Biostatistics and Biometrics
Open Access Journal
ISSN: 2573-2633

Juniper Publishers
key to the Researchers

Short Communication
Volume 8 Issue 4 - November 2018
DOI: 10.19080/BBOAJ.2018.08.555741

Biostat Biometrics Open Acc J

On a statistical approximation model of probabilitydensity function of non-negative random variables

Tibor KPogany*

Faculty od Maritime Studies, University of Rijeka, Croatia

Institute of Applied Mathematics, Obuda University, Hungary

Submission: July 31, 2018; **Published:** November 13, 2018

***Corresponding author:** Tibor K Pogany, Faculty od Maritime Studies, University of Rijeka, Studentska 2, 51000 Rijeka, Croatia; Institute of Applied Mathematics, Obuda University, Becsiut 96/b, 1034 Budapest, Hungary

Abstract

The probability density function of a non–negative random variable is approximated by virtue of a result by Totik [1], when the support domain is truncated to a finite interval. This result is transferred to the case of three–parameter Weibull distribution.

2010 Mathematics Subject Classification: Primary: 60E05, 62E17; Secondary: 41A10, 42C05, 44A60

Keywords: Christoffel function; moments; statistical approximation of probability density; three-parameter Weibull distribution

Introduction and preparation

In continuation we expose a method to estimate any probability density function (pdf) $f(x)$ having finite, compact support set $\mathrm{supp}(f)=[a,b]$, $-\infty<a<b<\infty$, by means of a random sample Ξ such that consists from n i.i.d. replicae of the initial $rv\xi \sim f(x)$. Moreover, when the $\mathrm{supp}(f)$ is infinite, we approximate f with its truncated variant so, that the corresponding approximation error does not skip the prescribed error level $\varepsilon>0$.

Denote (Ω, F, P) a given standard probability space, and let $L^p(\Omega,F,P)$ be the class of all complex–valued random variables $rv\xi\sim f(x)$ having finite moment of pth order, p positive integer, that is

$$m_p = m_p(\xi) := \int_{\mathbb{R}} x^p f(x)\,dx.$$

Let us consider the sequence of orthogonal polynomials $\{p_j(x)\}_{j\in\mathbb{N}_0}$ with respect to a pdf $f(x)$, $x\in I$, which can be expressed in terms of moments of $f(x)$ [2][7], [5] [3] as follows:

$$p_j(x)=p_j(f;x):=\frac{D_j(x)}{\sqrt{M_{2j}M_{2j-2}}}(j\in\mathbb{N}_0), \quad (1.1)$$

Where

$$D_j(x)=\begin{vmatrix} 1 & m_1 & \cdots & m_j \\ \vdots & \vdots & \ddots & \vdots \\ m_{j-1} & m_j & \cdots & m_{2j-1} \\ 1 & x & \cdots & x^j \end{vmatrix} \text{ and } M_{2j}=\begin{vmatrix} 1 & \cdots & m_j \\ \vdots & \ddots & \vdots \\ m_j & \cdots & m_{2j} \end{vmatrix} \quad (1.2)$$

with $M_{-2}=M_0=1$.

Our second tool is the Christoffel function [1, Eq. (4)]

$$\lambda_k(f,x)=\frac{1}{\sum_{j=0}^{k}|p_j(x)|^2}.$$

In his excellent article, solving the so–called Szego˝–problem on the real line, Totik proved [1] that for almost all $x\in I$ it holds

$$\lim_{k\to\infty} k\lambda_k(v,x)=\pi v'(x)\sqrt{(b-x)(x-a)},$$

being $\lambda_k(v,x)$ the Christoffel function for orthogonal polynomials associated to a positive measure $v>0$ giving the Radon–Nikodym derivative v', when $\mathrm{In}v'$ is integrable on I. Now, pointing out that for all B Borel

$$\int_B f(x)\,dx \qquad (B\in B_{\mathbb{R}})$$

is a probability measure, therefore it can be easily indentified with v, when f is bounded and positive. Thus, we have re–formulate Totik's asymptotic result into our setting (also see [1, p. 10, Theorem 4.1]).

Theorem A. Let $f(x)$ be a bounded pdf of $rv\ \xi$, $\mathrm{supp}(f)=[a,b]$ and $\inf\in L^1[a,b]$. Then we have

$$\lim_{k\to\infty} k\lambda_k(f,x)=\pi(f,x)\sqrt{(b-x)(x-a)}$$

for almost all $x\in[a,b]$.

The first approximation of the underlying pdf we deduce from (1.3) for almost all $x\in[a,b]$:

$$f(x)\approx f_k(x)=\frac{k\lambda_k(f,x)}{\pi\sqrt{(b-x)(x-a)}} \quad (1.4)$$

where the approximant $f_k(x)$ is not a pdf, so it has to be re–normalized.

Main results

Another problem arises with a pdf $f(x)$ having infinite support set $\mathrm{supp}(f)=[a,\infty), a\ge 0,$ which turns out to be only incidentally connected with the so–called Stieltjes moment problem, consult for instance in this respect [3, 4, 6][4,5,6]. To avoid the infinite support interval we truncate it to the finite $[a,b]$, $a<b$ so, that the truncation error should be less then $\varepsilon>0$.

Proposition 1. Let $\xi\sim f(x)$ be a rv such that $\mathrm{supp}(f)=[a,\infty)$, $a\ge 0$ and assume $\xi\in L^p(\Omega,F,P)$. Let $b=b_p(\varepsilon)$, $\varepsilon>0$ be the unique zero of the equation

$$\int_b^\infty x^p f(x)\,dx=\varepsilon b^p. \quad (2.1)$$

Then we have

$$P\{a\le\xi<b\}\ge 1-\varepsilon. \quad (2.2)$$

Proof: Because $\int_b^\infty x^p f(x)dx$ is a decreasing function of b, and εb^p monotone increases on the same domain $b>a$, , (2.1) has an unique root in b. Thus

$$P\{\xi\ge b\}=\int_b^\infty f(x)dx\le\frac{1}{b^p}\int_b^\infty x^p f(x)dx=\varepsilon$$

by the assumption of the proposition. Accordingly

$$1=P\{a\le\xi<\infty\}=P\{a\le\xi<b\}+P\{\xi\ge b\}\le P\{a\le\xi<b\}+\varepsilon,$$

which is equivalent to the claim (2.2).

Having in mind the approximation of an initial pdf having infinite support by means of (2.2) and in the same time renormalizing the resulting function approximants (1.4) so, that $\int_b^\infty f_k(x)dx=1-\varepsilon,$ we get the sequence of approximants

$$\tilde f_k(x)=\frac{\dfrac{(1-\varepsilon)\lambda_k(f,x)}{\sqrt{(b-x)(x-a)}}}{\displaystyle\int_a^b\frac{\lambda_k(f,x)\,dx}{\sqrt{(b-x)(x-a)}}}. \quad (2.3)$$

It remains to estimate b by means of the random sample Ξ in approximation formula (2.3).

In this goal we have to estimate all moments m_j, $j=\overline{1,2k}$ involved in $p_k(f,x)$ by means of certain statistics/Borel functions of Ξ. Let us recall that

$$A_k=\frac{1}{n}\sum_{j=1}^n X_j^k \qquad (k\in\mathbb{N})$$

is the so–called kth sample moment, the analogue of theoretical moment m_k. It is a standard result of mathematical statistics, that A_k is an unbiased estimator for m_k, being $EA_k=m_k$;

moreover, by the Khintchine's Law of Large Numbers for all k such that $m_k<\infty$, we have that $A_k\xrightarrow{P} m_k$, when n runs to infinity, that is, is a consistent estimator of A_k m_k.

Now, we transform the approximation procedure (2.3) to an infinite domain pdf, prescribing the admissible approximation error level $\varepsilon>0$, taking finite support interval $[a,\hat b]$, where

$$\hat b\ge\left(\frac{\sum_{j=1}^n X_j^p}{n\varepsilon}\right)^{1/p} \qquad (p\in\mathbb{N}).$$

By some practical reasons, we make use of the first sample mean A_1 or the maximum–order statistic $X_{(n)}-\max_{1\le j\le n}X_i$ in evaluating $\hat b$.

Proposition 2: Replacing m_j with A_j Aj in (1.1) and (1.2), by means of (2.3) we get the finite pdf approximating sequence

$$\hat f_k(x)=\frac{\dfrac{(1-\varepsilon)\hat\lambda_k(\Xi,x)}{\sqrt{(b-x)(x-a)}}}{\displaystyle\int_a^b\frac{\hat\lambda_k(\Xi,x)\,dx}{\sqrt{(b-x)(x-a)}}} \qquad (0\le k\le n-1),$$

Where

$$\hat\lambda_k(\Xi,x)=\frac{1}{\sum_{j=0}^k\left|\hat p_j(\Xi,x)\right|^2},\quad \hat p_j(\Xi,x)=\frac{\hat D_j(x)}{\sqrt{\hat M_{2j}\hat M_{2j-2}}},$$

$$\hat D_j(x)=\begin{vmatrix}1 & A_1 & \cdots & A_j\\ \vdots & \vdots & \ddots & \vdots\\ A_{j-1} & A_j & \cdots & A_{2j-1}\\ 1 & x & \cdots & x^j\end{vmatrix},\quad \hat M_{2j}=\begin{vmatrix}1 & \cdots & A_j\\ \vdots & \ddots & \vdots\\ A_j & \cdots & A_{2j}\end{vmatrix}.$$

Proof: Let us define the matrices

$$T_k:=\begin{bmatrix}1 & 1 & \cdots & 1\\ x_1 & x_2 & \cdots & x_n\\ x_1^2 & x_2^2 & \cdots & x_n^2\\ \vdots & \vdots & & \vdots\\ x_1^k & x_2^k & \cdots & x_n^k\end{bmatrix},\quad (k\in\mathbb{N}).$$

Then we have

$$T_kT_k^\tau=\begin{bmatrix}1 & 1 & \cdots & 1\\ x_1 & x_2 & \cdots & x_n\\ x_1^2 & x_2^2 & \cdots & x_n^2\\ \vdots & \vdots & & \vdots\\ x_1^k & x_2^k & \cdots & x_n^k\end{bmatrix}\cdot\begin{bmatrix}1 & x_1 & x_1^2 & \cdots & x_1^k\\ 1 & x_2 & x_2^2 & \cdots & x_2^k\\ 1 & x_3 & x_3^2 & \cdots & x_3^k\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & x_n & x_n^2 & \cdots & x_n^k\end{bmatrix}$$

$$=\begin{bmatrix}n & S_1 & S_2 & \cdots & S_k\\ S_1 & S_2 & S_3 & \cdots & S_{k+1}\\ S_2 & S_3 & S_4 & \cdots & S_{k+2}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ S_k & S_{k+1} & S_{k+2} & \cdots & S_{2k}\end{bmatrix},$$

Where

$$S_k:=nA_k=\sum_{j=1}^n X_j^k \qquad (X_1<\cdots<X_n).$$

How to cite this article: Tibor KPogany. On a statistical approximation model of probabilitydensity function of non-negative random variables. Biostat Biometrics Open Acc J. 2018; 8(4): 555741. DOI: 10.19080/BBOAJ.2018.08.555741

First, consider the case $k \geq n$. The columns of T_k are linearly dependent so rank $(T_k) = n$. The matrix $T_k T_k^\tau$ has $k+1 > n$ columns, and every column of $T_k T_k^\tau$ is a linear combination of columns of T_k. therefore $T_k T_k^\tau$ possesses at most n linearly independent columns. Thus, being the columns of $T_k T_k^\tau$ linear dependent, we conclude that

$$\det\left(T_k T_k^\tau\right) = n^{k+1} \hat{M}_{2k} = 0.$$

Now, it remains the case $k < n$. The matrix T_k has linear independent row–vectors such that we denote

$$a^i := \left(x_1^{i-1}, x_2^{i-1}, \ldots, x_n^{i-1}\right) \quad \left(i = \overline{1, k+1}\right).$$

It turns out that $T_k T_k^\tau = \Gamma\left(a^1, \ldots, a^{k+1}\right)$, where Γ stands for the Gramm determinant of the vector $\left(a^1, \ldots, a^{k+1}\right)$. Therefore,

$$\hat{M}_{2k} = \det\left(T_k T_k^\tau\right) > 0.$$

Obvious steps finish the proof.

Moreover, we quote an alternative approach to this problem.

$$P\{a \leq \xi < b\} \geq 1 - \varepsilon.$$

So, we have $P\left\{a \leq X_{(n)} < b\right\} = F^n(b) - F^n(a+) \geq 1 - \varepsilon$, where F is the cumulative distribution function of the rv ξ with the related pdf $f = F'$. Then

$$F(b) \geq \left(F^n(a+) + 1 - \varepsilon\right)^{1/n}.$$

The Lagrange theorem applied to F results in

$$F(b) = F\left(\bar{X}_n + \left(b - \bar{X}_n\right)\right) = F\left(\bar{X}_n\right) + f\left(X_{(n)}\right)\left(b - \bar{X}_n\right),$$

Where $\bar{X}_n = A_1 < X_{(n)} = \max_{1 \leq j \leq n} X_j < b$. This implies

$$b \geq \bar{X}_n + \frac{\left(F^n(a+) + 1 - \varepsilon\right)^{1/n} - F\left(\bar{X}_n\right)}{f\left(X_{(n)}\right)}.$$

Assuming that a is the continuity point of F, it is $F(a+) = 0$ and we have to solve the problem with respect to b. By the way $f\left(X_{(n)}\right) \approx \hat{f}_k\left(X_{(n)}\right)$, while

$$F\left(\bar{X}_n\right) \approx \int_a^{\bar{X}_n} \hat{f}_k(x) dx.$$

The value $\hat{f}_k(x)$ we approximate with the associated Christoffel function etc. However, if this procedure is not efficient, take $X_{(n)}$ instead of $\bar{X}_{(n)}$. Namely, in that case the lower bound for b will be moved to the right as $X_{(n)} > \bar{X}_{(n)}$.

Application to $W(\alpha, \beta, \eta)$ distribution

The random variable ξ possesses three–parameter Weibull distribution when the corresponding pdf reads

$$f_w(x) := \frac{\beta}{\eta}\left(\frac{x-\alpha}{\eta}\right)^{\beta-1} \exp\left\{-\left(\frac{x-\alpha}{\eta}\right)^{\beta}\right\} \chi(\alpha, \infty)(x),$$

with the parameter $\theta := (\alpha, \beta, \eta)$; $\alpha \geq 0$ while $\beta, \eta > 0$. This correspondence we write $\xi \sim W(\alpha, \beta, \eta)$. Here, $\chi_A(x)$ denotes the indicator of the event $\{x \in A\}$.

The question of estimating the parameter vector $\theta := (\alpha, \beta, \eta)$ by means of random sample $\Xi := (X_1, \ldots, X_n)$, being $X_i, j = \overline{1, n}$ the i.i.d. replica of $\xi \sim W(\theta)$ arises frequently. The most common way of given some estimator $\hat{\theta}$ is the maximum likelihood method. However, the ML estimator for the three–parameter Weibull distribution does not necessarily exists, and it is not necessarily unique, see [7] and the references therein.

Corollary 2.1. Let $\xi \sim W(\alpha, \beta, \eta)$, with all three parameters positive [8]. Consider the sample Ξ of size $|\Xi| = n$. Then

$$\hat{f}_{w,k}(x) = \frac{\dfrac{(1-\varepsilon)\hat{\lambda}_k(\Xi, x)}{\sqrt{\left(\hat{b} - x\right)(x-a)}}}{\displaystyle\int_a^b \frac{\hat{\lambda}_k(\Xi, x)(x-a)}{\sqrt{\left(\hat{b} - x\right)(x-a)}}} \qquad (0 \leq m \leq n-1),$$

Where

$$\hat{b} \geq \frac{1}{\varepsilon n} \sum_{j=1}^n X_j =: \hat{b}_1.$$

Proof. It is enough to show that the Szeg˝o–type local integrability condition is satisfied, that is, $\ln f_W(x) \in L^1\left(\alpha, \hat{b}\right]$. Indeed, as $f_W(x) > 0$ for all $x > \alpha > 0$, equivalently $x \geq \alpha + \delta$, $\delta > 0$, we have

$$\int_{\alpha+\delta}^{\hat{b}} \ln f_W(x) dx = \left(\hat{b} - \alpha - \delta\right) \ln \frac{\beta}{\eta\beta} + (\beta - 1) \ln \frac{\left(\hat{b} - \alpha\right)^{\hat{b}-\alpha}}{\delta^\delta e^{\hat{b}-\alpha-\delta}} - \frac{\left(\hat{b} - \alpha\right)^{\beta+1} - \delta^{\beta+1}}{\eta^\beta(\beta+1)}.$$

Obviously, the integral is convergent.

Remark. Instead of the lower bound $\hat{b}_1$, the estimate

$$\hat{b} \geq \frac{1}{\varepsilon} X_{(n)} := \hat{b}_\infty$$

can be used.

References

1. Totik V (2000) Asymptotics for Christoffel functions for general measures on the real line. J Anal Math 81(1): 283-3003.
2. Szego G (1939) Orthogonal Polynomials. American Mathematical Society Colloquium Publications 23, 2nd Edition, AMS: Providence, Rhode Island.
3. Shohat JA, Tamarkin D (1943) The Problem of Moments. American Mathematical Society Mathematical Surveys, Vol. II, AMS: New York.
4. Kopanov P, Stoyanov J (2017) Lin's condition for functions of random variables and moment determinacy of probability distributions. C R Acad Bulgare Sci 70(5): 611-618.
5. Gwo Dong Lin (2017) Recent developments on the moment problem. Journal of Statistical Distributions and Applications 4(5): 17.
6. Stoyanov J (2004) Stieltjes classes for moment–indeterminate probability distributions. J Appl Probab 41: 281-294.

How to cite this article: Tibor KPogany. On a statistical approximation model of probabilitydensity function of non-negative random variables. Biostat Biometrics Open Acc J. 2018; 8(4): 555741. DOI: 10.19080/BBOAJ.2018.08.555741

7. Jukic D, Bensic M, Scitovski R (2008) On the existence of the nonlinear weighted least squares estimate for a three–parameter Weibull distribution. Comput Statist Data Anal 52(9): 4502-4511.

8. Gavriliadis PN, Athanassoulis GA (2009) Moment information for probability distributions, without solving the moment problem, II: Main-mass, tails and shape approximation. J Comput Appl Math 229(1): 7-15.