On The Continuous Poisson Distribution

Salah H Abid[*], Sajad H Mohammed

Mathematics Department, Education College, Al-Mustansiriya University, Baghdad, Iraq
*Corresponding author: abidsalah@gmail.com

Abstract There are no scientific works deal directly and Extensively with the continuous Poisson distribution (CPD). There are some of rare allusions here and there. In this paper we will take this issue on our responsibility. We consider here the continuous Poisson distribution. Different methods to estimate CPD parameters are studied, Maximum Likelihood estimator, Moments estimator, Percentile estimator, least square estimator and weighted least square estimator. An empirical study is conducted to compare among these methods performances. We also consider the generating issue. Other empirical experiments are conducted to build a model for bandwidth parameter which is used for Poisson density estimation.

Keywords: Continuous Poisson distribution, MLE, Percentile estimator, bandwidth selection, density estimation, AR(1) model

Cite This Article: Salah H Abid, and Sajad H Mohammed, "On The Continuous Poisson Distribution." *International Journal of Data Envelopment Analysis and *Operations Research**, vol. 2, no. 1 (2016): 7-15. doi: 10.12691/ijdeaor-2-1-2.

1. Introduction

The Poisson distribution is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time and/or space, distance, area and volume, if these events occur with a known average rate and independently of the time since the last event.

Because the processes that generate the Poisson distribution are Countless, this distribution is one of the famous widely used distributions in practical applications [8,12].

In various applied research papers, many authors extensively use what they call a continuous Poisson distribution, providing this term with very different, not always correct meanings.

For example, by the "continuous Poisson distribution (CPD)", Some authors use to this end the Gamma distribution, Alissandrakis, Dialetis and Tsiropoula, (1987) [2] and Herzog et al. (2010) [9] or even simply the exponential distribution, Webb (2000) [24] as exact functions of CPD.

From the strictly formal point of view, the above distributions can not be regarded as genuine continuous analogues of the classical Poisson law, since these distributions have probabilistically little in common with this law.

Some other authors, Ilienko, and Podkalyuk, (2010) [10] and Turowski (2010) [21] defined an absolutely continuous distribution with the density of the form,

$$f(x) = c_\lambda \frac{e^{-\lambda}\lambda^x}{\Gamma(x+1)}, x > 0 \tag{1}$$

where c_λ is a normalizing constant. Actually, there is no one determined c_λ exactly. This task will be considered as an essential aim of this paper. The CPD properties will surely follow this step.

In this paper we will refer to continuous Poisson distribution by $X \sim CP(\lambda)$, which is mean that the random variable X follow continuous Poisson distribution with parameter λ.

2. Continuous Poisson Distribution

By continuous Poisson distribution with parameter $\lambda > 0$, Ilienko in 2013 [10] defined the probability measure π_λ supported by $[0, \infty)$ with distribution function of the form,

$$F(x) = \frac{\Gamma(x, \lambda)}{\Gamma(x)}, x > 0 \tag{2}$$

Where $\Gamma(x, \lambda) = \int_\lambda^\infty e^{-t} \, t^{x-1} \, dt$, is an upper incomplete gamma function and $\Gamma(x, 0) = \Gamma(x)$ is the ordinary gamma function.

Now, by using the following facts,

1. $\dfrac{\partial \Gamma(s, y)}{\partial s} = \Gamma(s, y).Ln(y) + yT(3, s, y)$, where

$T(m, s, y) = G_{m,m-1}^{m,0}\left(\begin{array}{c} 0,0,\ldots,0 \\ s-1,-1,\ldots,-1 \end{array}\Big| x\right)$ is the meijer G-function.

2. $\dfrac{\partial}{\partial x} Ln\Gamma(x) = \psi(x) = \dfrac{\Gamma^{'}(x)}{\Gamma(x)}$ is the digamma function.

One can derived the probability density function of $X \sim CP(\lambda)$ as follows,

$$f(x) = \frac{\left(\Gamma(x,\lambda).Ln(\lambda) + \lambda T(3,x,\lambda)\right)\Gamma(x) - \Gamma(x,\lambda)\Gamma'(x)}{\Gamma^2(x)}$$

$$= \frac{1}{\Gamma(x)}\left\{\Gamma(x,\lambda)\left[Ln(\lambda) - \psi(x)\right] + \lambda\, T(3,x,\lambda)\right\} \tag{3}$$

$$= F(x)\left[Ln(\lambda) - \psi(x)\right] + \lambda.\frac{T(3,x,\lambda)}{\Gamma(x)}, x > 0.$$

For a fixed $k \in N$, Ilienko in 2013 [10] defined the k-order moment function of the continuous Poisson distribution as follows,

$$E(x^k) = k\int\limits_0^\lambda e^{-t}\left(\int\limits_0^\infty \frac{x^{k-1}t^{x-1}}{\Gamma(x)}dx\right)dt. \tag{4}$$

Since $\gamma(x,\lambda) = \Gamma(x) - \Gamma(x,\lambda) = \int_0^\lambda e^{-t}\,t^{x-1}\,dt$ is a lower incomplete gamma function, then one can develop the above formula in term of the cumulative distribution function $P(x,\lambda) = \gamma(x,\lambda)/\Gamma(x)$ of Gamma random variable at λ with shape parameter x and scale parameter 1,

$$E(x^k) = k\int\limits_0^\infty \frac{x^{k-1}}{\Gamma(x)}\left(\int\limits_0^\lambda e^{-t}t^{x-1}dt\right)dx$$
$$= k\int\limits_0^\infty \frac{x^{k-1}}{\Gamma(x)}\gamma(x,\lambda)dx. \tag{5}$$

Now, since, $\gamma(x,\lambda) = \lambda^x \sum\limits_{r=0}^\infty \frac{(-\lambda)^r}{r!(x+r)}$, it is Possible to simplify (5) to be,

$$E(x^k) = k\int\limits_0^\infty \frac{x^{k-1}}{\Gamma(x)}\lambda^x \sum\limits_{r=0}^\infty \frac{(-\lambda)^r}{r!(x+r)}dx$$
$$= k\sum\limits_{r=0}^\infty \frac{(-\lambda)^r}{r!}\int\limits_0^\infty \frac{\lambda^x}{(x+r)}\frac{x^{k-1}}{\Gamma(x)}dx. \tag{6}$$

It is appropriate to state that for calculation purposes of $E(x^k)$, we used one of the most powerful Monte Carlo variance reduction techniques to solve the above integral empirically. This technique is the Correlated Sampling.

The reliability function of the continuous Poisson random variable X is,

$$R(x) = 1 - \frac{\Gamma(x,\lambda)}{\Gamma(x)} = \frac{\gamma(x,\lambda)}{\Gamma(x)}. \tag{7}$$

And the hazard function is,

$$h(x) = \gamma^{-1}(x,\lambda)\begin{Bmatrix}\Gamma(x,\lambda)\left[Ln(\lambda) - \psi(x)\right]\\ + \lambda\, T(3,x,\lambda)\end{Bmatrix}. \tag{8}$$

3. Parameter Estimation of CPD

The main aim of this section is to study some of different estimators of the unknown parameter λ of CPD.

(1) *Maximum Likelihood estimators(**MLE**).*

If $x_1, x_2, ..., x_n$ is a random sample from $CP(\lambda)$, then the likelihood function is,

$$L = \prod\limits_{i=1}^n f(x_i) = \prod\limits_{i=1}^n \frac{1}{\Gamma(x_i)}\begin{Bmatrix}\Gamma(x_i,\lambda)\left[Ln(\lambda) - \psi(x_i)\right]\\ + \lambda\, T(3,x_i,\lambda)\end{Bmatrix}.$$

And then the log-likelihood function is,

$$\text{Ln } L = \sum\limits_{i=1}^n -Ln - Ln\left(\Gamma(x_i)\right)$$
$$+ \sum\limits_{i=1}^n Ln\{\Gamma(x_i,\lambda)[Ln(\lambda) - \psi(x_i)] + \lambda\, T(3,x_i,\lambda)\}$$

So, the normal equation become,

$$\frac{\partial Ln\, L}{\partial \lambda} = \sum\limits_{i=1}^n \frac{\begin{Bmatrix}\Gamma(x_i,\lambda).(1/\lambda) - [Ln(\lambda) - \psi(x_i)]\lambda^{x_i-1}e^{-\lambda}\\ + \lambda T'(3,x_i,\lambda) + T(3,x_i,\lambda)\end{Bmatrix}}{\{\Gamma(x_i,\lambda)[Ln(\lambda) - \psi(x_i)] + \lambda\, T(3,x_i,\lambda)\}}.$$

After equating the above equation with zero, we using the numerical solution to obtain $\hat{\lambda}_{MLE}$ as ML estimator of λ.

(2)*The exact moments estimators (**EME**).*

The *method of moments* is a technique for constructing estimators of the parameters that is based on matching the sample moments with the corresponding distribution moments.

Here we provide the method of moments estimator of λ parameter of a

CPD when it is unknown. If X follows $CP(\lambda)$, then By equation E(x), in equation (6) which represent the population mean with the sample mean $\frac{\sum_{i=1}^n x_i}{n}$ and using the numerical solution for the resulting equation, one can get $\hat{\lambda}_{MLE}$ which is the moments estimator of λ.

(3) *Estimators based on percentiles (**PE**).*

Kao in (1959) [11] originally explored this method by using the graphical approximation to the best linear unbiased estimators. The estimators can be obtained by fitting a straight line to the theoretical points obtained from the distribution function and the sample percentile points. In the case of a CPD, it is possible to use the same concept to obtain the estimator of λ based on percentiles because of the structure of its distribution function. Since $F(x)$ defined in (2), Firstly, we find numerically the value of x where $x = F^{-1}(x,\lambda)$, since p_i is the estimate of $F(x_{(i)},\lambda)$. $\hat{\lambda}_{PE}$ can be obtained by minimizing

$$\sum\limits_{i=1}^n \left[x_{(i)} - F^{-1}(P_i,\lambda)\right]^2 \text{ with respect to } \lambda,$$

$$\text{where } E\left[F\left(x_{(i)}\right)\right]$$

$$= p_i = \frac{i}{n+1}\text{is the most used estimator of } F\left(x_{(i)}\right).$$

For some other choices of $p_i's$, see Mann, Schafer and Singpurwalla (1974)[13].

(4) *Least squares (**LSE**) and Weighted least squares (**WLSE**) estimators*

This method was originally suggested by Swain, Venkatraman and Wilson (1988) [20] to estimate the parameters of Beta distribution. Suppose $x_1, x_2, ..., x_n$ is a random sample of size n from a distribution function $G(.)$

and suppose $x_{(i)}$ $(i = 1,2, ..., n)$ denotes the ordered sample. This method uses the distribution of $G(x_{(i)})$. For a sample of size n, we have [13],

$$Var\left(G(x_{(i)})\right) = j(n - j + 1) / \left((n+1)^2 (n+2)\right)$$

and

$$Cov\left[\left(G(x_{(i)})\right), \left(G(x_{(k)})\right)\right]$$
$$= j(n - k + 1) / \left((n+1)^2 (n+2)\right) \text{ for } j < k.$$

So, in the case of CPD, one can obtain the **LS** estimator of λ, say $\hat{\lambda}_{LSE}$ by minimizing, $\sum_{j=1}^{n}\left(F(x_{(i)} - j/(n+1)\right)^2$ with respect to the unknown parameter λ.

The weighted least squares estimator of λ, say $\hat{\lambda}_{WLSE}$ can be obtained by minimizing, $\sum_{j=1}^{n} w_j \left(F(x_{(i)} - j/(n+1))\right)^2$ with respect to λ, where

$$w_j = 1/Var\left(F(y_{(i)})\right) = (n+1)^2 (n+2) / \left(j(n - j + 1)\right).$$

3.1. Generating Continuous Poisson-Distributed Random Variables

It is well known that if the time intervals between like events are exponentially distributed, the number of events occurring in a unit interval of time has the Discrete Poisson distribution. We can use this relationship between the exponential and Discrete Poisson distributions to generate deviates from the Discrete Poisson distribution. A Discrete Poisson deviate x can be defined in the following manner,

$$\sum_{i=1}^{x} y_i \leq 1 \leq \sum_{i=1}^{x+1} y_i$$

Where $y_1, y_2, ..., y_{x+1}$ are random deviates from the exponential distribution having mean $1/\lambda$ and are generated by inverse transform technique $y_i = -Ln(U_i)/\lambda$, where U_i is from the standard continuous uniform distribution. In summary, the cumulative sums are generated until the inequality holds. When this occurs, x is the Discrete Poisson random deviate desired.

By using the above argument, one can generate deviates from the continuous Poisson distribution x^* as,

$$x^* = x + \left(1 - \sum_{i=1}^{x} y_i\right) / y_{x+1} \qquad (9)$$

3.2. The Empirical Study and Discussions

We conduct extensive simulations to compare the performances of the different methods, stated in section 5, for estimating unknown parameter λ of CPD, mainly with respect to their mean square errors (**MSE**) for different sample sizes and different parameter values. The experiments are conducted according to run size $K = 1000$.

The results are reported in table (1). From the table, we observe that,

1) The MSE's decrease as sample size increases in all methods of estimation. It verifies the asymptotic unbiasedness and consistency of all the estimators.

2) The performances of **ELSE**, **LSE** and **EME** are according to their order.

3) For small (n=10) sample size and moderate (n=20) sample size, it is observed that **PE** method works the best whereas the second best method is **MLE**.

4) For large (n=50, 100) sample size, it is observed that **MLE** works the best from all other methods to estimate λ whereas the second best method is **PE**.

Table 1. Empirical MSE to estimate the CP distribution parameter λ

parameter		λ				
sample size	The method	0.3	0.6	0.9	1	1.2
10	MLE	9.805536	7.8975	7.701642	7.600554	7.189884
	EME	11.770434	9.875034	9.66654	9.552816	9.110556
	PE	8.131266	6.204276	6.027372	5.913648	5.446116
	LSE	11.53035	9.63495	9.426456	9.300096	8.870472
	WLSE	10.854324	8.933652	8.737794	8.598798	8.200764
20	MLE	8.074404	6.185322	5.97051	5.825196	5.40189
	EME	11.599848	9.704448	9.489636	9.356958	8.921016
	PE	7.979634	6.046326	5.850468	5.6862	5.351346
	LSE	11.448216	9.552816	9.344322	9.25587	8.819928
	WLSE	10.829052	8.927334	8.712522	8.611434	8.18181
50	MLE	7.594236	5.743062	5.52825	5.382936	4.984902
	EME	11.53035	9.628632	9.439092	9.325368	8.857836
	PE	7.859592	5.964192	5.755698	5.641974	5.187078
	LSE	10.942776	9.015786	8.819928	8.706204	8.301852
	WLSE	10.765872	8.870472	8.674614	8.541936	8.143902
100	MLE	7.411014	5.515614	5.300802	5.206032	4.744818
	EME	11.454534	9.559134	9.344322	9.262188	8.838882
	PE	7.733232	5.850468	5.679882	5.547204	5.098626
	LSE	10.797462	8.87679	8.668296	8.56089	8.099676
	WLSE	10.734282	8.788338	8.586162	8.504028	8.061768

4. Empirically Selection of Bandwidth for CP Kernel Density Estimation

Let $X_1, X_2, \ldots, X_n$ denote a random sample of size n from a random variable with density $f(.)$. The kernel density estimate of f at the point x is given by[18],

$$\widehat{f}_h(x) == \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) \quad (10)$$

where the smoothing parameter h is known as the bandwidth and the kernel K is generally chosen to be a unimodal probability density symmetric about zero. In this case, K satisfies the following conditions [16,17],

$$\int K(y)dy = 1,$$
$$\int y\, K(y)dy = 0,$$
$$\int y^2 K(y)dy > 0.$$

Actually there are a lot of popular choices of kernel function K; uniform, triangular, biweight, triweight, Epanechnikov, normal, and others [18]. The Epanechnikov kernel is optimal in a mean square error sense [6].

The bandwidth controls the smoothness of the density estimate and highly influence its appearance. Selecting a suitable h is a pivotal step in estimating $f(x)$. There has been great advancement in recent years in data-based bandwidth selection for kernel density estimation. Some "second generation" methods, including plug-in [16,22,23] and smoothed bootstrap techniques [5,7,19], have been developed that are far superior to well-known "first generation" methods, such as rules of thumb [17,23], least squares cross-validation [3,15], and biased cross-validation [4,15]. A lot of authors recommend a "solve-the-equation" plug-in bandwidth selector as being most trustworthy in terms of comprehensive rendering [17,18,19,21]. First generation methods for bandwidth selection were mostly proposed before 1990.

4.1. The Empirical Study and Discussions

A simulation Experiment was conducted to find a suitable model to represent bandwidth series for Poisson density estimation, by using Matlab software according to the following assumptions,

a. Generate observations from continuous Poisson distribution with $\lambda = 1$. In simulation experiments, it is commonly that the investigator take the parameter (s) values (real) around (up and down) some critical values to get a wide range of conclusions. Here we considered $\lambda = 1$ as critical point, so it's estimates surely will be near to 1, a little higher (smaller) than 1. This is without loss of generality, our aim.
b. The sample sizes $4 \leq n \leq 150$.
c. The Run size value was $k = 500$.
d. The bandwidth value h was calculated according to the closing window algorithm.
e. For reasons of computational efficiency, we use the Epanechnikov kernel, $K(x) = (3/4)\,(1 - x^2)$, $|x| < 1$
f. To determine h, one can use the **window closing method**. This method permits to detect progressively about the shape of the density. The idea beyond the mechanism of "window closing" is to begin with an arbitrary but big h and compute the resulting density estimate. We then calculate a further series of density estimates by closing the window, i.e. by gradually decreasing h. The incipient estimate based on a large h will be a very smooth function, but as we decrease h, the estimates will gradually display more detaile and become more "eccentric" in form. By examining the manner in which the estimates change as we decrease h, we may be able to detect the point at which the smoothing has been relaxed "too far", and this will then enable us to select a suitable value for bandwidth.
g. we will compute an empirical values of smoothing parameter (bandwidth) h, mean μ, standard deviation σ, interquartile range Q, skewness s_k and kurtosis k_r.

The results of simulaion experiment are recorded in Table 2.

Table (2). empirical values of smoothing parameter (bandwidth) $\breve{h}$, mean $\breve{\mu}$, standard deviation $\breve{\sigma}$, interquartile range $\breve{Q}$, skewness $\breve{s}_k$ and kurtosis $\breve{k}_r$

n	$\breve{h}$	$\breve{\mu}$	$\breve{\sigma}$	$\breve{Q}$	$\breve{s}_k$	$\breve{k}_r$
4	0.235985	0.8772	0.5497	0.7031	0.8664	1.647
5	0.1344	0.7978	0.5081	0.7292	1.177	2.374
6	0.25666	0.8571	0.4772	0.6545	0.8445	1.99
7	0.216244	0.7827	0.4781	0.6643	0.9554	2.33
8	0.175217	0.7423	0.4571	0.495	1.152	2.786
9	0.291795	1.099	1.152	0.9381	1.803	4.771
10	0.257345	1.08	1.088	0.6742	1.956	5.47
11	0.38594	1.251	1.178	0.9381	1.356	3.317
12	0.325977	1.244	1.124	0.9325	1.441	3.683
13	0.475411	1.311	1.103	1.1921	1.254	3.34
14	0.560527	1.361	1.076	1.5096	1.129	3.196
15	0.515786	1.311	1.055	1.346	1.23	3.453
16	0.472587	1.3	1.02	1.3199	1.302	3.741
17	0.430748	1.266	0.9975	1.1562	1.395	4.024
18	0.405257	1.281	0.9697	1.1302	1.385	4.156

19	0.380526	1.267	0.9443	1.0538	1.459	4.449
20	0.428974	1.339	0.9745	1.2611	1.207	3.581
21	0.388909	1.307	0.9609	1.0974	1.278	3.773
22	0.366488	1.308	0.9378	1.0645	1.307	3.956
23	0.36626	1.275	0.9304	1.0538	1.36	4.109
24	0.389239	1.294	0.9148	1.124	1.311	4.095
25	0.375581	1.294	0.8956	1.0911	1.339	4.273
26	0.381801	1.318	0.8862	1.1178	1.259	4.146
27	0.411955	1.387	0.94	1.2813	1.093	3.389
28	0.431641	1.415	0.9339	1.3228	1.006	3.243
29	0.426072	1.388	0.9284	1.3295	1.053	3.336
30	0.418114	1.371	0.9172	1.288	1.105	3.466
31	0.397222	1.379	0.903	1.2709	1.092	3.525
32	0.407839	1.411	0.9063	1.3124	0.9853	3.268
33	0.400562	1.397	0.8956	1.2952	1.032	3.384
34	0.39331	1.409	0.8847	1.2537	1	3.394
35	0.385441	1.404	0.872	1.2319	1.03	3.512
36	0.369776	1.399	0.8599	1.1694	1.059	3.63
37	0.354016	1.388	0.8505	1.1477	1.1	3.746
38	0.377685	1.44	0.898	1.2102	1.023	3.307
39	0.385481	1.456	0.8912	1.2517	0.9758	3.257
40	0.383991	1.432	0.8923	1.2734	0.9964	3.292
41	0.392102	1.471	0.9156	1.3149	0.9038	2.96
42	0.385497	1.474	0.9045	1.2932	0.9058	3.022
43	0.386045	1.5	0.9101	1.3051	0.827	2.834
44	0.381994	1.485	0.9047	1.2752	0.8625	2.895
45	0.376581	1.477	0.8959	1.2633	0.8928	2.974
46	0.372976	1.492	0.8912	1.2573	0.8474	2.919
47	0.369719	1.477	0.8873	1.2454	0.8797	2.972
48	0.364993	1.469	0.8796	1.2335	0.9097	3.046
49	0.360574	1.46	0.8726	1.2217	0.9402	3.118
50	0.36185	1.437	0.8792	1.2335	0.9323	3.121
51	0.3568	1.437	0.8704	1.2216	0.944	3.185
52	0.357167	1.416	0.8746	1.2218	0.9445	3.193
53	0.358195	1.394	0.8805	1.2517	0.939	3.195
54	0.366844	1.427	0.9051	1.2932	0.8821	2.952
55	0.367869	1.449	0.911	1.2871	0.8193	2.796
56	0.369822	1.426	0.9192	1.3268	0.8078	2.799
57	0.367346	1.413	0.9162	1.3405	0.8341	2.838
58	0.37344	1.442	0.9347	1.3524	0.7813	2.657
59	0.3723	1.458	0.935	1.4227	0.7335	2.574
60	0.373231	1.437	0.9405	1.4005	0.736	2.582
61	0.368912	1.439	0.9327	1.3964	0.7375	2.62
62	0.364895	1.435	0.9256	1.3524	0.7549	2.671
63	0.36192	1.426	0.921	1.3347	0.7798	2.714
64	0.358409	1.42	0.9149	1.3228	0.8014	2.763
65	0.345426	1.418	0.9078	1.3051	0.8123	2.812
66	0.351998	1.409	0.9041	1.2932	0.8361	2.852
67	0.350851	1.395	0.9038	1.3109	0.8537	2.875
68	0.347493	1.4	0.8979	1.2694	0.8434	2.892
69	0.3461	1.388	0.8969	1.2871	0.8623	2.918
70	0.345056	1.375	0.8967	1.2537	0.8785	2.941
71	0.339948	1.372	0.8907	1.2497	0.8936	2.991
72	0.333689	1.358	0.8925	1.2084	0.9022	3.004
73	0.33121	1.346	0.8923	1.2255	0.9175	3.027
74	0.354472	1.382	0.9391	1.288	0.9769	3.174
75	0.361197	1.404	0.9517	1.3539	0.9259	2.999
76	0.357859	1.405	0.9454	1.3124	0.9277	3.033
77	0.353212	1.404	0.9393	1.2952	0.9385	3.079
78	0.352394	1.412	0.9358	1.3367	0.9148	3.055
79	0.350788	1.402	0.934	1.3539	0.9328	3.085
80	0.347828	1.399	0.9284	1.3124	0.947	3.132

81	0.3459	1.39	0.9256	1.2952	0.9658	3.168
82	0.339545	1.392	0.92	1.2537	0.9652	3.197
83	0.329495	1.388	0.9152	1.2497	0.9819	3.243
84	0.341927	1.404	0.9216	1.2912	0.9358	3.105
85	0.340301	1.396	0.9194	1.2871	0.9532	3.137
86	0.334191	1.395	0.914	1.2456	0.9608	3.176
87	0.326408	1.385	0.9132	1.2497	0.9748	3.2
88	0.316542	1.39	0.9088	1.1872	0.9644	3.207
89	0.306728	1.389	0.9037	1.1831	0.9726	3.247
90	0.329546	1.395	0.9006	1.2456	0.9534	3.23
91	0.334203	1.415	0.9153	1.2725	0.9181	3.078
92	0.332926	1.423	0.9138	1.2989	0.8916	3.037
93	0.331672	1.415	0.9124	1.3022	0.9069	3.061
94	0.329783	1.409	0.9091	1.2758	0.9234	3.096
95	0.32746	1.412	0.9046	1.2725	0.9188	3.114
96	0.327287	1.423	0.906	1.2989	0.8849	3.037
97	0.325327	1.418	0.9025	1.2955	0.9003	3.072
98	0.323319	1.422	0.8987	1.2691	0.8892	3.075
99	0.321065	1.42	0.8943	1.2547	0.8987	3.111
100	0.318857	1.419	0.8899	1.2396	0.9085	3.148
101	0.316712	1.417	0.8857	1.2245	0.9191	3.184
102	0.316615	1.427	0.8872	1.2404	0.8868	3.109
103	0.315189	1.433	0.8849	1.2668	0.8679	3.089
104	0.314643	1.424	0.8851	1.2811	0.8781	3.101
105	0.317492	1.439	0.8948	1.3075	0.8466	2.982
106	0.315426	1.441	0.8907	1.2932	0.8452	3.003
107	0.316015	1.452	0.894	1.3024	0.813	2.919
108	0.314568	1.447	0.8916	1.2947	0.8276	2.945
109	0.312686	1.444	0.8879	1.2855	0.8389	2.977
110	0.310976	1.44	0.8846	1.2778	0.852	3.008
111	0.31015	1.433	0.8839	1.2668	0.8641	3.024
112	0.308872	1.428	0.8818	1.2591	0.878	3.049
113	0.307142	1.425	0.8784	1.2514	0.8899	3.081
114	0.305486	1.422	0.8752	1.225	0.9022	3.112
115	0.305741	1.412	0.8775	1.2327	0.9027	3.114
116	0.303885	1.412	0.8737	1.2176	0.9056	3.14
117	0.303886	1.403	0.8752	1.2253	0.9092	3.145
118	0.304168	1.394	0.8775	1.2389	0.909	3.147
119	0.307747	1.409	0.8893	1.2547	0.8856	3.039
120	0.309012	1.421	0.8945	1.2811	0.8553	2.947
121	0.309543	1.411	0.8975	1.2955	0.8527	2.948
122	0.30856	1.405	0.8961	1.2691	0.865	2.968
123	0.311276	1.4185	0.9055	1.3178	0.8399	2.872
124	0.31099	1.4264	0.9061	1.327	0.8154	2.823
125	0.311404	1.4168	0.9088	1.3311	0.815	2.826
126	0.30968	1.4177	0.9052	1.3219	0.8153	2.845
127	0.308036	1.4159	0.9018	1.3229	0.8237	2.871
128	0.306804	1.4115	0.8996	1.2914	0.8359	2.894
129	0.305333	1.4086	0.8967	1.2873	0.8466	2.92
130	0.303609	1.408	0.8933	1.2609	0.8519	2.945
131	0.302607	1.4033	0.8914	1.2507	0.8639	2.966
132	0.302135	1.3965	0.8914	1.2499	0.8725	2.977
133	0.30067	1.399	0.8884	1.2458	0.8671	2.986
134	0.300076	1.3926	0.888	1.2389	0.8767	2.999
135	0.29965	1.386	0.8881	1.2423	0.8847	3.01
136	0.29266	1.384	0.885	1.1912	0.8928	3.036
137	0.287077	1.377	0.8861	1.1945	0.8967	3.042
138	0.28676	1.3703	0.8862	1.1515	0.9042	3.053
139	0.300051	1.3893	0.9111	1.2497	0.9413	3.149
140	0.3076	1.401	0.9183	1.2648	0.9153	3.054
141	0.306073	1.4017	0.915	1.2725	0.916	3.073
142	0.303714	1.401	0.9118	1.2574	0.9218	3.097

143	0.303576	1.405	0.9101	1.2725	0.9083	3.083
144	0.302854	1.4	0.9092	1.2758	0.918	3.098
145	0.301453	1.3983	0.9063	1.2725	0.9259	3.124
146	0.30056	1.3937	0.9048	1.2456	0.9363	3.143
147	0.29488	1.3947	0.9018	1.254	0.9358	3.159
148	0.28775	1.3923	0.8992	1.1912	0.9452	3.184
149	0.29873	1.401	0.902968	1.257364	0.9196	3.107
150	0.298	1.3966	0.90182	1.24814	0.9295	3.124

Now, we will analyze and discuss the valuable contents of table (2),

1. It is easily to show that the model that represents bandwidth h as a function of mean μ, standard deviation σ, interquartile range Q and sample size n is,

$$h = \frac{\sigma^{0.67} Q^{1.557}}{\mu^{0.728} n^{0.264}} = 1000 \sqrt{\frac{\sigma^{670} Q^{1557}}{\mu^{728} n^{264}}}$$

for practical applications μ, σ and Q can be replaced respectively by sample mean $\bar{X}$, sample standard deviation s, sample interquartile range q.

Following some notices and indicators related with the above model,

a. $\bar{X}$ can be calculated from equation (9) as follows,

$$\bar{X} = \bar{X}_{DP} + \sum_{l=1}^{n} y_{x_j+1}^{-1} + \left(\frac{1}{n\,\lambda}\right) \sum_{j=1}^{n} \sum_{i=1}^{x_j} y_{x_j+1}^{-1} Ln(U_i)$$

Then the model will be in terms of parameter λ also.

b. The model perfectly fits the data by the coefficient of determination, $R^2 = 0.989$.

c. The ANOVA table is,

Source	Sum of Squares	DF	Mean Square	F-Ratio
Model	176.409	4	44.1022	3119.88
Error	2.02143	143	0.0141359	
Total	178.430	147		

It is clear that the model is significant since F calculated is equal to 3119.88 which is more than F tabulated under all common significant levels 0.01, 0.05 and 0.10, where $F_{4,143,0.01} = 3.46$, $F_{4,143,0.05} = 2.44$ and $F_{4,143,0.10} = 1.99$.

The mean square error is $s_e^2 = MSE = 0.014136$ which is mean that the model is quite suitable.

d. The above model is developed one for Silverman's rule of thumb [17,18] which is $h_{SROT} = 0.9 A n^{-1/5}$ where $A = Min\{s, q/1.34\}$. Actually, a lot of computer packages used this rule to determine h, so it is very popular and reliable rule.

2. Among many of time series models, an AR(1) model is the best one to represent the bandwidth h, as follows,

$$Ln(h_t) = 0.99375 Ln(h_{t-1}) \rightarrow h_t = h_{t-1}^{0.99375}.$$

Following some notices and indicators related with the above model,

a. We chose the above model as the best one according to Akaike criterion and the behaviour of autocorrelation function and partial autocorrelation function (see Figure 2).

b. Estimated white noise variance = 0.0249264 with 146 degrees of freedom.

c. For diagnostic checking of residuals, portmanteau chi-square test statistic on first 20 residual autocorrelations = 20.9827 with probability of a larger value given white noise = 0.337758. Since 0.337758 > 0.1, 0.05 and 0.01 (the common significant levels) then the series of residuals over time are random and independent, so it has no model (see Figure 3).

d. Table 3 contains the Pessimistic, optimistic and usual forecasts for bandwidth series. Figure 1 represents these forecasts and the original series of bandwidth.
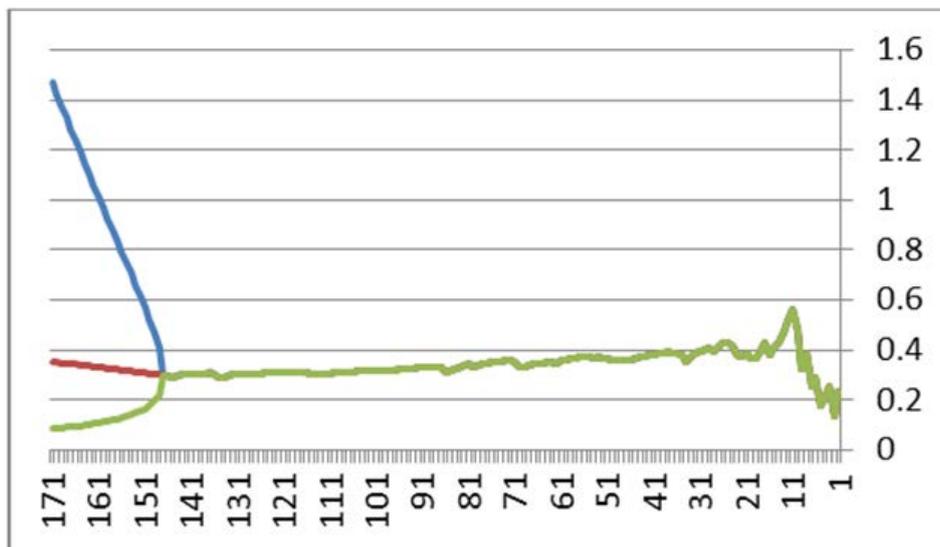

Figure 1. bandwidth series and it's Pessimistic, optimistic and usual forecasts
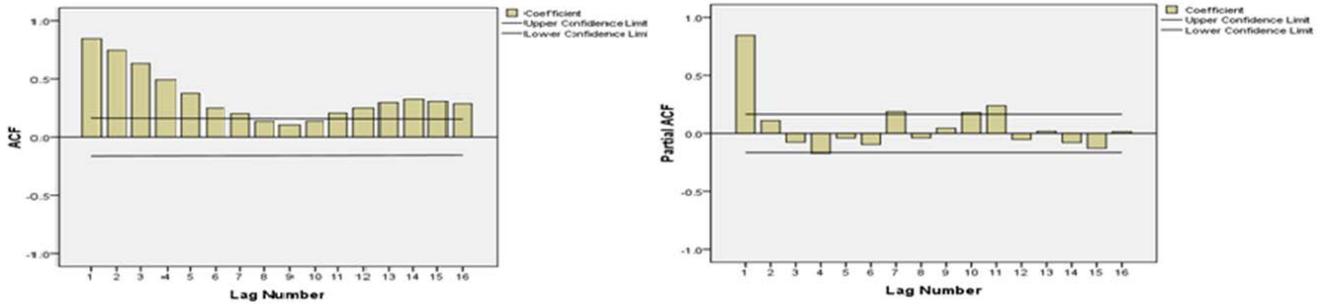
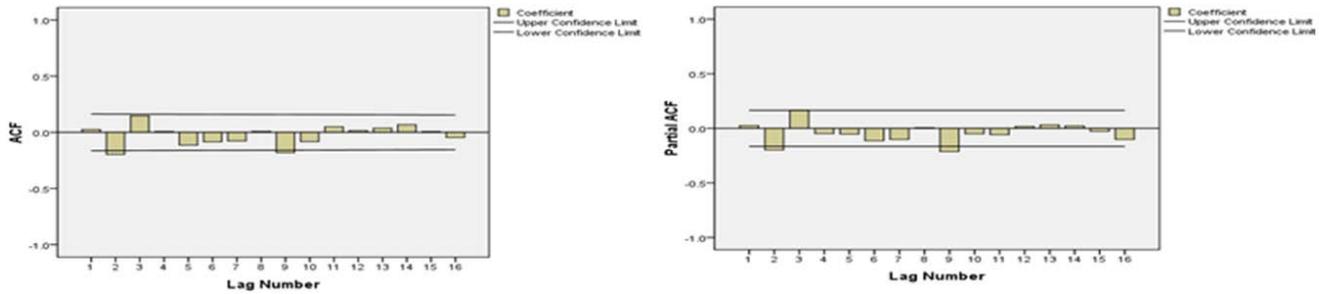**Figure 2.** autocorrelation function (ACF) and partial autocorrelation function (PACF) for bandwidth series

Figure 3. autocorrelation function (ACF) and partial autocorrelation function (PACF) for residuals series

Table 3. Pessimistic, optimistic and usual forecasts

Optimistic forecasts	usual forecasts	Pessimistic forecasts
0.410247	0.300265	0.219766
0.46974	0.302531	0.194843
0.52158	0.304803	0.178121
0.569905	0.307076	0.165458
0.616277	0.30935	0.155284
0.66148	0.311626	0.146811
0.705965	0.31391	0.13958
0.750019	0.316191	0.133299
0.793835	0.318475	0.127768
0.837549	0.320761	0.122844
0.88126	0.323049	0.118423
0.92504	0.325341	0.114425
0.968948	0.327633	0.110784
1.013024	0.329928	0.107453
1.057304	0.332223	0.10439
1.101812	0.33452	0.101564
1.146569	0.336819	0.098945
1.191591	0.339117	0.096511
1.236891	0.341421	0.094242
1.282478	0.343723	0.092122
1.328357	0.346023	0.090136
1.374537	0.348328	0.088271
1.421019	0.350631	0.086517
1.467805	0.352939	0.084864

5. Summary and Conclusions

In view of the great importance of the Poisson distribution in statistical analysis, the continuous Poisson distribution (CPD) is considered here. For CPD, Different methods to estimate continuous Poisson distribution parameter λ are studied, Maximum Likelihood estimator, Moments estimator, Percentile estimator, least square estimator and weighted least square estimator. An empirical study was conducted to compare among these methods. It seemed to us that the Percentile estimator is

the best one for small moderate whereas the maximum likelihood estimator is the best to estimate the parameter for moderate and large samples. Other empirical experiments are conducted to build a model for bandwidth parameter which is used for Poisson density estimation. Two models we obtained to represent bandwidth with high quality properties.

References

[1] Abramowitz, M. and Stegun, I. (1970)." Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables", 9-Revised edition, Dover Publications, USA.

[2] Alissandrakis, C., Dialets, D. and Tsiropoula, G. (1987) "Determination of the mean lifetime of solar features from photographic observations", Astronomy and Astrophysics journal , vol. 174, no. 1-2, March, p. 275-280.

[3] Bowman, A. (1984)." An alternative method of cross-validation for the smoothing of density estimates"; Biometrika, 71; 353-360.

[4] Bowman, A. and Azzalini, A. (1997). "Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations"; Oxford Univ. Press.

[5] Deheuvels, P. (1977). "Estimation nonparamétrique de la densité par histogrammesgeneralizes"; Rev. Statist. Appl. 25 5-42.

[6] Epanechnikov, V.A. (1969). "Non-parametric estimation of a multivariate probability density". Theory of Probability and its Applications 14: 153-158.

[7] Faraway, J. and Jhun, M. (1990). "Bootstrap choice of bandwidth for density estimation", J. Amer. Statist. Assoc. 85 (1990) 1119-1122.

[8] Haight, F. (1967) "Handbook of the Poisson Distribution"; 1st Edition, Wiley, USA.

[9] Herzog, A., Binder, T. , Friedl, F., Jahne, B. and Kostina, A. (2010). "Estimating water-sided vertical gas concentration profiles by inverse modeling" ; International Conference on Engineering Optimization, September 6-9, Lisbon, Portugal.

[10] Ilienko, A. and Klesov, O. (2013). "Continuous counterparts of Poisson and binomial distributions and their properties"; Annals Univ. Sci. Budapest., Sect. Comp. 39 (2013) 137-147.

[11] Kao J (1959). A graphical estimation of mixed Weibull parameters in life testing electron tubes., Technometrics, 1, 389-407.

[12] Kingman, J. (1993). "Poisson Processes", 1st Edition, Oxford Studies in Probability (Book 3), Clarendon Press, UK.

[13] Mann N, Schafer R and Singpurwalla N (1974). Methods for Statistical Analysis of Reliability and Life Data., New York, Wiley.

[14] Scott, D. (1979). "On optimal and data-based histograms"; Biometrika; 66; 605-610.

[15] Scott, D. and Terrell, G. (1987). "Biased and unbiased cross-validation in density estimation"; J. Amer. Statist. Assoc.; 82; 1131-1146.

[16] Sheather, S. and Jones, M. (1991). "A reliable data-based bandwidth selection method for kernel density estimation"; J. Roy. Statist. Soc. Ser.B 53, 683-690.

[17] Sheather, S. (2004). "Density Estimation"; Statistical Science, 19(4), pp. 588-597.

[18] Silverman, B. (1986). "Density Estimation for Statistics and Data Analysis"; Chapman and Hall, London.

[19] Simonoff, J. (1996). "Smoothing Methods in Statistics"; Springer, New York.

[20] Swain J, Venkatraman S and Wilson J (1988). Least squares estimation of distribution function in Johnson's translation system, Journal of Statistical Computation and Simulation, 29, 271-297.

[21] Turowski, J. (2010). "Probability distributions of bed load transport rates: A new derivation and comparison with field data"; water resources research, Volume 46, Issue 8.

[22] Venables, W. and Ripley, B. (2002). "Modern Applied Statistics with S", 4th ed. Springer, New York.

[23] Wand, M. and Jones, M. (1995). "Kernel Smoothing"; London: Chapman & Hall/CRC.

[24] Webb, G. (2000). "MultiBoosting: A Technique for Combining Boosting and Wagging", Machine Learning, 40, 159-39.