X)$ are considered in Tables 1-3 and Figures 1-3.

Table 1. The Minimum χ^2-divergence Probability Distributions Given: A Prior Beta Distribution and Arithmetic Mean.

$g(x)$	$E(X^2)$	$f(x)$	$\chi^2_{\min}$	*Figure*
$Beta(0.2,1)$	$\begin{cases}0.1667\\0.3561\\0.5455\end{cases}$	$\begin{cases}Beta(0.2,1)\\\frac{1}{2}\left[Beta(0.2,1)+Beta(1.2,1)\right]\\Beta(1.2,1)\end{cases}$	$\begin{cases}0\\0.5618\\2.2727\end{cases}$	$1a$
$Beta(0.5,0.5)$	$\begin{cases}0.5\\0.625\\0.75\end{cases}$	$\begin{cases}Beta(0.5,0.5)\\\frac{1}{2}\left[Beta(0.5,0.5)+Beta(1.5,0.5)\right]\\Beta(1.5,0.5)\end{cases}$	$\begin{cases}0\\0.125\\0.5\end{cases}$	$1b$
$Beta(1,1)$	$\begin{cases}0.5\\0.5833\\0.6667\end{cases}$	$\begin{cases}Beta(1,1)\\\frac{1}{2}\left[Beta(1,1)+Beta(2,1)\right]\\Beta(\;2,1)\end{cases}$	$\begin{cases}0\\0.0833\\0.3333\end{cases}$	$1c$
$Beta(2,4)$	$\begin{cases}0.3333\\0.3809\\0.4286\end{cases}$	$\begin{cases}Beta(\;2,4)\\\frac{1}{2}\left[Beta(2,4)+Beta(3,4)\right]\\Beta(3,4)\end{cases}$	$\begin{cases}0\\0.0714\\0.28557\end{cases}$	$1d$

Table 2. The Minimum χ^2-divergence Probability Distributions Given: A Prior Beta Distribution and Second Moment.

$g(x)$	$E(X^2)$	$f(x)$	$\chi^2_{\min}$	*Figure*
$Beta(0.2,1)$	$\begin{cases}0.09091\\0.30736\\0.52381\end{cases}$	$\begin{cases}Beta(0.2,1)\\\frac{1}{2}\left[Beta(0.2,1)+Beta(2.2,1)\right]\\Beta(2.2,1)\end{cases}$	$\begin{cases}0\\1.1905\\4.7618\end{cases}$	$2a$
$Beta(0.5,0.5)$	$\begin{cases}0.375\\0.55208\\0.72916\end{cases}$	$\begin{cases}Beta(0.5,0.5)\\\frac{1}{2}\left[Beta(0.5,0.5)+Beta(2.5,0.5)\right]\\Beta(2.5,0.5)\end{cases}$	$\begin{cases}0\\0.2361\\0.9444\end{cases}$	$2b$
$Beta(1,1)$	$\begin{cases}0.33333\\0.46667\\0.6\end{cases}$	$\begin{cases}Beta(1,1)\\\frac{1}{2}\left[Beta(1,1)+Beta(3,1)\right]\\Beta(\;3,1)\end{cases}$	$\begin{cases}0\\0.2\\0.8\end{cases}$	$2c$
$Beta(2,4)$	$\begin{cases}0.14285\\0.21031\\0.27777\end{cases}$	$\begin{cases}Beta(\;2,4)\\\frac{1}{2}\left[Beta(2,4)+Beta(4,4)\right]\\Beta(4,4)\end{cases}$	$\begin{cases}0\\0.2361\\0.9443\end{cases}$	$2d$

Table 3. The Minimum χ^2-divergence Probability Distributions Given: A Prior Beta Distribution, First and Second Moments.

$g(x)$	$E(X)$, $E(X^2)$	$f(x)$	*Figure*
$Beta(0.5, 0.5)$	$\begin{cases} 0.5, 0.375 \\ 0.625, 0.5521 \\ 0.75, 0.7292 \end{cases}$	$\begin{cases} Beta(0.5, 0.5) \\ \left[(-2.258x10^{-4})Beta(0.5, 0.5) + (0.089)Beta(1.5, 0.5) \right. \\ +(0.1592)Beta(2.5, 0.5) \\ \left[(-2.4842x10^{-3})Beta(0.5, 0.5) + (0.0967)Beta(1.5, 0.5) \right. \\ +(0.1965)Beta(2.5, 0.5) \end{cases}$	$3a$
$Beta(1, 1)$	$\begin{cases} 0.5, 0.3333 \\ 0.5833, 0.4667 \\ 0.6667, 0.6 \end{cases}$	$\begin{cases} Beta(1, 1) \\ [(0.00738)Beta(1, 1) + (0.0498)Beta(2, 1) \\ +(0.08487)Beta(3, 1) \\ [(0.0111)Beta(1, 1) + (0.0443)Beta(2, 1) \\ +(0.0959)Beta(3, 1) \end{cases}$	$3b$
$Beta(2, 4)$	$\begin{cases} 0.3333, 0.1428 \\ 0.3809, 0.2103 \\ 0.4286, 0.2778 \end{cases}$	$\begin{cases} Beta(\ 2, 4) \\ [(0.00765)Beta(\ 2, 4) + (0.0625)Beta(\ 3, 4) \\ +(0.9392)Beta(\ 4, 4) \\ [(0.00127)Beta(\ 2, 4) + (0.0586)Beta(\ 3, 4) \\ +(0.10237)Beta(4, 4) \end{cases}$	$3c$

From these tables and figures, there is evidence that the new (minimum chi-square divergence) distribution, given a *prior* distribution as beta and the new (current) information on moments, is not the beta 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 beta distribution, in addition to the partial information in terms of the moments, the minimum chi-square divergence principle provides a useful methodology for characterizing beta probability distributions.

REFERENCES

Kullback, S. *Information Theory and Statistics.* 1959, New York: John Wiley.

Kumar, Pranesh and Taneja, I. J. *Chi-square divergence and minimization problem for continuous probability distributions.* (submitted for publication).

Pearson, K. *On the Criterion that a given system of deviations from the probable in the case of correlated system of variables is such that it can be reasonable supposed to have arisen from random sampling. Phil. Mag.*,1900, 50, 157-172.

Shore, J.E. and Johnson, R.W. *Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross entropy. IEEE Trans. Information Theory*, 1980, IT-26,1,26-37.

Taneja, I. J. *Generalized Information Measures and their Applications.* On line book: *http://www.mtm.ufsc.br/∼taneja/book/book.html*, 2001.