



Prediction of the Chances of Death in Covid-19 Data Using the Poisson Process

Rizky Ashgi^{1*}, Sudradjat Supian¹, Subiyanto²

¹*Department of Mathematics, Faculty of Mathematics and Natural Sciences,
Universitas Padjadjaran, Indonesia*

²*Department of Marine Sciences, Faculty of Fishery and Marine Sciences,
Universitas Padjadjaran, Indonesia*

Corresponding author e-mail address: rizky19016@mail.unpad.ac.id

Abstract

Covid-19 has brought about major changes for all people in various countries, for example creating vaccines, wearing masks and predicting the predictive state of death that will occur. In this paper, we will predict cases of covid-19 deaths using data taken from the worldometer website by taking data on daily covid-19 deaths worldwide in the period January 23rd- April 16th, 2020. Then the data is processed using the Poisson process that has been transformed using SPSS computer programming, namely the daily mortality rate in the period January 23rd - March 16th, 2020 using descriptive statistics, it was found that the death rate was 4 people in one day, then the Kolmogorov test followed the Poisson distribution, because it met the requirements for the P-value. value > 5 . Furthermore, it is calculated by using the death process, which is the chance of an event with the chance of death of all the corona suspects in the next 5 days, namely April 21 because the data has been transformed, so $e^{-0.9} = 0.406$. the chance that no one will die within the next 45 days, namely April 30th, 2020 is close to $e^0 = 1$. In the period of January 23rd - April 16th, 2020 using descriptive statistics, it was found that the death rate was 6 people in one day, then the Kolmogorov test was carried out with the results following the Poisson distribution, because it fulfilled the requirements for a P-value > 5 . Furthermore, it is calculated using the death process, which is the chance of an event with the chance of death of all the corona suspects in the next 5 days, namely April 21st, 2020 because the data has been transformed, so $e^{-0.8} = 0.449$. The chance that no one will die within the next 45 days, namely May 31, is close to $e^0 = 1$.

Keywords: Poisson Process, Poisson Distribution, Covid-19

1. Introduction

The Covid-19 virus outbreak in various countries is in the spotlight of the world, because it has claimed many lives and caused difficulties in facing economic demands that must be met on a daily basis, one of the main treatments in dealing with the Covid-19 virus is medical personnel who are at the forefront. In fact, they have also become victims of Covid-19, and all countries have made various efforts to create a vaccine that aims to eradicate the Covid-19 virus. This virus originated in Wuhan, China in one of the traditional markets, until recently as of November 21st, 2020, about 57 million of people worldwide had been infected with severe acute respiratory syndrome coronavirus 2 (SARS-coV-2), and about 1.4 million had died from coronavirus disease 2019 (COVID-19) (Grassin-Delyle et al., 2020). In Diagnosis and treatment plan of corona virus disease 2019 (COVID-19) (Tentative Third Edition) issued by National Health Commission (NHC) of the PRC in January 2020, it is specified that COVID-19 is in the category of pestilence in traditional Chinese medicine (TCM) (Gong et al., 2020).

Calculating the number of human deaths infected with the Covid-19 virus and estimating the number of deaths in a few days can be calculated by applying a mathematical field, namely the stochastic process, the Poisson distribution and the Poisson process which can calculate data on deaths from the Covid-19 virus, the stochastic process. is a collection of random variables at any given point in time, the Poisson process is an example of a stochastic process with the realization in the form of a counting process, meaning that during the Poisson process, the process of counting the number of events is carried out at a certain time. Offshore fire incidents can be modelled as random (stochastic) events by assuming their occurrence follows a Poisson process where the events are independent of each

other and the average time between events is known (Halim et al., 1920). Poisson process is used to model the occurrences of events and the time points at which the events occur in a given time interval, such as the occurrence of natural disasters and the arrival times of customers at a service center (Tse, 2014). For simplicity, it is assumed that the Poisson process has a constant failure/incident rate, expressed by the parameter (Lawrence et al., 2017; Leonenko et al., 2014; Grabski, 2018). A Poisson distribution is the number of successes in a Poisson experiment and is known as a Poisson random variable. The Poisson distribution is useful for modeling and finding the probability for the number of events occurring in a specific interval of time or within a given area of space (Grech and Calleja, 2018). Therefore, the preceding theorem leads to some natural measures of discrepancy with respect to the Poisson distribution (Puig and Weiß, 2020). The Poisson distribution with parameter λ (Pinelis, 2020). In this study, the researchers wanted to estimate the daily case data of deaths from the covid-19 virus in the world taken from worldometer site data using only the pure death model in the average or mortality rate in one day.

2. Methodology

2.1. Materials

In this study, the data of this study are daily data on the death of suspect corona in the world, for the period 23 January 2020 - 16 March 2020 and data with the period 23 January 2020 - 16 April obtained from <https://www.worldometers.info/coronavirus/> where all data around the world directly updated every day.

Table 1. Data on Covid-19 Deaths in the World in One Day

Date	Dead	Date	Dead
Jan-23	8	Mar-06	107
Jan-24	16	Mar-07	105
Jan-25	15	Mar-08	228
Jan-26	24	Mar-09	198
Jan-27	26	Mar-10	271
Jan-28	26	Mar-11	332
Jan-29	38	Mar-12	353
Jan-30	43	Mar-13	447
Jan-31	46	Mar-14	405
Feb-01	45	Mar-15	687
Feb-02	58	Mar-16	642
Feb-03	64	Mar-17	817
Feb-04	66	Mar-18	972
Feb-05	73	Mar-19	1079
Feb-06	73	Mar-20	1356
Feb-07	86	Mar-21	1625
Feb-08	89	Mar-22	1629
Feb-09	97	Mar-23	1873
Feb-10	108	Mar-24	2381
Feb-11	97	Mar-25	2388
Feb-12	146	Mar-26	2791
Feb-13	122	Mar-27	3270
Feb-14	143	Mar-28	3518
Feb-15	143	Mar-29	3204
Feb-16	106	Mar-30	3709
Feb-17	98	Mar-31	4535
Feb-18	136	Apr-01	4889
Feb-19	117	Apr-02	5979
Feb-20	121	Apr-03	5714
Feb-21	113	Apr-04	5799
Feb-22	100	Apr-05	4737
Feb-23	158	Apr-06	5227
Feb-24	81	Apr-07	7631
Feb-25	64	Apr-08	6630
Feb-26	37	Apr-09	7434
Feb-27	58	Apr-10	7165
Feb-28	65	Apr-11	6275
Feb-29	54	Apr-12	5604

Mar-01	73	Apr-13	5600
Mar-02	67	Apr-14	7129
Mar-03	85	Apr-15	8096
Mar-04	83	Apr-16	6813
Mar-05	102		

2.2. Method

2.2.1. Poisson distribution

The Poisson probability of any given number of the variable X calculated with the following formula:

$$P(X) = \frac{\lambda^x e^{-\lambda}}{x!} \quad (1)$$

Where X is the number of the events whose probability is calculated, λ is the average number of events occurring per unit of time or space, e is a constant that is approximately equal to 2.71828, and $x!$ is the factorial number of x (VanPool and VanPool, 2016).

Theorem

An infinitely divisible random variable X has Poisson distribution if and only if

$$P\{X < 0\} = 0, PP\{X = 0\} > 0, E(X - EX)^2 = EX,$$

$$\lim_{n \rightarrow \infty} \frac{\ln P\{X > n\}}{n \ln n} = -1$$

The Poisson distribution is a widely used probability model in the description of count related data that arise very frequently across many disciplines (Rakitzis et al., 2016).

2.2.2 Poisson Process

A Poisson process with parameter (rate) $\lambda > 0$ is a family of random variables $\{N, t \geq 0\}$ satisfying the following properties:

1. $N = 0$
2. $N_{t_1} - N_{t_0}, N_{t_2} - N_{t_1}, \dots, N_{t_n} - N_{t_{n-1}}$ are independent random variables where $0 = t_0 < t_1 < \dots < t_n$.
3. $P(N_t - N_s = n)$

$$\frac{e^{-\lambda(t-s)}}{n!} (\lambda(t-s))^n \text{ for } t > s \quad (2)$$

$N_t = N(0, t]$ can be thought of the number of arrivals up to time t or the number of occurrences up to time t .

We give some properties associated with the Poisson process. The proofs can be found in 1 or 2. if we let $W_n, n \geq 1$ be the time of the n^{th} arrival ($W_0 = 0$), and we let $X_{n+1} = W_{n+1} - W_n, n \geq 1$, be the interarrival time ($X_1 = W_1$).

2.2.3 Stochastic Process

If the stochastic process $\{X(t), t \geq 0\}$ is a Markov chain with the probability of a stationary, state space transition $\{k, k = 0, 1, 2, \dots, n\}$, Fulfill :

1. $X(0) = n$
2. $P(X(t+h) - X(t) = -1 | X(t) = k) = \mu_k h + o(h)$
3. $P(X(t+h) - X(t) \geq -2 | X(t) = k) = o(h)$

Then the process is called a pure death process with parameters $\mu_k, k = 1, 2, \dots, n$.

Write the transition opportunities with the initial conditions $X(0) = n$.

$P_k(t) = P(N(t) = k | X(0) = n) (k = 0, 1, 2, \dots)$ (represents the probability that there are k units remaining in the interval $(0, t]$). Like a pure birth process, the Kolmogorov equation is obtained:

$$P'_n(t) = -\mu_n P_n(t)$$

$$P'_n(t) = -\mu_n P_n(t) + \mu_{k+1} P_{k+1}(t) (k = 1, 2, \dots, n-1)$$

$$P'_0(t) = \mu_1 P_1(t)$$

Especially if $\mu_k = \mu (k = 1, 2, \dots, n)$, $P_n(0) = 1$, $P_k(0) = 0 (k = 1, 2, \dots, n-1)$, so

$$P_0(t) = \sum_{k=1}^n \frac{(\mu t)^{n-k} e^{-\mu t}}{(n-k)!}, (k = 1, 2, \dots, n) \quad (3)$$

and

$$P_0(t) = 1 - \sum_{k=1}^n P_k(t) = 1 - \sum_{k=1}^n \frac{(\mu t)^{n-k} e^{-\mu t}}{(n-k)!} = 1 - \sum_{k=0}^{n-1} \frac{(\mu t)^k e^{-\mu t}}{k!} \quad (4)$$

Flow chart of this research can be seen in Figure 1.

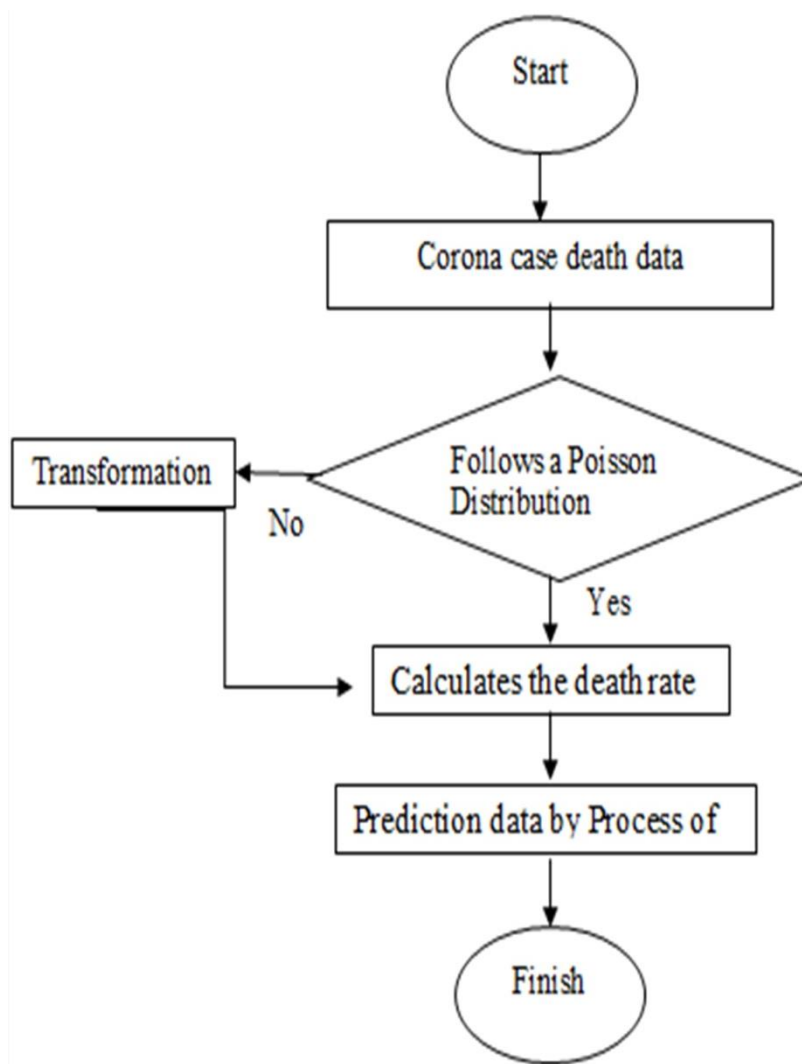


Figure 1. Flow chart of the research

3 Results and Discussion

From calculations using real data, it is known by the Kolmogorov test that the data has not followed the Poisson distribution. This has an impact on the calculation of odds, where the numbers obtained are very large and difficult to calculate, and seem not suitable for describing the real odds, so we transformed the data using the ' \ln ' method, and rounded the results, because the poisson distribution is discrete and states the odds. a number of events. Transformation of data can be seen in Figure 2.

Table 2. Transformation Data on Covid-19 Deaths in the World in One Day

Date	Rounded	Date	Rounded
Jan-23	2	Feb-19	5
Jan-24	2	Feb-20	5
Jan-25	2	Feb-21	5
Jan-26	3	Feb-22	5
Jan-27	3	Feb-23	5
Jan-28	3	Feb-24	4
Jan-29	3	Feb-25	4
Jan-30	3	Feb-26	3
Jan-31	3	Feb-27	4
Feb-01	3	Feb-28	4
Feb-02	4	Feb-29	4
Feb-03	4	Mar-01	4
Feb-04	4	Mar-02	4
Feb-05	4	Mar-03	4
Feb-06	4	Mar-04	4
Feb-07	4	Mar-05	5
Feb-08	4	Mar-06	5
Feb-09	5	Mar-07	5
Feb-10	5	Mar-08	5
Feb-11	5	Mar-09	5
Feb-12	5	Mar-10	6
Feb-13	5	Mar-11	6
Feb-14	5	Mar-12	6
Feb-15	5	Mar-13	6
Feb-16	5	Mar-14	6
Feb-17	5	Mar-15	6
Feb-18	5	Mar-16	6

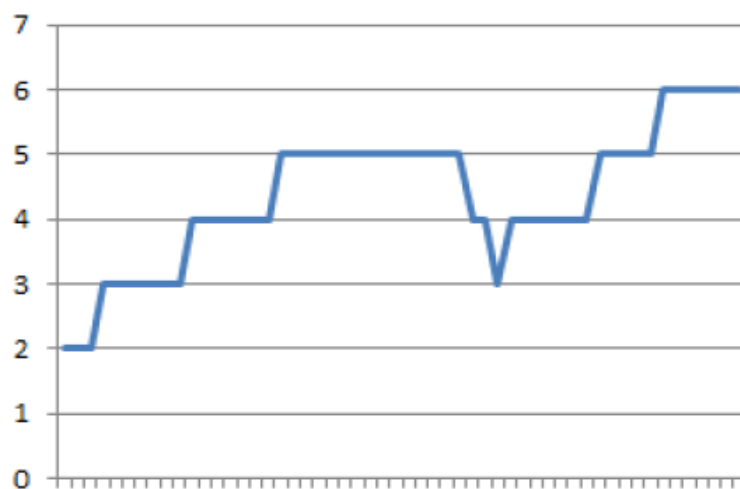


Figure 2. Transformation of data

Table 3. Descriptive Statistics

	Lots of Data	Minimum	Maximum	Mean	Std. Deviation
Death Data	55	2	6	4.29	1.0692

With an average death rate of

$$\mu = \frac{236}{55} = 4.29 = 4 \text{ People/day}$$

Kolmogorov-Sminov test with a hypothesis:

H_0 : P-value > 0.05 then the data is Poisson distributed.

H_1 : P-value > 0.05 then the data is not Poisson distributed.

From the Komogorov-Smirnov test using SPSS with a 95% confidence level, the following results were obtained:

		VAR00003
N		54
Poisson Parameter ^{a,b}	Mean	4.3704
Most Extreme Differences	Absolute	.161
	Positive	.153
	Negative	-.161
Kolmogorov-Smirnov Z		1.183
Asymp. Sig. (2-tailed)		.122

Figure 2. One-Sample Kolmogorov-Smirnov Test

Because P-value > 0,05, then accept H_0 so that the data on the death of suspect corona follows the Poisson distribution;

What is the chance that all corona virus suspects will die within the next 5 days?

$$N = 236$$

$$\mu = 4$$

$$P_0(5) = 1 - \sum_{N=1}^{236} \frac{(5.4)^{236-n} e^{-5.4}}{(236-n)!} = 1 - 1.9 = -0.9$$

So the chances of death of all suspected corona in the next 5 days, April 21st, 2020 because the data has been transformed $e^{-0.9} = 0.406$.

What is the chance that there will be no death in the suspect corona virus within the next 45 days?

Answer

$$N = 236$$

$$\mu = 4$$

$$P_{236}(45) = \frac{(4.45)^{236-236} e^{-4.45}}{(236-236)!} = \frac{1}{e^{180}} = 0$$

So, the chance that no one will die within the next 45 days, namely April 30, 2020 is close $e^0 = 1$.

But if we shorten the time, for example in the next 15 days, the following calculations will be obtained

$$P_{236}(15) = \frac{(4.45)^{236-236} e^{-4.45}}{(236-236)!} = \frac{1}{e^{60}} = 0$$

Then obtained 8.522998×10^{-27} the chance if no one dies within the next 15 days, March 30th, 2020 is 1. Previously we took data that was transformed in the period January 23rd – March 16th, 2020, so now we do the calculations taken in the period January 23rd - April 16th, 2020 (See Table 4).

Table 4. Transformation Data on Covid-19 Deaths in the World in One Day

Date	Rounded	Date	Rounded
Jan-23	2	Mar-04	4
Jan-24	2	Mar-05	5
Jan-25	2	Mar-06	5
Jan-26	3	Mar-07	5
Jan-27	3	Mar-08	5
Jan-28	3	Mar-09	5
Jan-29	3	Mar-10	6
Jan-30	3	Mar-11	6
Jan-31	3	Mar-12	6
Feb-01	3	Mar-13	6
Feb-02	4	Mar-14	6
Feb-03	4	Mar-15	6
Feb-04	4	Mar-16	6
Feb-05	4	Mar-20	7
Feb-06	4	Mar-21	7
Feb-07	4	Mar-22	7
Feb-08	4	Mar-23	8
Feb-09	5	Mar-24	8
Feb-10	5	Mar-25	8
Feb-11	5	Mar-26	8
Feb-12	5	Mar-27	8
Feb-13	5	Mar-28	8
Feb-14	5	Mar-29	8
Feb-15	5	Mar-30	8
Feb-16	5	Mar-31	8
Feb-17	5	Apr-01	8
Feb-18	5	Apr-02	9
Feb-19	5	Apr-03	9
Feb-20	5	Apr-04	9
Feb-21	5	Apr-05	8
Feb-22	5	Apr-06	9
Feb-23	5	Apr-07	9
Feb-24	4	Apr-08	9
Feb-25	4	Apr-09	9
Feb-26	3	Apr-10	9
Feb-27	4	Apr-11	9
Feb-28	4	Apr-12	9
Feb-29	4	Apr-13	9
Mar-01	4	Apr-14	9
Mar-02	4	Apr-15	9
Mar-03	4	Apr-16	9

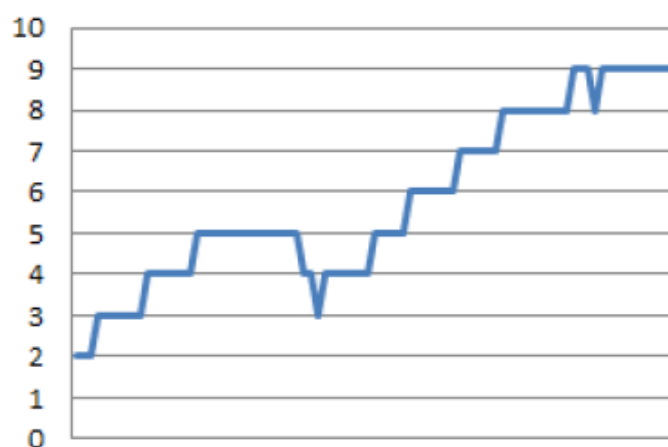
**Figure 3.** Transformation of Data

Table 5. Descriptive Statistics

	Lots of Data	Minimum	Maximum	Mean	Std. Deviation
Death Data	85	2	9	5.78	2.116

With an average death rate of

$$\mu = \frac{492}{85} = 5.78 = 6 \text{ People/days}$$

Kolmogorov-Sminov test with a hypothesis:

H_0 : P-value > 0.05 then the data is Poisson distributed.

H_1 : P-value > 0.05 then the data is not Poisson distributed.

From the Komogorov-Smirnov test using SPSS with a 95% confidence level, the following results were obtained:

		VAR00002
N		85
Poisson Parameter ^{a,b}	Mean	5.7882
Most Extreme Differences	Absolute	.073
	Positive	.073
	Negative	-.067
Kolmogorov-Smirnov Z		.670
Asymp. Sig. (2-tailed)		.760

Figure 4. One-Sample Kolmogorov-Smirnov Test

Because P-value > 0,05, then accept H_0 so that the death data suspect corona follows the Poisson distribution

What is the chance that all coronavirus suspects will die within the next 5 days?

Answer

$$N = 492$$

$$\mu = 6$$

$$P_0(5) = 1 - \sum_{N=1}^{492} \frac{(5.6)^{(492)} e^{-5.6}}{(492-n)!} = -0.8$$

So the chances of death of all suspected corona in the next 5 days, April 21st, 2020 because the data has been transformed $e^{-0.8} = 0.449$.

What is the chance that there will be no death in the suspect corona virus within the next 45 days?

Answer

$$N = 492$$

$$\mu = 6$$

$$P_{492}(45) = \frac{(6.45)^{492-492} e^{-6.45}}{(492-492)!} = \frac{1}{e^{270}} = 0$$

So, the chances of no one dying by May 31st 2020 are 1.

But if we shorten the time, for example in the next 15 days, the following calculations will be obtained

$$P_{492}(15) = \frac{(6.15)^{492-492} e^{-6.15}}{(492-492)!} = \frac{1}{e^{90}} = 0$$

So the chances are that if no one dies until May 1st, 2020 is 1.

4 Conclusion

The chances of all coronavirus suspects dying within the next 5 days are increasingly increasing, it can be seen that the increase in chances from the calculation of the period January 23rd - March 16th, 2020 with the period January 23rd - April 16th, 2020. In the future days, from the calculation for the period of January 23rd - March 16th, 2020 with the period January 23rd - April 16th, 2020 still needs a lot of improvement, or in other words it cannot be calculated using this method, because it has a very large error with real data. So it can be said that the chance of death for suspect corona in the world increases when viewed from the processed data.

References

- Al-Shimmary, A. F. A. (2017). Solving Initial Value Problem Using Runge-Kutta 6th Order Method. *ARPJ Journal of Engineering and Applied Sciences*, 12, 3953-3961.
- Gong, Y. B., Yang, Z. L., Liu, Y., Zhang, Y., Jiang, K., Shi, X. J., & Zhou, J. (2020). Two cases of corona virus disease 2019 (COVID-19) treated with the combination of acupuncture and medication in bedridden patients 针药结合治疗新冠肺炎卧床患者 2 例. *World Journal of Acupuncture-Moxibustion*, 30(3), 171-174.
- Grabski, F. (2018). Nonhomogeneous stochastic processes connected to Poisson process. *Scientific Journal of Polish Naval Academy*, 213(2), 5-15.
- Grassin-Delye, S., Roquencourt, C., Moine, P., Saffroy, G., Carn, S., Heming, N., & Exhalomics Collaborators. (2020). Metabolomics of exhaled breath in critically ill COVID-19 patients: A pilot study. *EBioMedicine*, 63, 103154.
- Grech, V., & Calleja, N. (2018). WASP (Write a Scientific Paper): Probability–Poisson and binomial distributions. *Early human development*, 123, 46-47.
- Halim, S. Z., Quddus, N., & Pasman, H. (1920). Time-trend analysis of offshore fire incidents using nonhomogeneous Poisson process through Bayesian inference. *Process Safety and Environmental Protection*, 147, 421-429.
- Lawrence, E., Vander Wiel, S., Law, C., Spolaor, S. B., & Bower, G. C. (2017). The nonhomogeneous poisson process for fast radio burst rates. *The Astronomical Journal*, 154(3), 117.
- Leonenko, N., Scalas, E., & Trinh, M. (2019). Limit theorems for the fractional nonhomogeneous Poisson process. *Journal of Applied Probability*, 56(1), 246-264.
- Pinelis, I. (2020). Monotonicity properties of the Poisson approximation to the binomial distribution. *Statistics & Probability Letters*, 167, 108901.
- Puig, P., & Weiß, C. H. (2020). Some goodness-of-fit tests for the Poisson distribution with applications in Biodosimetry. *Computational Statistics & Data Analysis*, 144, 106878.
- Rakitzis, A. C., Castagliola, P., & Maravelakis, P. E. (2016). A two-parameter general inflated Poisson distribution: properties and applications. *Statistical Methodology*, 29, 32-50.
- Tse, K. K. (2014). Some applications of the Poisson process. *Applied Mathematics*, 5(19), 3011.
- VanPool, C. S., & VanPool, T. L. (2006). Gender in middle range societies: A case study in Casas Grandes iconography. *American Antiquity*, 53-75.