ORİJİNAL ARAŞTIRMA ORIGINAL RESEARCH

DOI: 10.5336/biostatic.2016-52392

Survival Data Analysis by Minminxent and Maxminxent Methods

Minminxent ve Maxminxent Yöntemleri ile Sağkalım Veri Analizi

Aladdin SHAMILOV,[a]
Sevda ÖZDEMİR[b]

[a]Department of Statistics,
Anadolu University Faculty of Science,
Eskişehir
[b]Department of Accounting and
Tax Administration,
Yüzüncü Yıl University
Özalp Vocational School, Van

Geliş Tarihi/*Received:* 11.07.2016
Kabul Tarihi/*Accepted:* 29.09.2016

Yazışma Adresi/*Correspondence:*
Sevda ÖZDEMİR
Yüzüncü Yıl University
Özalp Vocational School,
Department of Accounting and
Tax Administration, Van,
TÜRKİYE/TURKEY
sevdaozdemir@yyu.edu.tr

ABSTRACT Objective: Entropy Optimization Methods (EOM) have important applications, especially in statistics, economy, engineering, survival data analysis and etc. There are several examples in the literature that known statistical data do not conform to theoretical distributions, however do conform the entropy optimization distributions well. In the present study, survival data of male patients with localized cancer of a rectum diagnosed in Connecticut from 1935 to 1944 is analyzed by using Generalized Entropy Optimization Methods (GEOM) in the form of the MinMinxEnt and the MaxMinxEnt methods. **Material and Methods:** The MinMinxEnt and the MaxMinxEnt methods have suggested distributions in the form of the MinMinxEnt, the MaxMinxEnt distributions which are closest to statistical data and furthest from statistical data in the sense of Kullback-Leibler measure, respectively. **Results:** The results are acquired by using statistical software MATLAB. The performances of MinMixEnt and MaxMinxEnt methods are established by Chi-Square, Root Mean Square Error (RMSE) and Kullback-Leibler criteria. It is shown that $(MinMinxEnt)_4$ is better than $(MaxMinxEnt)_4$ distribution in the sense of Kullback-Leibler measure to mentioned data. Furthermore, in the sense of RMSE criteria $(MaxMinxEnt)_4$ distribution is more suitable for statistical data than $(MinMinxEnt)_4$ distribution. These results are also corroborated by graphical representation. **Conclusion:** In this study, it is shown that $(MinMinxEnt)_4$ and $(MaxMinxEnt)_4$ distributions more successfully represent Survival Data. Our investigation indicates that GEOM in survival data analysis yields reasonable results.

Keywords: Generalized entropy optimization methods; MinMinxEnt; MaxMinxEnt distributions; Kullback-Leibler measure

ÖZET Amaç: Entropi Optimizasyon Yöntemleri (EOY) özellikle istatistik, ekonomi, mühendislik, sağkalım veri analizi vb. alanlarda önemli uygulamalara sahiptir. Literatürde istatistiksel verinin bilinen teorik dağılımlara uymadığı ancak entropi optimizasyon dağılımlarına iyi bir şekilde uyduğunu gösteren çeşitli örneklerin varlığı mevcuttur. Bu çalışmada, Connecticut şehrinde 1935-1944 yılları arasındaki bağırsak kanseri tanısı konulmuş erkek hastaların sağkalım verileri MinMinxEnt ve MaxMinxEnt yöntemleri şeklinde Genelleştirilmiş Entropi Optimizasyon Yöntemleri (GEOY) kullanılarak analiz edilmiştir. **Gereç ve Yöntemler:** MinMinxEnt ve MaxMinxEnt yöntemleri Kullback-Leibler ölçümüne göre sırasıyla, istatistiksel veriye en yakın ve en uzak MinMinxEnt ve MaxMinxEnt dağılımların bulunmasını sunmaktadır. **Bulgular:** Sonuçlar MATLAB Programını uygulamakla elde edilmiştir. MinMinxEnt ve MaxMinxEnt yöntemlerinin performansı Ki-Kare, Hata Kareler Ortalamasının Kökü (RMSE), ve Kullback-Leibler ölçümü kriterleri kullanılarak belirlenmiştir. Kullbak-Leibler ölçümüne göre bahsi geçen veri için $(MinMinxEnt)_4$ dağılımının $(MaxMinxEnt)_4$ dağılımından daha iyi olduğu gösterilmiştir. Dahası, RMSE Kriterine göre, veriye $(MaxMinxEnt)_4$ dağılımı $(MinMinxEnt)_4$ dağılımından daha iyi uyum sağlamaktadır. Sonuçlar grafiksel olarak da gösterilmiştir. **Sonuç:** Bu çalışmada, $(MinMinxEnt)_4$ ve $(MaxMinxEnt)_4$ dağılımlarının Sağkalım verilerini başarılı bir şekilde temsil ettiği gözlenmiştir. Araştırmalarımız göstermiştir ki sağkalım veri analizinde GEOY başarılı sonuçlar vermektedir.

Anahtar Kelimeler: Genelleştirilmiş entropi optimizasyon yöntemleri; MinMinxEnt; MaxMinxEnt yöntemleri; Kullback-Leibler ölçümü

Entropy Optimization Methods (EOM) have important applications, especially in statistics, economy, engineering survival data analysis and so on. There are several examples in the literature that known statistical data do not conform to theoretical distributions, however do conform the entropy optimization distributions well.[1] Generalized Entropy Optimization Methods (GEOM) have suggested distributions in the form of MinMinxEnt, MaxMinxEnt which are closest to statistical data and furthest from mentioned data in the sense of information theory, respectively. For this reason, GEOM can be more successfully applied in Survival Data Analysis.[2-6]

Different aspects and methods of investigations of survival data analysis are considered in.[7-12]

In particular in the paper it is investigated several problems of hazard rate function estimation based on the maximum entropy principle.[10] The potential applications include developing several classes of the maximum entropy distributions which can be used to model different data-generating distributions that satisfy certain information constraints on the hazard rate function.

In order to represent the results of our investigations, we give some auxiliary concepts and facts first.

SURVIVAL DATA ANALYSIS

Survival Time: Survival time can be defined broadly as the time to the occurrence of a given event. This event can be the development of a disease, response to a treatment, relapse or death.[13]

Censoring: The techniques for reducing experimental time are known as censoring. In survival analysis, the observations are lifetimes which can be indefinitely long. So quite often the experiment is so designed that the time required for collecting the data is reduced to manageable levels.

Let T be a continuous, non-negative valued random variable representing the lifetime of a unit. This is the time for which an individual (or unit) carries out its appointed task satisfactorily and then passes into "failed" or "dead" state thereafter.[14]

GENERALIZED ENTROPY OPTIMIZATION METHODS (GEOM)

Entropy Optimization Problem (EOP) and Generalized Entropy Optimization problem (GEOP) can be formulated in the following form.[3,15]

EOP: Let $f^{(0)}(x)$ be given probability density function (p.d.f.) of random variable X, L be an entropy optimization measure and $g(x)$ be a given moment vector function generating m moment constraints. It is required to obtain the distribution corresponding to $g(x)$, which gives extreme value to L.[3]

GEOP: Let $f^{(0)}(x)$ be given probability density function of random variable X, L be an entropy optimization measure and K be a set of given moment vector functions. It is required to choose moment vector functions $g^{(1)}$, $g^{(2)} \in K$ such that $g^{(1)}$ defines entropy optimization distribution $f^{(1)}(x)$ closest to $f^{(0)}(x)$, $g^{(2)}$ defines entropy optimization distribution $f^{(2)}(x)$ furthest from $f^{(0)}(x)$ with respect to entropy optimization measure L. If L is taken as Shannon entropy measure, then $f^{(1)}(x)$ is called the MinMaxEnt distribution, and $f^{(2)}(x)$ is called the MaxMaxEnt distribution. If L is taken as Kullback - Leibler measure, then $f^{(1)}(x)$ is called the MinMinxEnt distribution, and $f^{(2)}(x)$ is called the MaxMinxEnt distribution.[3]

The method of solving GEOP is called as GEOM.

MinxEnt FUNCTIONAL

The problem of minimizing Kullback-Leibler functional (function)

$$D(p : q) = \sum_{i=1}^{n} p_i \ln \frac{p_i}{q_i}, p = (p_1, \cdots, p_n),\ q = (q_1, \cdots, q_n) \quad (1)$$

subject to constraints

$$\sum_{i=1}^{n} p_i g_j(x_i) = \mu_j\ ,\ j = 0,1,2, \cdots, m \quad (2)$$

where $\mu_0 = 1, g_0(x) = 1, p_i \geq 0, i = 1,2, \cdots, n\ , m + 1 < n$ and $q_i \geq 0, i = 1,2, \cdots, n$;

$\sum_{i=1}^{n} q_i \leq 1$ has solution

$$p_i = q_i e^{-\sum_{j=0}^{m} \lambda_j g_j(x_i)}\ , i = 1,2, \cdots, n \quad (3)$$

where $\lambda_j (j = 0,1,2, \cdots, m)$ are Lagrange multipliers. In the literature, there have been numerous studies that have calculated these multipliers.[2] In this study, we use the MATLAB program to calculate Lagrange multipliers.

If (3) is substituted in (1), the minimum value of $D(p : q)$ is obtained:

$$D(p : q) = \sum_{i=1}^{n} q_i e^{-\sum_{j=0}^{m} \lambda_j g_j(x_i)} \left(-\sum_{j=0}^{m} \lambda_j g_j(x_i)\right) = -\sum_{j=0}^{m} \lambda_j \mu_j\ . \quad (4)$$

If distribution $p^{(0)} = \left(p_1^{(0)}, \cdots, p_n^{(0)}\right)$ is calculated from the data, the moment vector value $\mu = (1, \mu_1, \cdots \mu_m)$ can be obtained for each moment vector function $g(x) = \left(1, g_1(x), \cdots, g_m(x)\right)$. Thus, $D(p : q)$ is considered as a functional of $g(x)$ and called the MinxEnt functional (function) and is noted by $U(g)$.

MinMinxEnt and MaxMinxEnt DISTRIBUTIONS

Let K be the compact set of moment vector functions $g(x)$. $U(g)$ reaches its least and greatest values in this compact set, because of its continuity property. For this reason,

$$\min_{g \in K} U(g) = U\left(g^{(1)}\right);\ \max_{g \in K} U(g) = U\left(g^{(2)}\right).$$

Consequently, $U\left(g^{(1)}\right) \leq U\left(g^{(2)}\right)$. Distribu-tions $p^{(1)} = \left(p_1^{(1)}, \cdots, p_n^{(1)}\right)$ and $p^{(2)} = \left(p_1^{(2)}, \cdots, p_n^{(2)}\right)$ corresponding to the $g^{(1)}(x)$ and $g^{(2)}(x)$, respectively, are called MinMinxEnt and MaxMinxEnt distributions.[2,3]

MinMinxEnt and MaxMinxEnt METHODS FOR SURVIVAL DATA

MinMinxEnt and MaxMinxEnt DISTRIBUTIONS

In the present research, the data of male patients with localized cancer of a rectum diagnosed in Connecticut from 1935 to 1944 given in Table 1 is considered.[13]

In this investigation, the experiment is planned for 388 numbers of patients surviving at beginning of interval but the presence of censoring from the planning patients 52 individuals stay out the experiment. This situation is taken into account in Table 2.

The presence of censoring in the survival times leads to situation that for the survival data the sum of observation probabilities stands less than 1. For this reason in solving many problems it is required to

TABLE 1: The data of male patients with localized cancer of a rectum diagnosed in Connecticut from 1935 to 1944.

Survival Time (year) t	Number of patients surviving at beginning of interval n_i	Number of patients dying in interval (d_i)	Number of patients censoring in interval (c_i)
1	388	167	2
2	219	45	1
3	173	45	1
4	127	19	0
5	108	17	0
6	91	11	1
7	79	8	0
8	71	5	0
9	66	6	1
10	59	7	0

TABLE 2: Observation probabilities, Corrected probabilities.

t	n_i'	d_i	c_i	Observed probabilities q_i	Corrected probabilities p_i^*
0-1	336	167	2	0.4970	0.5030
1-2	167	45	1	0.1339	0.1369
2-3	121	45	1	0.1339	0.1369
3-4	75	19	0	0.0565	0.0565
4-5	56	17	0	0.0506	0.0506
5-6	39	11	1	0.0327	0.0357
6-7	27	8	0	0.0238	0.0238
7-8	19	5	0	0.0149	0.0149
8-9	14	6	1	0.0179	0.0208
9-10	7	7	0	0.0208	0.0208

Source: In here, n_i' denotes that number of patients surviving except for the presence of censoring from the planning patients 52 individuals at beginning of interval

supplement the sum of observation probabilities up to 1. Consequently, the basing of the admissibility of this method of supplementation acquires a new significance. Mentioned problem is solved by applying the MinMinxEnt and MaxMinxEnt methods so that among Entropy Optimization Distributions with respect to the Kullback-Leibler measure, it is chosen distribution which is closest to observed probability distribution or furthest from this distribution. Since the sum of observed probabilities $q_i, i = 1,2,\cdots,10$ is 0.9821 in Table 2 , according to the number of censoring, supplementary probability $1 - 0.9821 = 0.0179$ is uniformly distributed to each censoring data and corrected probabilities $p_i^*, i = 1,2,\cdots,10$ are obtained.

Let $K_0 = \{g_1,\cdots,g_r\}$ be the set of characterizing moment vector functions and all combinations of r elements of K_0 taken m elements at a time be $K_{0,m}$. We note that, each element of $K_{0,m}$ is vector g with m components.

Solving the MinMinxEnt and the MaxMinxEnt problems require to find vector functions $\left(g_0(x), g^{(1)}(x)\right)$, $\left(g_0(x), g^{(2)}(x)\right)$, where $g_0(x) = 1, g^{(1)} \in K_{0,m}$, $g^{(2)} \in K_{0,m}$ which give minimum and maximum values, respectively to $U(g)$ with respect to Kullback-Leibler measure. It should be noted that $U(g)$ reaches its minimum (maximum) value subject to constraints generated by function $g_0(x)$ and all m −dimensional vector functions $g(x), g \in K_{0,m}$. In other words, minimum (maximum) value of $U(g)$ is the least (the greatest) value of D_{min} values corresponding to $g(x), g \in K_{0,m}$. If $\left(g_0(x), g^{(1)}(x)\right)$ $\left(\left(g_0(x), g^{(2)}(x)\right)\right)$ gives the minimum (maximum) value to $U(g)$ then distribution $p^{(1)} = \left(p_1^{(1)}, \cdots, p_n^{(1)}\right)$ $\left(p^{(2)} = \left(p_1^{(2)}, \cdots, p_n^{(2)}\right)\right)$ corresponding to $\left(g_0(x), g^{(1)}(x)\right)$ $\left(\left(g_0, g^{(2)}(x)\right)\right)$ is called the MinMinxEnt (the MaxMinxEnt) distribution.

Corollary: If by $(MinMinxEnt)_m$ $((MaxMinxEnt)_m)$ denote the MinMinxEnt (the MaxMinxEnt) distribution corresponding to m moment conditions generated by moment functions $g(x), g \in K_{0,m}$ and $q = (q_1, \cdots, q_n)$ is given distribution then inequalities

$D\left((MinMinxEnt)_{m_1} : q\right) \leq D\left((MinMinxEnt)_{m_2} : q\right)$

$D\left((MaxMinxEnt)_{m_1} : q\right) \leq D\left((MaxMinxEnt)_{m_2} : q\right)$

are fulfilled, when $m_1 < m_2$. In other words, Kullback -Leibler measure of the MinMinxEnt distribution depending on the number m of moment conditions increases. Moreover for any m the inequality

$D((MaxMinxEnt)_m : q) > D((MinMinxEnt)_m : q)$

takes place.

This result shows that both distributions can be applied in solving proper problems in survival data analysis.

In our investigation as components of K_0 characterizing moment vector functions $g_1(x) = x, g_2(x) = x^2, g_3(x) = ln\, x, g_4(x) = (ln\, x)^2, g_5(x) = ln(1 + x^2)$ are chosen.

Consequently, $K_0 = \{g_1, \cdots, g_5\}$. For example, if $m = 3$ then $\left(g_0, g^{(1)}\right) = (1, x, x^2, lnx), g^{(1)} \in K_{0,3}$ gives the least value to $U(g)$ and $\left(g_0, g^{(2)}\right) = (1, x^2, lnx, ln(1 + x^2))$, $g^{(2)} \in K_{0,3}$ gives the greatest value to $U(g)$.

The MinxEnt distributions corresponding to $(g_0, g), g_0(x) = 1, g \in K_{0,m}$, $m = 1, \cdots, 4$ and $D(p : q)$ values are shown in Table 3-6. By virtue of these tables are also obtained $(MinMinxEnt)_m$, $(MaxMinxEnt)_m$, $m = 1, \cdots, 4$ distributions and $D_{min} = D((MinMinxEnt)_m : q)$, $D_{max} = D((MaxMinxEnt)_m : q)$, $m = 1, \cdots, 4$ which are shown in Table 7, 8. It should be noted that the MinxEnt, $(MinMinxEnt)_m$, $(MaxMinxEnt)_m, m = 1, \cdots, 4$ distributions for the investigated data are determined by using MATLAB.

In order to obtain the performance of the mentioned distributions, we use various criteria as Root Mean Square Error (RMSE), Chi-Square, and Kullback-Leibler measures of distributions. The acquired results are demonstrated in Table 9 and Table 10.

In the sense of RMSE criteria each $(MinMinxEnt)_m$ $(m = 1,2,3)$ distribution is better than corresponding $(MaxMinxEnt)_m$ $(m = 1,2,3)$ distribution but, $(MaxMinxEnt)_4$ distribution is more suitable for statistical data than $(MinMinxEnt)_4$ distribution. These results also are corroborated by graphical representation (see Figure 1 (a), 1 (b)-4 (a), 4 (b)).

This result shows that $(MinMinxEnt)_4$ is better than the $(MaxMinxEnt)_4$ distribution in the sense of Kullback -Leibler measure.

TABLE 3: The $(MinxEnt)_1$ distribution corresponding to $(g_0,\ g),\ g_0(x)=1,\ g\in K_{0,1}$ and $D(p:q),\ D_{min},\ D_{max}$ values.

(g_0, g)	(g_0, g_1)	(g_0, g_2)	(g_0, g_3)	(g_0, g_4)	(g_0, g_5)
$(MinxEnt)_1$ Distribution	0.5029 0.1361 0.1366	**0.5039 0.1359 0.1362**	**0.5025 0.1366 0.1371**	0.5040 0.1355 0.1362	0.5027 0.1363 0.1369
	0.0579 0.0520 0.0338	**0.0577 0.0518 0.0337**	**0.0581 0.0520 0.0337**	0.0578 0.0520 0.0338	0.0580 0.0521 0.0338
	0.0247 0.0155 0.0187	**0.0246 0.0155 0.0187**	**0.0246 0.0154 0.0185**	0.0247 0.0155 0.0187	0.0246 0.0154 0.0185
	0.0219	**0.0221**	**0.0216**	0.0219	0.0216
$D(p:q)$	0.0260533472	**0.0260546553**	**0.0260387930**	0.0260478159	0.0260435739

TABLE 4: The $(MinxEnt)_2$ distribution corresponding to $(g_0,\ g),\ g_0(x)=1,\ g\in K_{0,2}$ and $D(p:q),\ D_{min},\ D_{max}$ values.

(g_0, g)	(g_0, g_1, g_2)	(g_0, g_1, g_3)	(g_0, g_1, g_4)	(g_0, g_1, g_5)	(g_0, g_2, g_3)
$(MinxEnt)_2$ Distribution	0.5034 0.1360 0.1363	0.5034 0.1359 0.1364	0.5028 0.1361 0.1367	**0.5036 0.1359 0.1362**	0.5033 0.1360 0.1364
	0.0578 0.0519 0.0337	0.0578 0.0520 0.0338	0.0579 0.0520 0.0338	**0.0577 0.0519 0.0338**	0.0578 0.0519 0.0337
	0.0247 0.0155 0.0187	0.0247 0.0155 0.0187	0.0247 0.0155 0.0187	**0.0247 0.0155 0.0187**	0.0247 0.0155 0.0187
	0.0220	0.0219	0.0218	**0.0220**	0.0220
$D(p:q)$	0.0260556309	0.0260545291	0.0260534019	**0.0260565687**	0.0260559433
(g_0, g)	(g_0, g_2, g_4)	(g_0, g_2, g_5)	(g_0, g_3, g_4)	(g_0, g_3, g_5)	(g_0, g_4, g_5)
$(MinxEnt)_2$ Distribution	0.5039 0.1361 0.1362	0.5035 0.1360 0.1363	0.5033 0.1358 0.1365	**0.5032 0.1359 0.1367**	0.5034 0.1357 0.1364
	0.0576 0.0517 0.0336	0.0578 0.0519 0.0337	0.0579 0.0521 0.0338	**0.0580 0.0521 0.0338**	0.0579 0.0520 0.0338
	0.0246 0.0155 0.0188	0.0247 0.0155 0.0187	0.0247 0.0155 0.0186	**0.0247 0.0155 0.0186**	0.0247 0.0155 0.0186
	0.0221	0.0220	0.0218	**0.0217**	0.0218
$D(p:q)$	0.0260553122	0.0260553622	0.0260499182	**0.0260457605**	0.0260492914

TABLE 5: The $(MinxEnt)_3$ distribution corresponding to $(g_0,\ g),\ g_0(x)=1, g\in K_{0,3}$ and $D(p:q),\ D_{min},\ D_{max}$ values.

(g_0, g)	(g_0, g_1, g_2, g_3)	(g_0, g_1, g_2, g_4)	(g_0, g_1, g_2, g_5)	(g_0, g_1, g_3, g_4)	(g_0, g_1, g_3, g_5)
$(MinxEnt)_3$ Distribution	**0.5032 0.1362 0.1364**	0.5030 0.1365 0.1365	0.5037 0.1357 0.1360	0.5031 0.1366 0.1364	0.5030 0.1368 0.1363
	0.0578 0.0518 0.0337	0.0577 0.0518 0.0336	0.0577 0.0520 0.0339	0.0576 0.0517 0.0336	0.0576 0.0517 0.0336
	0.0246 0.0155 0.0187	0.0246 0.0155 0.0187	0.0248 0.0156 0.0187	0.0246 0.0155 0.0188	0.0246 0.0155 0.0188
	0.0220	0.0221	0.0219	0.0221	0.0221
$D(p:q)$	**0.0260561615**	0.0260580951	0.0260581141	0.0260617141	0.0260634885
(g_0, g)	(g_0, g_2, g_3, g_5)	(g_0, g_2, g_3, g_5)	(g_0, g_2, g_3, g_5)	(g_0, g_2, g_3, g_5)	(g_0, g_2, g_3, g_5)
$(MinxEnt)_3$ Distribution	0.5031 0.1367 0.1364	0.5031 0.1364 0.1365	0.5030 0.1366 0.1365	0.5032 0.1364 0.1365	**0.5030 0.1369 0.1362**
	0.0576 0.0517 0.0336	0.0577 0.0518 0.0336	0.0577 0.0517 0.0336	0.0577 0.0518 0.0336	**0.0575 0.0516 0.0336**
	0.0246 0.0155 0.0188	0.0246 0.0155 0.0187	0.0246 0.0155 0.0187	0.0246 0.0155 0.0187	**0.0246 0.0155 0.0188**
	0.0221	0.0221	0.0221	0.0221	**0.0221**
$D(p:q)$	0.0260624197	0.0260574362	0.0260587230	0.0260571954	**0.0260667248**

From Table 3 it is seen that the MinMaxEnt (the MaxMaxEnt) distribution is realized by vector function $(g_0, g_3) = (1, lnx)$ $\left((g_0, g_2) = (1, x^2)\right)$ and

$D_{min} = D((MinMaxEnt)_1 : q) = 0.0260387930,$

$D_{max} = D((MaxMaxEnt)_1 : q) = 0.0260546553.$

From Table 4 it is seen that the MinMaxEnt (the MaxMaxEnt) distribution is realized by vector function $(g_0, g_3, g_5) = (1, lnx, ln(1+x^2))$ $\left((g_0, g_1, g_1) = (1, x, ln(1+x^2))\right)$ and

$D_{min} = D((MinMaxEnt)_2 : q) = 0.0260457605,$

$D_{max} = D((MaxMaxEnt)_2 : q) = 0.0260565687.$

TABLE 6: The $(MinxEnt)_4$ distribution corresponding to $(g_0,\ g), g_0(x) = 1, g \in K_{0,4}$ and $D(p:q),\ \ D_{min},\ D_{max}$ values.					
(g_0, g)	$(g_0, g_1, g_2, g_3, g_4)$	$(g_0, g_1, g_2, g_3, g_5)$	$(g_0, g_1, g_2, g_4, g_5)$	$(g_0, g_1, g_3, g_4, g_5)$	$(g_0, g_2, g_3, g_4, g_5)$
$(MinxEnt)_4$ Dist.	**0.5029**	0.5029	0.5029	**0.5029**	0.5029
	0.1375	0.1375	0.1375	**0.1374**	0.1375
	0.1354	0.1354	0.1354	**0.1355**	0.1355
	0.0572	0.0573	0.0573	**0.0573**	0.0573
	0.0517	0.0518	0.0518	**0.0518**	0.0518
	0.0339	0.0339	0.0339	**0.0339**	0.0339
	0.0250	0.0249	0.0249	**0.0249**	0.0249
	0.0157	0.0157	0.0157	**0.0156**	0.0157
	0.0188	0.0188	0.0188	**0.0188**	0.0188
	0.0217	0.0218	0.0218	**0.0218**	0.0218
$D(p:q)$	**0.0260843817**	0.0260829118	0.0260833086	**0.0260786357**	0.0260807828

TABLE 7: Distributions of $(MinMinxEnt)_m,\ \ m = 1,2,3,4.$					
t	p_i^*	$(MinMinxEnt)_1$	$(MinMinxEnt)_2$	$(MinMinxEnt)_3$	$(MinMinxEnt)_4$
0-1	0.5030	0.5025	0.5032	0.5032	0.5029
1-2	0.1369	0.1366	0.1359	0.1362	0.1374
2-3	0.1369	0.1371	0.1367	0.1364	0.1355
3-4	0.0565	0.0581	0.0580	0.0578	0.0573
4-5	0.0506	0.0520	0.0521	0.0518	0.0518
5-6	0.0357	0.0337	0.0338	0.0337	0.0339
6-7	0.0238	0.0246	0.0247	0.0246	0.0249
7-8	0.0149	0.0154	0.0155	0.0155	0.0156
8-9	0.0208	0.0185	0.0186	0.0187	0.0188
9-10	0.0208	0.0216	0.0217	0.0220	0.0218

TABLE 8: Distributions of $(MaxMinxEnt)_m,\ \ m = 1,2,3,4.$					
t	p_i^*	$(MaxMinxEnt)_1$	$(MaxMinxEnt)_2$	$(MaxMinxEnt)_3$	$(MaxMinxEnt)_4$
0-1	0.5030	0.5039	0.5036	0.5030	0.5029
1-2	0.1369	0.1359	0.1359	0.1369	0.1375
2-3	0.1369	0.1362	0.1362	0.1362	0.1354
3-4	0.0565	0.0577	0.0577	0.0575	0.0572
4-5	0.0506	0.0518	0.0519	0.0516	0.0517
5-6	0.0357	0.0337	0.0338	0.0336	0.0339
6-7	0.0238	0.0246	0.0247	0.0246	0.0250
7-8	0.0149	0.0155	0.0155	0.0155	0.0157
8-9	0.0208	0.0187	0.0187	0.0188	0.0188
9-10	0.0208	0.0221	0.0220	0.0221	0.0217

TABLE 9: The obtained results for $(MinMinxEnt)_m,\ \ m = 1,2,3,4.$				
$(MinMinxEnt)_m$ **Distribution**	$D(p:q)$	**Calculated value of Chi - Square**	**Table value of Chi – Square**	**RMSE**
$(MinMinxEnt)_1$	0.0260387930	0.0118	$\chi^2_{8,\alpha} = 15.51$	0.0338
$(MinMinxEnt)_2$	0.0260457605	0.0116	$\chi^2_{7,\alpha} = 14.07$	0.0333
$(MinMinxEnt)_3$	0.0260561615	0.0112	$\chi^2_{6,\alpha} = 12.59$	0.0329
$(MinMinxEnt)_4$	0.0260786357	0.0103	$\chi^2_{5,\alpha} = 11.07$	0.0317

TABLE 10: The obtained results for $(MaxMinxEnt)_m,\ \ m = 1,2,3,4.$				
$(MaxMinxEnt)_m$ **Distribution**	$D(p:q)$	**Calculated value of Chi – Square**	**Table value of Chi – Square**	**RMSE**
$(MaxMinxEnt)_1$	0.0260546553	0.0115	$\chi^2_{8,\alpha} = 15.51$	0.0332
$(MaxMinxEnt)_2$	0.0260565687	0.0112	$\chi^2_{7,\alpha} = 14.07$	0.0329
$(MaxMinxEnt)_3$	0.0260667248	0.0107	$\chi^2_{6,\alpha} = 12.59$	0.0323
$(MaxMinxEnt)_4$	0.0260843817	0.0104	$\chi^2_{5,\alpha} = 11.07$	0.0320

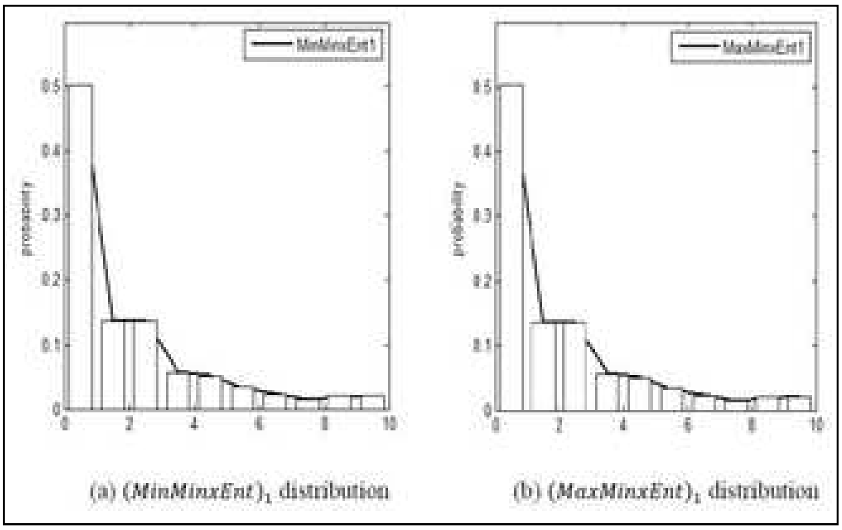

FIGURE 1: (a) $(MinMinxEnt)_1$ distribution; **(b)** $(MaxMinxEnt)_1$ distribution.

From Table 5 it is seen that the MinMaxEnt (the MaxMaxEnt) distribution is realized by vector function $(g_0, g_1, g_2, g_3) = (1, x, x^2, lnx)\big((g_0, g_2, g_3, g_5) = (1, x^2, lnx, ln(1 + x^2))\big)$ and

$D_{min} = D((MinMaxEnt)_3: q) = 0.0260561615,$

$D_{max} = D((MaxMaxEnt)_3: q) = 0.0260667248.$

TABLE 11: Survival Analysis by $(MinMinxEnt)_4$.

t	n_i	d_i	c_i	$\hat{f}(t) = (MinMinxEnt)_4$	$\hat{F}(t)$	$\hat{S}(t)$	$\hat{h}(t) = \frac{\hat{f}(t)}{\hat{S}(t)}$
1	388	167	2	0.5029	0.5029	0.4971	1.0117
2	219	45	1	0.1374	0.6403	0.3597	0.3820
3	173	45	1	0.1355	0.7758	0.2242	0.6044
4	127	19	0	0.0573	0.8331	0.1669	0.3433
5	108	17	0	0.0518	0.8849	0.1151	0.4500
6	91	11	1	0.0339	0.9188	0.0812	0.4175
7	79	8	0	0.0249	0.9437	0.0563	0.4423
8	71	5	0	0.0156	0.9593	0.0407	0.3833
9	66	6	1	0.0188	0.9781	0.0219	0.8584
10	59	7	0	0.0218	0.9999	0.0001	--

TABLE 12: Survival Analysis by $(MaxMinxEnt)_4$.

t	n_i	d_i	c_i	$\hat{f}(t) = (MaxMinxEnt)_4$	$\hat{F}(t)$	$\hat{S}(t)$	$\hat{h}(t) = \frac{\hat{f}(t)}{\hat{S}(t)}$
1	388	167	2	0.5029	0.5029	0.4971	1.0117
2	219	45	1	0.1375	0.6404	0.3596	0.3824
3	173	45	1	0.1354	0.7758	0.2242	0.6039
4	127	19	0	0.0572	0.8330	0.1670	0.3425
5	108	17	0	0.0517	0.8847	0.1153	0.4484
6	91	11	1	0.0339	0.9186	0.0814	0.4165
7	79	8	0	0.0250	0.9436	0.0564	0.4433
8	71	5	0	0.0157	0.9593	0.0407	0.3857
9	66	6	1	0.0188	0.9781	0.0219	0.8584
10	59	7	0	0.0217	0.9998	0.0002	--

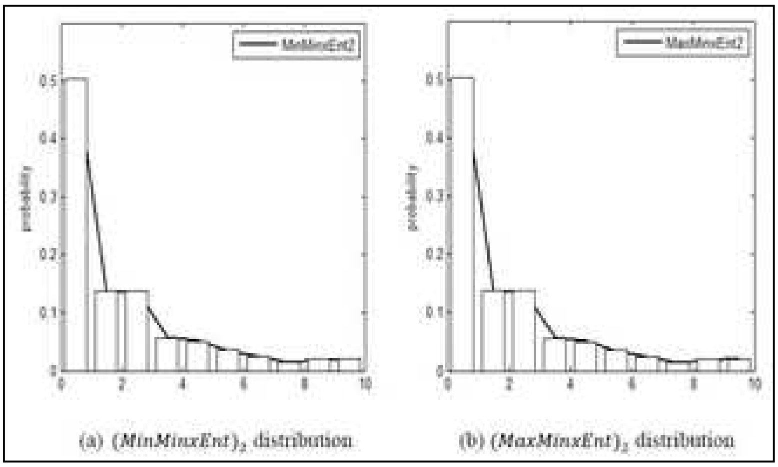

FIGURE 2: (a) $(MinMinxEnt)_2$ distribution; **(b)** $(MaxMinxEnt)_2$ distribution.

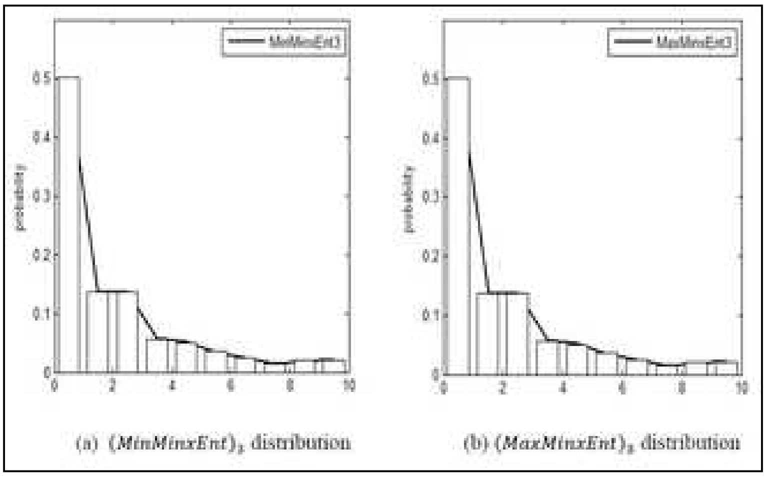

FIGURE 3: (a) $(MinMinxEnt)_3$ distribution; **(b)** $(MaxMinxEnt)_3$ distribution.

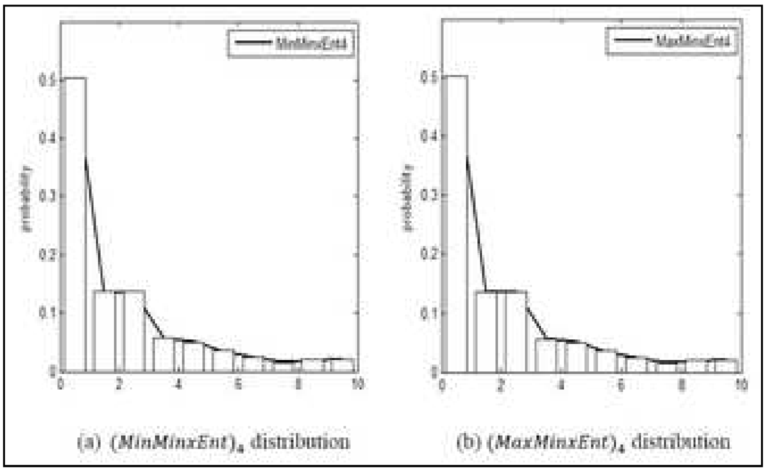

FIGURE 4: **(a)** $(MinMinxEnt)_4$ distribution; **(b)** $(MaxMinxEnt)_4$ distribution.

From Table 6 it is seen that the MinMaxEnt (the MaxMaxEnt) distribution is realized by vector function $(g_0, g_1, g_3, g_4, g_5) = (1, x, ln\,x, (ln\,x)^2, ln(1+x^2))$ $\left((g_0, g_1, g_2, g_3, g_4) = (1, x, x^2, ln\,x, (ln\,x)^2)\right)$ and

= (()$_4$:) =0.0260786357,

= ((()$_4$:) = 0.0260843817).

PROBABILITY DENSITY FUNCTION, CUMULATIVE DISTRIBUTION FUNCTION, SURVIVAL FUNCTION AND HAZARD FUNCTION BY GEOM

In this section survival data analysis is conducted by ()$_4$ (()$_4$) distribution since the above acquired investigations ()$_4$ (()$_4$) is more presentable for survival data among ()(()), = 1,2,⋯,4 distributions.

Table 11 (Table 12) shows the MinMinxEnt (the MaxMinxEnt) estimators of Probability Density Function̂(), Cumulative Distribution Function̂(), Survival Function̂() and Hazard Function. On basis of the results given in Table 11 (Table 12), graphics of̂(),̂(),̂() and $\hat{h}$() are demonstrated in Figure 5 (a) -5 (d) (Figure 6 (a)-6 (d)).

CONCLUSION

As above it is noted that, Entropy Optimization Methods (EOM) have important applications, especially in statistics, economy, engineering and etc. There are several examples in the literature that known statistical data do not conform to theoretical distributions, however do conform the entropy optimization distributions well. Generalized Entropy Optimization Methods (GEOM) have suggested distributions in the form of the MinMinxEnt, the MaxMinxEnt which are closest to statistical data and furthest from statistical data in the sense of Kullback-Leibler measures, respectively.

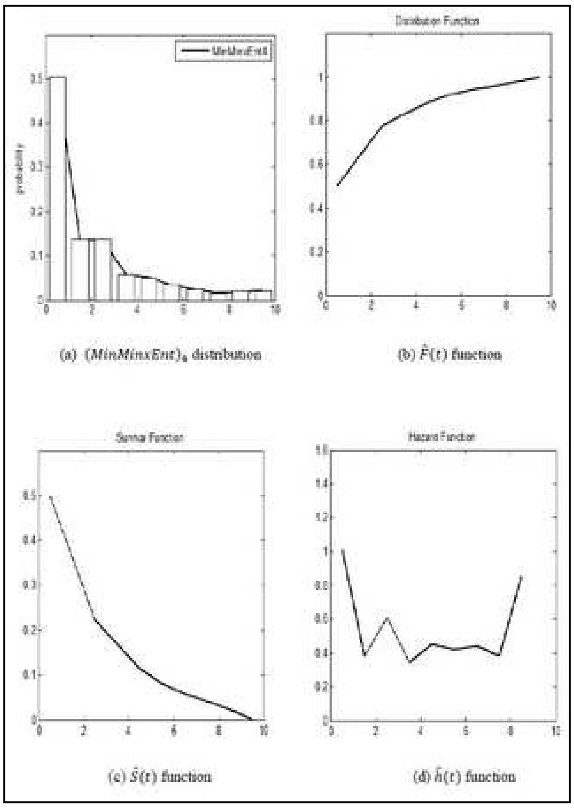

FIGURE 5: (a) $(MinMinxEnt)_4$ distribution **(b)** $\hat{F}(t)$ function **(c)** $\hat{S}(t)$ function **(d)** $\hat{h}(t)$ function

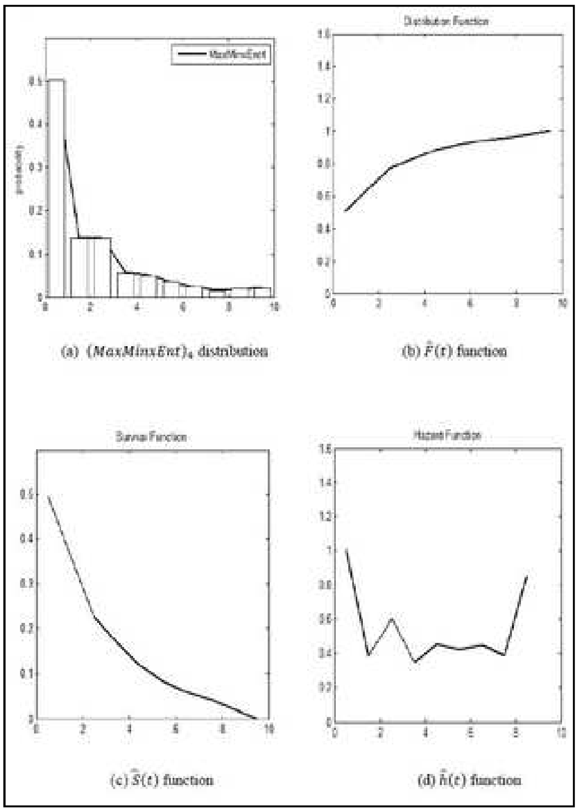

FIGURE 6: (a) $(MaxMinxEnt)_4$ distribution **(b)** $\hat{F}(t)$ function **(c)** $\hat{S}(t)$ function **(d)** $\hat{h}(t)$ function

The presence of censoring in the survival times leads to situation that for the survival data the sum of observation probabilities stands less than 1. For this reason in solving many problems it is required to supplement the sum of observation probabilities up to 1. Consequently, the basing of the admissibility of this method of supplementation acquires a new significance. Mentioned problem is solved by applying the MinMinxEnt and MaxMinxEnt methods so that among Entropy Optimization Distributions with respect to the Kullback-Leibler measure it is chosen such which is closest to observed probability distribution.

In this study, it is shown that $(MinMinxEnt)_4$ and $(MaxMinxEnt)_4$ distributions more successfully represent Survival Data. Furthermore, in the sense of RMSE criteria $(MaxMinxEnt)_4$ distribution is more suitable for statistical data than $(MinMinxEnt)_4$ distribution. These results are also corroborated by graphical representation. At the same time $(MinMinxEnt)_4$ is better than $(MaxMinxEnt)_4$ distribution in the sense of Kullback-Leibler measure. Our investigation indicates that GEOM in survival data analysis yields reasonable results.

Conflict of Interest

Authors declared no conflict of interest.

Financial Support

This study was supported by Anadolu University Scientific Research Project Council (AU BAP, Project No: 1208F134).

Authorship Contributions

Idea/Concept: Constructing the hypothesis or idea of research and/or article: *Aladdin Shamilov;* ***Design: Planning methodology to reach the conclusions:*** *Aladdin Shamilov, Sevda Özdemir;* ***Control/Supervision: Organizing, supervising the course of progress and taking the responsibility of the research/study:*** *Aladdin Shamilov, Sevda Özdemir;* ***Data Collection and/or Processing: Taking responsibility in patient follow-up, collection of relevant biological materials, data management and reporting, execution of the experiments:*** *Sevda Özdemir;* ***Analysis and/or Interpretation: Taking responsibility in logical interpretation and conclusion of the results;*** *Sevda Özdemir, Aladdin Shamilov;* ***Literature Review: Taking responsibility in necessary literature review for the study:*** *Aladdin Shamilov, Sevda Özdemir;* ***Writing the Article: Taking responsibility in the writing of the whole or important parts of the study:*** *Aladdin Shamilov, Sevda Özdemir;* ***Critical Review: Reviewing the article before submission scientifically besides spelling and grammar:*** *Aladdin Shamilov, Sevda Özdemir;* ***References and Fundings: Providing personnel, environment, financial support tools that are vital for the study:*** *Aladdin Shamilov, Sevda Özdemir.*

REFERENCES

1. Венцель ЕС. Теория вероятностей, МОСКВА 1969;148.
2. Shamilov A. A development of entropy optimization methods. WSEAS Transactions on Mathematics 2006;5:568-75.
3. Shamilov A. Generalized entropy optimization problems and the existence of their solutions. PHYSICA A 2007;382(2):465-72.
4. Shamilov A.. Entropy, Information and Entropy Optimization. 1.baskı. Nobel Publication, Ankara; 2015. p.1-183.
5. Shamilov A. Generalized entropy optimization problems with finite moment functions sets. J Stat Manag Syst 2010;13(3):595-603.
6. Shamilov A, Giriftinoglu C, Usta I, Mert Kantar Y. A new concept of relative suitability of moment function sets. JAMC 2008;206:521-9.
7. Kaminski D, Geisler C. Survival analysis of faculty retention in science and engineering by gender. Science 2012;335(6070):864-6.

8. Reingold EM, Reichle ED, Glaholt MG, Sheridan H. Direct lexical control of eye movements in reading: evidence from a survival analysis of fixation durations. Cogn Psychol 2012;65(2):77-206.
9. Wang H, Dai H, Fu B. Accelerated failure time models for censored survival data under referral bias. Biostatistics 2012;14(2):313-26.
10. Ebrahimi N. The maximum entropy method for lifetime distributions. The Indian Journal of Statistics 2000;62(2):236-43.
11. Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Med Res Methodol 2012;12:9.
12. Joly P, Gerds TA, Qvist V, Commenges D, Keiding N. Estimating survival of dental fillings on the basis of interval-censored data and multi-state models. Stat Med 2012;31(11-12):1139-49.
13. Lee ET, Wang JW. Statistical Methods for Survival Data Analysis. 3rded. Hoboken: Wiley-Interscience; 2003. p.534.
14. Deshpande JV, Purohit SG. Series on quality, reliability and engineering statistics. Life Time Data: Statistical Models and Methods. 1sted. Singapore: World Scientific Publ; 2005. p.11.
15. Kapur JN, Kesavan HK. Entropy Optimization Principles with Applications. 1sted. Boston: Academic Press; 1992. p.408.