Iraqi Journal for Administrative Sciences
المجلة العراقية للعلوم الإدارية
ISSN P.1818-1074 ISSN E.2958-6941
Vol. 19 issue76 Jun. 2023

تعظيم الاحتمال لتقدير بعض توزيعات العائلة الأسية

Maximization of Likelihood Estimations for Some Exponential Distributions Family

أ.م.د.عباس كولمراد بك مراد
Dr. Abbas Gulmurad Beg Murad
Email: abbas.beg@univsul.edu.iq
Phone: +9647702671373

أ.م.د.سميرة محمد صالح
Dr. Samira Muhamad salh
Email: samira.muhamad@univsul.edu.iq
Phone: +9647701441226

جامعة السليمانية
University of Sulaimaniyah

الملخص

في هذه الورقة البحثية، نحاول تشخيص بعض التوزيعات الأسية ، والتي تتمثل بالتوزيع الطبيعي كمعيار مع توزيعين آخرين (توزيع جاما والتوزيع الأسي)، حيث نحاول تحليل تأثير عدد المعلمات سواء كانت معلمتين اثنين أو معلمة واحدة في التوزيع على تعظيم احتمالية التقدير للتوزيعات الثلاثة و كذلك تاثير حجم العينة على تعظيم احتمال التقدير (MLE's) للمعلمات؛تم الحصول على النتائج باستعمال بعض الحزم مثل (fitdistrplus) بالاضافة لبعض الخوارزميات و الدوال في لغة (R) التي تعتمد على سلوك بيانات المحاكاة مثل عدد المعلمات في التوزيع وكذلك حجم العينة، وتم إجراء مقارنة بين التوزيعات المذكورة بناءً على اختبار kolmogrov-smirnov وكذلك اختبار Grammer von، للمفاضلة على التوزيع المناسب لسلوك البيانات بشكل أفضل من التوزيعات الأخرى، حيث تمثلت مقدمة جيدة لبعض التوزيعات في تلخيص العلاقات فيما بينها وكيفية حساب ورسم دالة الكثافة الاحتمالية في التوزيع الطبيعي.

الكلمات المفتاحية: تعظيم التقديرات الاحتمالية (MLE، توزيعات العائلة الأسية ، دالة الكثافة الاحتمالية ، الحزمة البرمجية fitdistrplus، اختبار .Grammer von)

ABSTRUCT

In this paper, we attempt to diagnosis of some of exponential distributions, represent by the normal distribution as a criteria with two others distributions (Gamma & Exponential) distributions we attempt to analyze the effect of the number of parameter(s) two or one parameter(s) in the distribution on the maximize likelihood estimation for the three distributions. That proved the maximum likelihood estimates (MLE's) of the parameters; it's obtained by using some packages (e.g. fitdistrplus) insides some algorithms in (R) that based on behavior of the simulated data like parameter and sample size. A comparison is carried out between the mentioned distributions based on the classical kolmogrov-smirnov distance to minimize distance estimation test statistic and Grammer von misses distance, to emphasize that which distribution fitted to the data better than the other models. After completing this, A good preface to some distributions is to summarize the relationships between observations. How to calculate and Plot Probability and Density Functions the normal distribution.

Keywords: Maximization of likelihood estimations (MLE), exponential family, density functions, fitdistrplus package, Grammer von misses distance.

Aim of research

1) Diagnosis of some of exponential distributions, represent by the normal (Gaussian) distribution as a criteria with two others distributions (Gamma & Exponential) distributions.

2) Attempt to analyze the effect of a high or low number of the parameters in the distribution on the maximum likelihood estimation of the distribution for the three distributions.

Literature Review

In 2006 (Rameshwar D. Gupta and Debasis Kundu) worked on a distribution named as (generalized function distribution) or exponential distribution. It is notice in many cases, the generalized exponential distribution can be substituted for the gamma distribution. It is detect that for a precondition gamma distribution there exists a general function distribution so that the two distribution functions are almost identical, we observed for practical purposes, approximate gamma random samples can be generated using the generalized exponential distribution, and statistical tests cannot be used to distinguish random samples obtained in this way. Moreover, if skewed data set found where gamma distribution fits very well, the generalized exponential distribution also used.

In 2013 researcher's (Gauss M. Cordeiroa, Edwin M.M. Ortegab and Artur J. Lemontea) used moment method and mle with exponential-Weibull distribution derive obvious expressions for the generalized ordinary moments and a general formula for the incomplete moment based on infinite sums of Meijer's G functions.at the end concluded that the resilience of the new distribution it is clarifying by mean of the an application to real data .

Also in 2014 two researchers (Eugene F. Fama and Hichard) takes several steps to alleviate data analysis problems caused by the fact that the underlying expressions for density and cumulative displacement functions (c.d.f.) apply to the most stable distributions. Develop numerical approximations of c.d.f. and inverse functions of c.d.f. of symmetric stable distributions.

Also (jimmy Reyes, Emilio Gómez-Déniz, Héctor W. Gómez and Enrique Calderín-Ojeda), in 2021 used multi Classical exponential distribution function generalizations in the statistical literature have proven useful in many situations, and in this work an attempt was made to fill in the gaps in the literature by presenting a series of distributions that could be unimodal or bimodal, and overlapping null exponential distributions.

In 2021 (Mathias Niepert,Pasquale Minervini and Luca Franceschi) published the paper and demonstrate that Implicit Maximum Likelihood Estimation (I-MLE) is a comprehensive learning framework for models that combine discrete family distributions with differentiable neural components, using experiments on multiple datasets, and demonstrate that I-MLE is competitive Powerful and generally better than current methods relying on problem-specific relaxations, we show that I-MLE minimizes likelihood estimates when used in some learning contexts related to recently studied synthetic solutions.

1-Introduction

There are most powerful distributions in statistical literature available to use for analyzing of data in various scientific fields that require an appropriate model.
The data modeling are very important, in which the quality of the outputs of statistical analysis depends heavily on the assumed model. Therefore, it is necessary to find more nearer distributions to get accurate result for this. Multi studies have been made to define new categories of univariate continuous distributions by adding some parameters to the base model that extend known distributions thereby providing more flexibility in the model data. A statistical distribution is a mathematical increasing function has parameter(s) that gives the probability outcomes for a random variable. There are two types of it (discrete and continuous distributions) based on the random values it represents. This work will introduce the three important statistical distributions, show there, and discuss the relationships among different distributions and their applications [4,9] .

2- Density Functions and probability distributions [9,14]

The distribution provides a statistical distribution function that can be used to calculate the probability of any single observation in the sample space. This distribution describes the aggregation or density of observations and is called a probability density function. A summary of these relationships between observations is called a cumulative density function (C.D.F).A distribution is simply a collection of data, or scores, on a variable. Usually, these data are order from smallest to largest and then they can be presented graphically. The probability distribution is also a basic concept in statistics. They are used on both a theory and practical fields. Some practical uses of probability distributions are calculating critical regions for hypothesis testing it is often useful to model a reasonable distribution of data and calculate, confidence bounds of parameters.

3- Estimation and estimate [10,16]

Any statistic T(x) that used to estimate some function of unknown (θ) , say t(θ)is called an estimator and the value of that statistic is called an estimate. An estimator is always a statistic which is both random variable and function thus the word of estimator. Stand with the function for example ($\overline{X}$) is an estimator of μ, also (S) is an estimator of (σ^2).

3-1- Methods of estimation [10,15]

Let x_1 ,.... x_n be a random sample from a Distribution with probability density function p.d.f of [f (x: θ)] suppose it is required to estimate the parameter (θ) a several methods are variable for estimating the parameters same of them are:

1. Method of moments
2. Maximum likelihood method
3. Bayesian method
4. Ordinary least square method

In this paper the second method used:

3-1-1- Maximum likelihood estimation method [10,15,16]

This method is used to estimating the parameter (θ), the concept of this method depend on likelihood function which can be defined as follow:

Let $x_1, x_2, x_3 \ldots x_n$ be a random sample from Distribution That has p.d.f [f(x, θ)] ,the likelihood function L(θ ,x1 ,.... xn) is a joint probability density of the r.s (x1 ,.... xn).

$$L(\theta, x_1 \ldots x_n) = f(x_1 \ldots x_n; \theta)$$
$$= \mathrm{f}(\mathrm{x}_1, \theta).\mathrm{f}(\mathrm{x}_2, \theta).\mathrm{f}(\mathrm{x}_3, \theta)\ldots.\mathrm{f}(\mathrm{x}_n, \theta) \qquad \ldots(1)$$
$$= \prod_{i=1}^{n} f(x;\theta)$$

In this method we choose as estimate (value of the parameter θ) that maximize L(θ).

$L(\hat{\theta}, x_1 \ldots x_n) = \max[L(\beta; x_1 \ldots x_n)]$ Then $\hat{\theta}$ is called a MLE for (θ) .

To finding out MLE to the unknown parameter (θ), it will be the solution of the following equation.

$$\frac{\partial \ln L(\theta)}{\partial \theta} = 0 \qquad \text{and} \qquad \frac{\partial^2 \ln L(\theta)}{\partial \theta^2} < 0 \qquad \ldots.(2)$$

It was noticed that the logarithm easier to use for function of the likelihood, there more ln(θ) or L(θ) have their maximum at the same value of θ .and for the case of more than one parameter (two parameter) or multiple parameters ,we find the MLE for the parameters (θ_1, θ_2) by differentiate the likelihood function with respect to θ_1 and θ_2 put in equal to zero.

$$\frac{\partial\ L(\theta, x)}{\partial \theta_1} < 0 \qquad \frac{\partial\ L(\theta, x)}{\partial \theta_2} < 0 \qquad \ldots.(3)$$

The solution of these equation at one the MLE for θ_1 and θ_2

These solutions will be maximized if

$$\frac{\partial^2 L(\theta; x)}{\partial \theta_1} < 0 \qquad \frac{\partial^2 L(\theta; x)}{\partial \theta_2} < 0$$
$$and$$
$$\left(\frac{\partial^2 L(\theta; x)}{\partial \theta_1}\right)\left(\frac{\partial^2 L(\theta; x)}{\partial \theta_2}\right) > \left(\frac{\partial^2 L(\theta; x)}{\partial \theta_1 \partial \theta_2}\right)^2 \qquad \ldots(4)$$

4- The exponential family or exponential class of distribution [8,10,14]

In statistic most important family of distribution called the exponential family (class). This class possesses complete and sufficient statistics which are readily determined from the distribution. Consider a family {f(x; θ) : θ ∈ Ω} of probability density or mass functions, where Ω is the interval set $\Omega=\{\mu:\gamma\prec\mu\prec\delta\}$,where (γ and δ) are known constants between ($\pm\infty$)

$$f(x;\mu)=\begin{cases}\exp\left[p(\mu)K(x)+H(x)+q(\mu)\right] & x\in S\\ 0 & otherwise\end{cases} \quad ...(5)$$

Where (S) is the support of X. A pdf of the equation (5) is said to be a member of the regular exponential class of probability density or mass functions if:

1. S, the support of X, does not depend upon θ.

2. p(θ) is a nontrivial continuous function of θ ∈ Ω

3. Finally, (a) if X is a continuous random variable, then each of K (x) ≡ 0 and H(x) is a continuous function of x ∈ S.

(b) if X is a discrete random variable, then K(x) is a nontrivial function of x ∈ S.

In this work we concerned with same distribution which is a member of these class such as (normal distribution (Gaussian, exponential, and gamma distribution):

4-1- Gaussian (normal) distribution [4]

The normal distribution, named for Carl Friedrich Gauss, is the focus of much of the field of statistics. Surprisingly, data from many fields of study can be described using a Gaussian distribution, so much so that this distribution is often called the (normal) distribution because it is so common. The Gaussian distribution can be described by two parameters: mean: represented by the lowercase Greek letter mu, which is the expected value of the distribution.
The Variance of the data: It is denoted by the lowercase Greek letter sigma, raised to the power of the second (because the unit of the variable is squared), describing the distribution of observations from the mean. A common practice is to use a standard deviation calculation called the standard deviation: denoted by the Greek letter lowercase sigma, which describes the normal distribution of observations from the mean. The regular probability density function(pdf) can be used to create a Gaussian probability density function with a given space of the sample, the mean, and the variance.

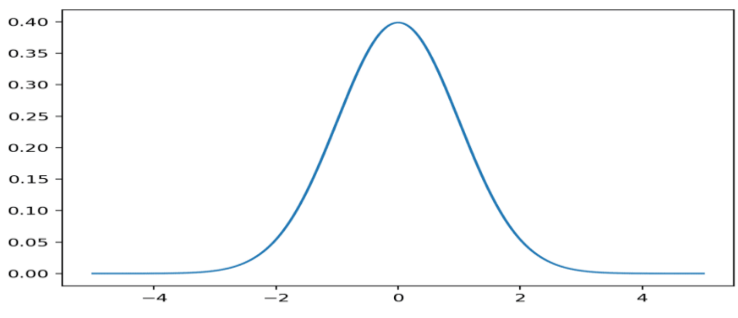


Fig. (1) Gaussian probability density function

The probability density function of this distribution writes as follow:

$$f(x,\mu,\sigma)=\frac{1}{\sqrt{2\pi\sigma}}e^{-\frac{(x-\mu)}{2\sigma}}\quad where -\infty\le x\le\infty$$
$$-\infty\le\mu\le\infty\quad and\quad \sigma\succ 0\quad ...(6)$$

4-2- Gamma Distribution [8,11,15]

In fact, the gamma distribution is derived from the gamma function, or what is occasionally named gamma integral, which is mentioned in many advanced mathematics books. This function of distribution can be considered as one of the major distributions employed to studying problems in which time is one of its factors, such as those specialized studies on the length of the operating time of equipment the particular factory. it is also considered one of the most important distributions that enter the subject of reliability and it is known as a mathematical quantization in the following form:

$$\Gamma\alpha=\int_0^\infty y^{\alpha-1}e^{-\lambda y}dx \qquad for\ \alpha,\lambda\succ 0 \qquad ...(7)$$

And we can show the gamma function in this figure below:

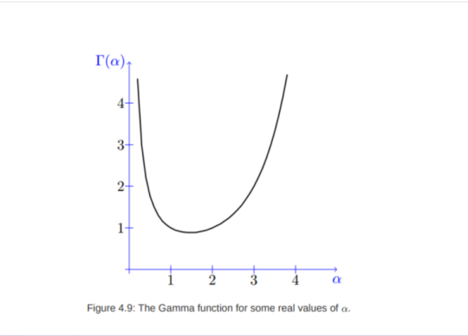


Fig. (2) shows the Gamma function for positive real values

Since (Γ(α)) represents the value of the gamma integral at a certain value of (α) and this integral is convergent for all the value of ($\alpha\le 0$) and divergent for the values of alpha, for example when ($\alpha=1$), so

$$\Gamma 1=\int_0^\infty e^{-x}dx=1 \qquad ...(8)$$

Note here that the integral is convergent and if ($\alpha=0$) then

$$\Gamma 0=\int_0^\infty y^{-1}e^{-x}dx \qquad ...(9)$$

And by using integration part, we notice that $\Gamma 0=\lim_{x\to\alpha}e^{-x}Lnx-\lim_{x\to\alpha}e^{-x}Lnx$

And this means that the form is indefinite, that is, the divergent integration. Also, $\Gamma\alpha, \alpha \succ 0$ is a positive number ,And by dividing both sides of the gamma function by $\Gamma\alpha$ we get : $1 = \frac{1}{\Gamma\alpha}\int_0^\infty y^{\alpha-1}e^{-\lambda y}dx$ ...(10)

This means (x) that a random variable has a probability density function :

$$f(x;\alpha) = \begin{cases} \frac{1}{\Gamma\alpha}\int_0^\infty y^{\alpha-1}e^{-\lambda x}dx & ; x \succ 0 \\ 0 & otherwise \end{cases} \quad .....(11)$$

This is called the first form of the gamma distribution function with the parameter ($\alpha \succ 0$).And with symbols $x \sim G(\alpha)$.

It is clear from this that the exponential distribution with the parameter ($\theta = 1$) is a special case of the gamma distribution when ($\alpha = 1$) . There is another form of the gamma distribution function that is derived from the first form, which is the following:

$$\Gamma\alpha = \int_0^\infty \left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\frac{x}{\beta}} \frac{1}{\beta} dx$$

$$= \frac{1}{\beta}\int_0^\infty x^{\alpha-1}e^{-\frac{x}{\beta}} dx \quad ...(12)$$

Divide two sides by ($\Gamma\alpha$) we get :

$$1 = \frac{1}{\Gamma\alpha\beta^\alpha} x^{\alpha-1}e^{-\frac{x}{\beta}} dx \;\; ; x \succ 0$$

This means that (x) a random variable has a probability density function

$$f(x;\alpha,\beta) = \frac{1}{\Gamma\alpha\beta^\alpha} x^{\alpha-1}e^{-\frac{x}{\beta}} dx \;\; ; x \succ 0 \quad ...(13)$$

And this is the second form of the gamma function with two parameters $(\alpha, \beta \succ 0)$ denoted by $x \sim G(\alpha, \beta)$, It is clear that the exponential distribution is a special case of the gamma distribution when $\alpha = 1, \theta = \frac{1}{\beta}$, It can be shown that the value of the gamma integral is $\Gamma\alpha = (n-1)\,!$ and as follows :

$\Gamma\alpha = \int_0^\infty y^{\alpha-1}e^{-\lambda x}dx$,by using integration part and assuming $u = y^{\alpha-1}$,

$dv = e^{-x}dx$ then $du = (\alpha-1)x^{\alpha-2} \; dx$, $v = -e^{-x}$

Consequently

$$\left\{\begin{aligned} \Gamma\alpha &= \int_0^\infty x^{\alpha-1}e^{-x}\,dx \\ &= \left[-x^{\alpha-1}e^{-y}\right]_0^\infty + (\alpha-1)\int_0^\infty x^{\alpha-2}e^{-x}\,dx \\ &= (\alpha-1)\int_0^\infty x^{\alpha-2}e^{-x}\,dx \\ &= (\alpha-1)\Gamma(\alpha-1) \end{aligned}\right\} \quad .....(14)$$

By same way can be observed $\Gamma(\alpha-1)=(\alpha-2)\Gamma(\alpha-2)$

i.e that $\Gamma(\alpha)=(\alpha-1)(\alpha-2)\Gamma(\alpha-2)$ And if the case continues with retail integration by part, then $\Gamma(\alpha)=(\alpha-1)(\alpha-2)(\alpha-3).....3.2.1.0 \quad =(\alpha-1)$

4-3- Exponential Distribution [2,13]

The exponential distribution is one of the widely used as continuous distributions. Often used to model the elapsed time between events we will attempt to define the exponential distribution mathematically and derive its expected value. we will then develop a distributional intuition and discuss its many interesting properties .A continuous random variable X is said to have an exponentially distribution with parameter λ > 0, shown as X ~ Exponential (λ), if its PDF is given by :

$$f(x)=\left\{\begin{matrix} \lambda e^{-\lambda x} & x \succ 0 \\ 0 & otherwise \end{matrix}\right\} \quad ...(15)$$

The PDF of the exponential distribution over multiple values λ.

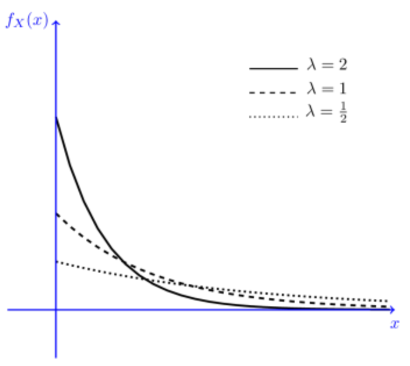


Fig. (3) Probability density function of exponential random variable

The exponential distribution has an important property is that it can be considered as a continuous dual of the distribution of geometric.

5- Convergence property of the normal distribution [6,10]

Assume that $X \sim G(\alpha)$ and that $Z=\frac{x-\alpha}{\sqrt{\alpha}}$ so $Z \sim N(0,1)$ where $\alpha \to \infty$

To approve that whereas $X \sim G(\alpha)$ and that $Z=\frac{x-\alpha}{\sqrt{\alpha}}$ standard normal in this distribution were $\mu_x = o$ & $\sigma^2{}_x = 1$.

Let it be

$$M_z(x)=E(e^{tz})=e^{-t\sqrt{\alpha}}.M_x(\frac{1}{\sqrt{\alpha}})$$

$$=e^{-t\sqrt{\alpha}}.(1-\frac{1}{\sqrt{\alpha}})^{-\alpha} \qquad ...(16)$$

And that $K_z(t)=\ln M_x(t)=-t\sqrt{\alpha}-\alpha\ln(1-\frac{1}{\sqrt{\alpha}})$

But according to Tyler's series then: $\ln(1+K)=K-\frac{K^2}{2}+\frac{K^3}{3}-...,|K|\prec 1$

And if we assume that when $(K=-\frac{1}{\sqrt{\alpha}})$ and $|K|\prec 1$ close to infinity then:

$$K_z(t)=-t\sqrt{\alpha}+\alpha(\frac{1}{\sqrt{\alpha}}+\frac{t^2}{2\alpha}+\frac{t^3}{3\alpha\sqrt{\alpha}}+...)$$

$$=\frac{t^2}{2}+\frac{t^3}{3\sqrt{\alpha}}+0(\frac{1}{\alpha}) \qquad ...(17)$$

Where ($0(\frac{1}{\alpha})$) represents last terms that include (α) in their denominators. So that $\lim_{\alpha\to\infty} K_z(t)=\frac{t^2}{2}$ that is meaning $\lim_{\alpha\to\infty} M_z(t)=\frac{1}{2}t^2$, last equation represent the moment generating function of standard normal distribution .so conclude that : $Z \sim N(0,1)$ Where $\lim_{\alpha\to\infty} G(\alpha)\to N(\alpha,\alpha)$ SO $\lim_{\alpha\to\infty} G(\alpha,\beta)\to N(\alpha\beta,\alpha\beta^2)$

6 – Methods of Comparisons

A comparison is carried out between the mentioned distributions based on the classical kolmogrov-smirnov distance to minimize distance estimation test statistic and Grammer von misses distance, to emphasize that which distribution fitted to the data better than the other models.

6-1- Kolmogorov-Smirnov Test [18]

This test is used to decide if a sample comes from a hypothesized continuous distribution. It is based on the empirical cumulative distribution function (ECDF). Assume that we have a random sample $x_1, \ldots, x_n$ from some distribution with CDF *F(x)*. The empirical CDF is denoted by:

$$Fn(x) = \tfrac{1}{n}\left[number\ of\ observations \le x\right] \quad \ldots.(18)$$

The Kolmogorov-Smirnov statistic (D) is based on the largest vertical difference between the theoretical and the empirical cumulative distribution function:

$$D = \max_{1\le i\le n}\left(F(x_i) - \tfrac{i-1}{n}, \tfrac{i}{n} - F(x_i)\right) \quad \ldots(19)$$

6-2- Gramer-mises criterion [14]

Is a criterion used for the goodness of fit of a (c.d.f) cumulative distribution function(F*) compared to a given(F_n) empirical distribution function, or for comparing two empirical distributions. It is also used as a part of other algorithms, such as minimum distance, It is defined as:

$$\omega^2 = \int_{-\infty}^{\infty}\left[Fn(x) - F^*(X)\right]^2 .dF^*(X) \quad \ldots(20)$$

In one-sample applications **F** is the theoretical distribution and **F*** is the empirical distribution function. Alternatively the two distributions can both be empirically estimated ones; this is called the two-sample case. For one sample the formula is defined as:

$$T = n\omega^2 = \tfrac{1}{12n} + \sum_{i=1}^{n}\left[\tfrac{2i-1}{2n} - F(x_i)\right]^2 \quad \ldots(21)$$

A modified version of the Gramer–von Mises test is the Watson test which uses the statistic U2, where

$$U^2 = T - (\bar{F} - \tfrac{1}{2})^2$$

$$where \quad \bar{F} = \tfrac{1}{n}\sum_{i=1}^{n} F(x_i)$$

And for two sample

$$T = \frac{NM}{N+M}\omega^2 = \frac{U}{NM(N+M)} - \frac{4MN-1}{6(M+N)} \quad (22)$$

$$where: U = N\sum_{i=1}^{N}(r_i - i)^2 + M\sum_{j=1}^{M}(Sj - j)^2$$

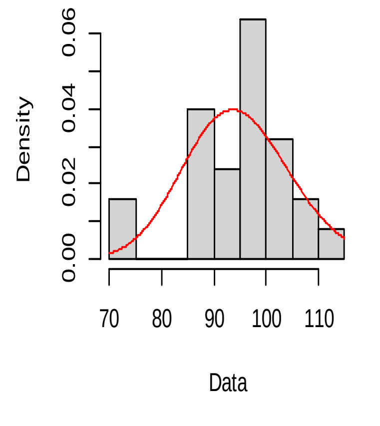
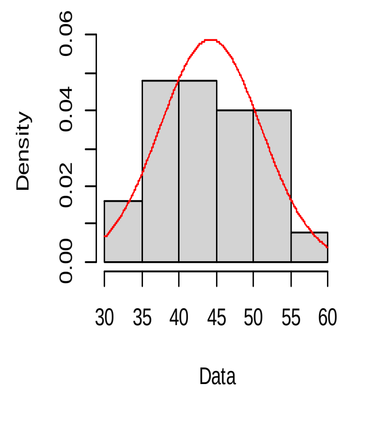
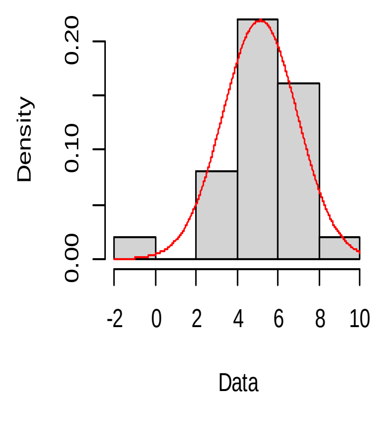
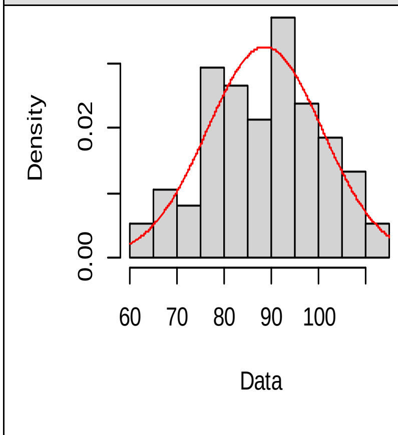
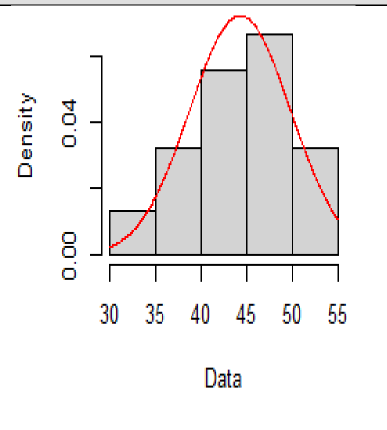
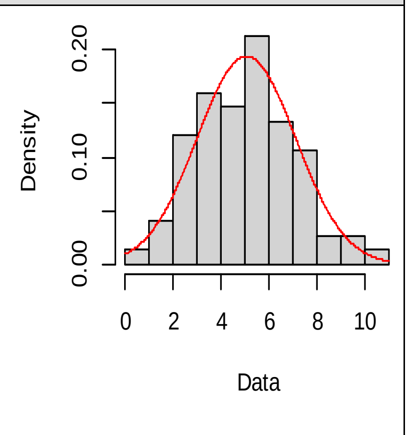
N and M are sample size, with combine (S) sample

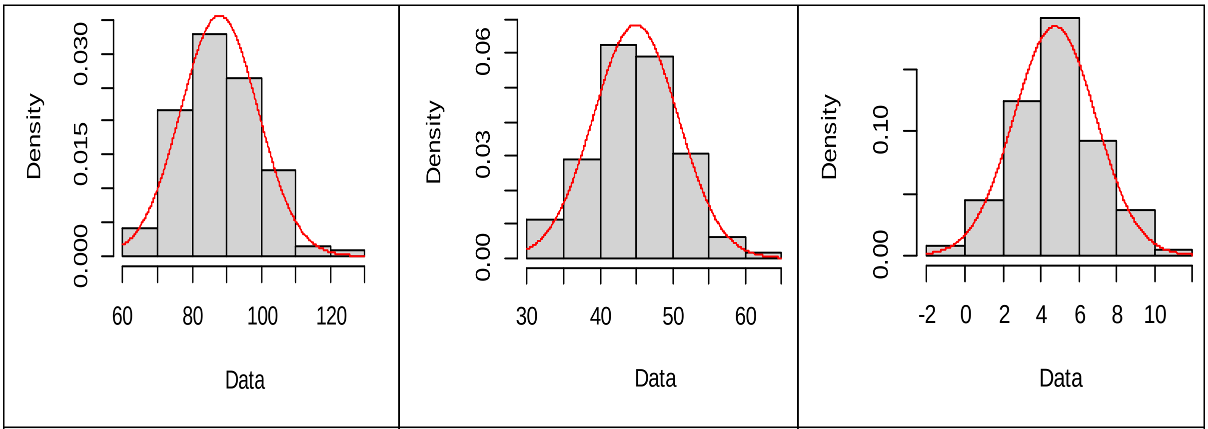
7- Simulation and data analysis:

To study maximizing the likelihood of distribution, simulated data were generated that follow three probability distributions: the Normal distribution, Exponential distribution , in addition to Gamma distribution with different sample sizes (25, 75 and 125) and different parameter sizes (three sizes for each distribution parameter(s)), The parameters of these distributions were estimated on real data.

The R programming language used and a set of packages (fitdistrplus) and functions (rnorm, rgamma, rexp, "mle" , fitdist, "mge" , gof="KS", gof="CvM" insides plot) that helped us get these results.

A comparison was made between the appropriateness of the distributions for the data generated by relying on two statistical tests (Grammer von mises distance, classical kolmogrov smirnov distance) and the results were as follows.

		
M=90 & SD=12 , N=25	M=45 & SD=6, N=25	M=5 & SD=2, N=25
		
M=90 & SD=12 , N=75	M=45 & SD=6, N=75	M=5 & SD=2 , N=75



M=90 & SD=12 , N=125	M=45 & SD=6, N=125	M=5 & SD=2 , N=125

Fig.(4) Shows the Empirical and theoretical density from generating data distributed Normal distribution

Parameter(s) value	Rank	N=25	K.S.	N=75	K.S.	N=125	K.S.
M=90 & SD=12	1	Normal	0.15098	Normal	0.04914	Gamma	0.06438
	2	Gamma	0.16494	Gamma	0.06699	Normal	0.08055
	3	Exponential	0.52547	Exponential	0.51266	Exponential	0.52627
M=45 & SD=6	1	Normal	0.08648	Normal	0.08088	Normal	0.05435
	2	Gamma	0.09737	Gamma	0.09749	Gamma	0.0706
	3	Exponential	0.51127	Exponential	0.51113	Exponential	0.50155
M=5 & SD=2	1	Normal	0.10052	Normal	0.04626	Normal	0.05889
	2	NFIT	NA	Gamma	0.0781	NA	NA
	3	NFIT	NA	Exponential	0.2852	NA	NA

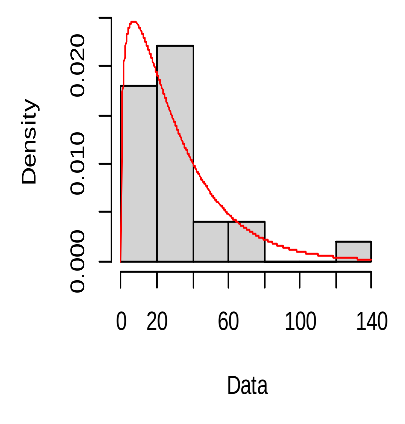
Table (1): Shows the rank of distributions fitting when the generating data follows Normal Dist by Kolmogorov Smirnov test.

Parameter(s) value	Rank	N=25	Chi	N=75	Chi	N=125	Chi
M=90 & SD=12	1	Gamma	0.0717	Normal	0.72547	Gamma	7.7928
	2	Normal	0.0758	Gamma	1.3736	Normal	8.3469
	3	Exponential	106.48	Exponential	294.37	Exponential	570.68
M=45 & SD=6	1	Normal	0.33531	Normal	2.4571	Normal	2.4006
	2	Gamma	0.34983	Gamma	4.89	Gamma	5.9406

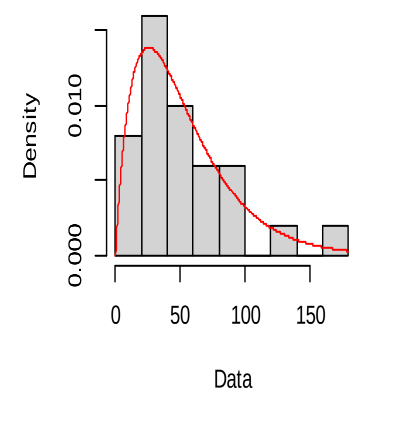
	3	Exponential	40.117	Exponential	334.52	Exponential	530.9
M=5 & SD=2	1	Normal	1.1865	Normal	1.4353	Normal	4.9187
	2	NFIT	NA	Gamma	3.1778	NA	NA
	3	NFIT	NA	Exponential	57.472	NA	NA

Table(2): Shows the rank of distributions fitting when the generating data follows Normal Dist by Grammer von test.

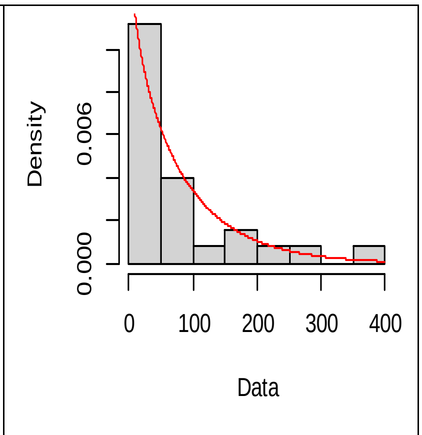
For the above two tables, which represent the fit of the data generated according to the normal distribution, it is not necessary that the normal distribution always be the best fit distribution, as sometimes the gamma distribution is the best and it has not been shown that the exponential distribution is the best for representing the data while both distributions (Gamma & Exponential) fail to fitting the data when the parameters size is small and for both tests (Kolmogorov Smirnov and Grammer von tests).



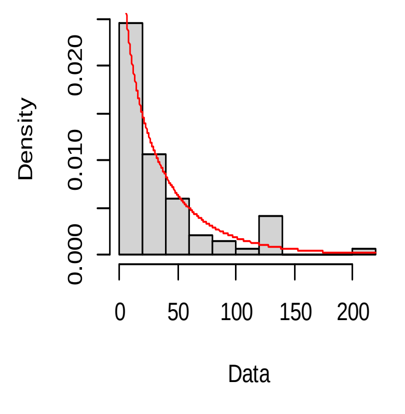
Lambda=0.310 , N=25



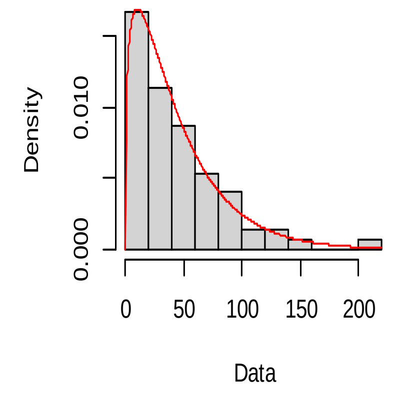
Lambda=0.213, N=25



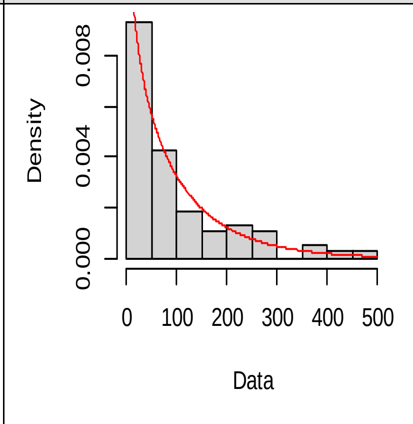
Lambda=0.11, N=25



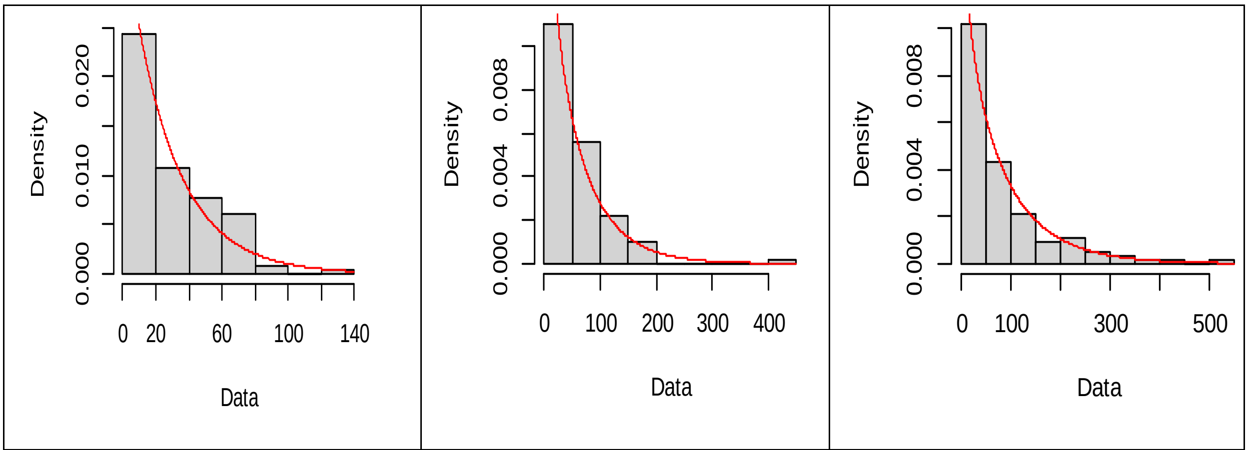
Lambda=0.310 , N=75



Lambda=0.213, N=75



Lambda=0.11, N=75



Lambda=0.310 , N=125	Lambda=0.213, N=125	Lambda=0.11, N=125

Fig.(5) Shows the Empirical and theoretical density from generating data distributed Exponential distribution

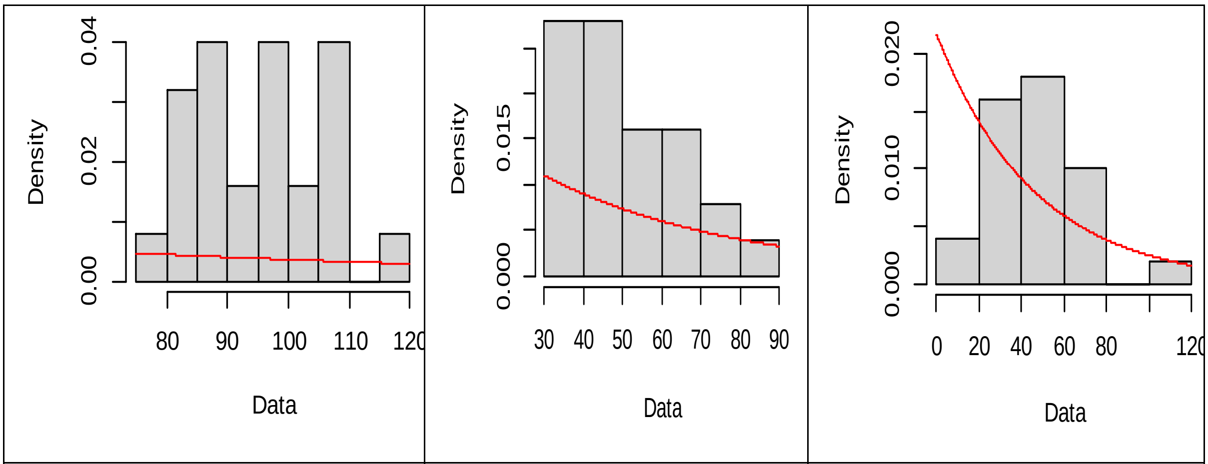
Parameter(s) value	Rank	N=25	K.S.	N=75	K.S.	N=125	K.S.
Lambda=0.310	1	Gamma	0.12166	Gamma	0.0696	Exponential	0.07382
	2	Exponential	0.16941	Exponential	0.12425	Gamma	0.09213
	3	Normal	0.2197	Normal	0.19733	Normal	0.13394
Lambda=0.213	1	Gamma	0.07532	Gamma	0.06926	Gamma	0.07787
	2	Normal	0.1484	Exponential	0.09197	Exponential	0.09092
	3	Exponential	0.18559	Normal	0.13702	Normal	0.1710
Lambda=0.11	1	Gamma	0.1503	Gamma	0.07361	Gamma	0.06376
	2	Exponential	0.21628	Exponential	0.11403	Exponential	0.0862
	3	Normal	0.24866	Normal	0.21423	Normal	0.21716

Table (3): Shows the rank of distributions fitting when the generating data follows Exponential Dist by Kolmogorov Smirnov test.

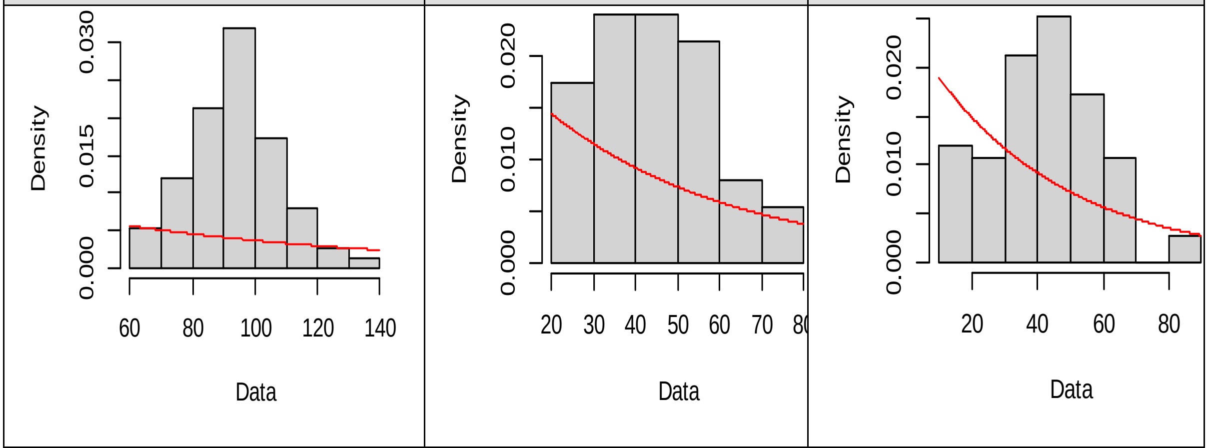
Parameter(s) value	Rank	N=25	Chi	N=75	Chi	N=125	Chi
Lambda=0.310	1	Exponential	1.6048	Gamma	2.1893	Exponential	4.4631
	2	Gamma	2.814	Exponential	3.0281	Gamma	12.197
	3	Normal	3.8467	Normal	14.483	Normal	15.51
Lambda=0.213	1	Gamma	0.37344	Gamma	2.5623	Gamma	8.9714
	2	Normal	0.7918	Exponential	4.266	Exponential	9.0204
	3	Exponential	0.88825	Normal	4.1684	Normal	10.399
Lambda=0.11	1	Gamma	3.0814	Normal	6.7321	Gamma	1.724
	2	Exponential	3.8795	Gamma	12.643	Exponential	2.4609
	3	Normal	4.2729	Exponential	16.808	Normal	22.8

Table (4): Shows the rank of distributions fitting when the generating data follows Exponential Dist by Grammer von test.

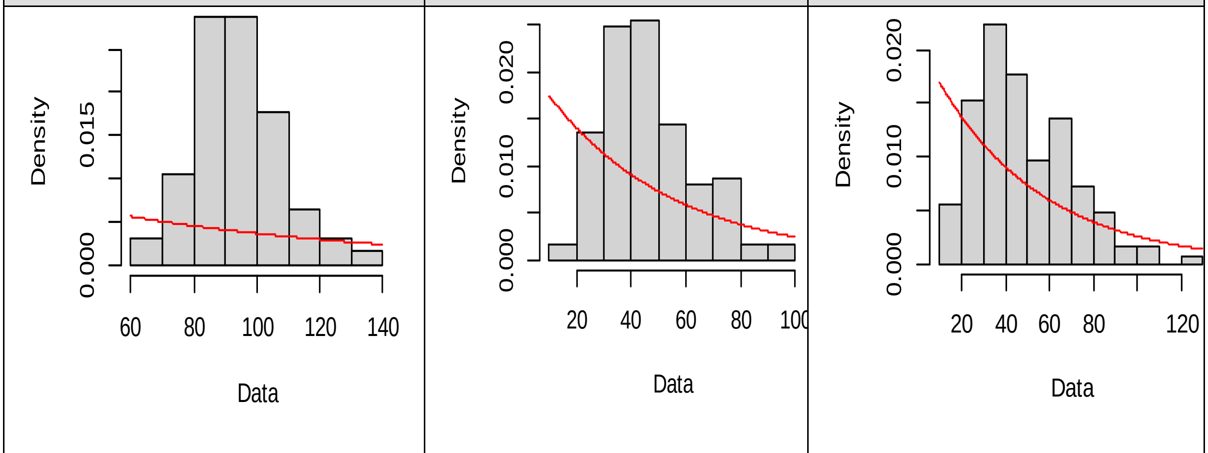
For the above two tables, which represent the fit of the data generated according to the exponential distribution, the exponential distribution does not have to be the best-fit distribution, as often the gamma distribution is best and for both tests (Kolmogorov Smirnov and Grammer von tests) and in rare cases a normal distribution when the parameter size is small.



Shape=50 & Rate=0.54 , N=25	Shape=10.15 & Rate=0.22, N=25	Shape=5.2 & Rate=0.12, N=25



Shape=50 & Rate=0.54 , N=75	Shape=10.15 & Rate=0.22, N=75	Shape=5.2 & Rate=0.12, N=75



Shape=50 & Rate=0.54 , N=125	Shape=10.15 & Rate=0.22, N=125	Shape=5.2 & Rate=0.12, N=125

Fig.(6) Shows the Empirical and theoretical density from generating data distributed Gamma distribution

Parameters value	Rank	N=25	K.S.	N=75	K.S.	N=125	K.S.
Shape=50 & Rate=0.54	1	Gamma	0.1185	Gamma	0.05755	Gamma	0.05485
	2	Normal	0.12814	Normal	0.073	Normal	0.06744
	3	Exponential	0.566	Exponential	0.49865	Exponential	0.50477
Shape=10.15	1	Gamma	0.10124	Gamma	0.07999	Gamma	0.05832

& Rate=0.22	2	Normal	0.14067	Normal	0.09215	Normal	0.09548
	3	Exponential	0.45288	Exponential	0.4009	Exponential	0.357
Shape=5.2 & Rate=0.12	1	Gamma	0.07723	Normal	0.05826	Gamma	0.06355
	2	Normal	0.13305	Gamma	0.08744	Normal	0.12346
	3	Exponential	0.34868	Exponential	0.30073	Exponential	0.29649

Table (5): Shows the rank of distributions fitting when the generating data follows Gamma Dist by Kolmogorov Smirnov test.

Parameters value	Rank	N=25	Chi	N=75	Chi	N=125	Chi
Shape=50 & Rate=0.54	1	Normal	0.93601	Normal	4.326	Gamma	6.6939
	2	Gamma	0.93827	Gamma	6.5927	Normal	8.4665
	3	Exponential	98.43	Exponential	274.01	Exponential	596.44
Shape=10.15 & Rate=0.22	1	Gamma	0.36269	Gamma	2.0272	Gamma	7.2576
	2	Normal	1.1542	Normal	4.1371	Normal	9.4773
	3	Exponential	21.996	Exponential	94.858	Exponential	134.15
Shape=5.2 & Rate=0.12	1	Gamma	0.01487	Normal	5.334	Gamma	7.1979
	2	Normal	3.5316	Gamma	6.1305	Normal	14.976
	3	Exponential	11.078	Exponential	73.569	Exponential	91.705

Table (6): Shows the rank of distributions fitting when the generating data follows Gamma Dist by Chi-square test.

For the above two tables, which represent the fit of the data generated according to the gamma distribution, it is not necessary for the gamma distribution to be the best fit distribution, sometimes the normal distribution is the best and no case appears that the exponential distribution is the best among the three distributions and for both tests (Kolmogorov Smirnov and Grammer von).

8- Conclusions

Through the analysis of the generated data, we noticed that there is a relative relationship between the three distributions (normal, gamma and exponential) and according to the different sample sizes (25, 75 and 125) and the values of the three different values of parameters for each distribution, maximizing the probability values of the distribution depends on three factors the behavior of the data, the size of the sample and the number of parameters in the specified distribution, we have concluded that it is not necessary that the data generated for a particular distribution gets the best fitting of the data when comparing it with other distributions, e.g. when generating data that follows the exponential distribution with a sample size (25,125) and for the three different parameter sizes, the best distribution for fitting the data is the gamma distribution not exponential distribution, while the exponential distribution is the best among the three distributions when the sample size is (125) and the value of parameter is (0.036) and the two statistical test (Grammer von mises distance, classical kolmogrov smirnov distance) not same always for distribution fitting.

In the case of generating data follow the normal distribution, and in the two cases of sample size (25, 75) and with three different sizes of parameters that were chosen, the best distribution suitable for the generating data is the normal distribution itself, while in the case of large sample size (125) and the values of the parameters are (mean=90, SD= 12) the best distribution fitting the data Is gamma distribution not normal.

When generating data that follows a gamma distribution, the most appropriate distribution of data with sample sizes (25,125) and with different parameter values is the gamma distribution, while in the case of a medium sample size (75) and small parameter size (shape=5.2, rate=0.12) the normal distribution is best for fitting the data.

It is not necessary for the data to be generated following a specific distribution (e.g. Normal dist.), it can follow the others (gamma or an exponential) with a lesser probability and fail to be gamma or exponential.

9- Recommendations

We recommend studying the similarities and differences between other distributions (bounded, unbounded, non negative), e.g. (Pert, Gumbel, levy) respectively in terms of their degree of agreement with the data.

We recommend to use other tests as a criteria for comparison among distributions to minimize the distance and which one better than others like (Akaki Information Criteria or Anderson Darling) tests.

We recommend using a single distribution with multiple parameters, for example, a gamma distribution with two, three, four parameters...etc. And studying the degree of complexity of the distribution and its impact on maximizing the likelihood of the distribution.

Reference

1) Christian walck ,2007,” statistical distribution for experimentalists “ university of Stockholm
2) David M. Blei ,2014 “the exponentials family “Columbia university .
3) Douglas G. Altman, J.Martin Bland,1995,” the normal distribution “Medical statistics Laboratory, London .
4) Eugene F. Fama & Richard Rolla (2014),’’ Some Properties of Symmetric Stable Distributions ‘’Journal of the American Statistical Association.
5) Eric Xing ,2009, probabilistic graphical models “School of Computer Science Carnegie Mellon
6) Eugenef. F. Fama,Richard Roll,(2014),” some properties of symmetric stable distributions“ journal of the American statistical association .
7) Gauss M. Cordeiro,, Edwin M.M. Ortega and Artur J. Lemonte,(2013),’’ The exponential–Weibull lifetime distribution’’ Journal of Statistical Computation and Simulation.
8) Jimmy Reyes, Emilio Gómez-Déniz,Héctor W.Gómez and Enrique Calderín-Ojeda,(2021) ‘’A Bimodal Extension of the Exponential Distribution with Applications in Risk Theory’’
9) Rameshwar D. Gupta and Debasis Kundu (2006), ’’Closeness of Gamma and Generalized Exponential Distribution. Communication in statistics Theory and Methods Vol. 32, No. 4, pp. 705–721.
10) Robert V.Hogg, Allent.T. Craig and . McKean W.Joseph (2019) “introduction of mathematical statistics “8th edition .Pearson Education Inc.USA.

11) Sivakumar Rathinam & Jordan,(2004),” the exponentials family “stat241 A statistical learning theory .lecture 11 .
12) Teffen Lauitzen ,Michaelmas team, ,2004,” maximum likelihood in exponentials family”lecture 6 ,university of oxford .
13) Thomos A.Severinl ,(2005),”Elements of distribution theory “Cambridge university press .Newwork .USA .
14)Laio, F. (2004),” Cramer–von Mises and Anderson-Darling goodness of fit tests for extreme value distributions with unknown parameters”, Water Resour. Res., 40 .

15) Mathias Niepert , Pasquale Minervini and Luca Franceschi.(2021) “Implicit MLE: Back propagating Through Discrete Exponential Family Distributions” .
16) Ogbonna C.Justin ,(2004),”statistical distribution theory “1th edition .Nkems Printing press.
17)Pishro Hossein-Nik ,(2021),”introduction to probability statistics and ransom process “
18)Weiwen Miao and Marjorie Hahn.(1997),”Existence of MLE for multi-Dimensonal exponential families “Scandinavian journal of statistics ,vol,24,No.3 pp(371-386).
19)W. J. Conover(1999),”Practical Nonparametric Statistical ” ,3rd edition, pp.428-433 (6.1), John Wiley & Sons, Inc. New York.