Vol.9.Issue.3.2021 (July-Sept)
©KY PUBLICATIONS

http://www.bomsr.com
Email:editorbomsr@gmail.com

BULLETIN OF MATHEMATICS AND STATISTICS RESEARCH

A Peer Reviewed International Research Journal

INTERNATIONAL STANDARD SERIAL NUMBER 2348-0580

RESEARCH ARTICLE

THE NORMAL APPROXIMATION TO THE POISSON DISTRIBUTION – A RANDOM SAMPLE APPROACH

RAMNATH TAKIAR
Scientist G – (Retired), National Centre for Disease Informatics and Research (NCDIR), Indian Council of Medical Research (1978-2013), Bangalore – 562110, Karnataka, India & Ph.D. Department, University of Finance and Economics, Ulaanbaatar - 13381 - Mongolia
Email: ramnathtakiar@gmail.com, ramnath_takiar@yahoo.co.in
DOI:10.33329/bomsr.9.3.44

RAMNATH TAKIAR

ABSTRACT

The Poisson distribution is a discrete distribution and is characterized by only one parameter namely "λ", the mean number of successes per given unit of time and space. In literature, while there are reports about the size of the λ that should be acceptable for the initiation of the Normal approximation to the Poisson distribution, there is hardly any mention on the role of the size of the sample in the Normal approximation. In the present study, therefore an attempt is made to check the role of the sample size, if any, in Normal approximation of the Poisson distribution? Two random samples of Poisson distribution are generated for each of the selected λ (2,3,5 and 10) value and for the sample size (50, 100 and 200). Correspondingly, the Normal sample is also generated. It is known that if X follows the Poisson distribution with mean = λ, then Z = $\frac{x-\lambda}{\sqrt{\lambda}}$ is assumed to follow the standard Normal variate. To assess the validation of the Normal approximation, the χ^2 approach is utilized. In general, for each λ (2,3,5 and 10) value and the sample size (50,100 and 200), the Normal approximation to the Poisson distribution is found to be valid. The study has brought out clearly that for the sample size as low as 50 and λ =2, the Normal approximation to the Poisson distribution can be carried out.

Keywords: Poisson distribution, Normal approximation, Random Sample

INTRODUCTION

Poisson distribution is a discrete distribution. In the Poisson distribution, the primary interest lies in the random occurring of an even, often termed as the success. Thus, the Poisson distribution is characterized by only one parameter namely "λ", the mean number of successes per given unit of time and space.

The Probability mass function of the Poisson distribution is given by

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

where x = 0,1,2,3……… n and denotes the number of successes

λ = Average number of successes based on the historical data

In text books, generally, it is mentioned that as $\lambda \rightarrow \infty$, the standard Poisson variate tends to follow the standard normal variate (Gupta and Kapoor 2001, Gupta 2012). Symbolically, If X follows the Poisson distribution with mean = λ, then Z = $\frac{x-\lambda}{\sqrt{\lambda}}$ is assumed to follow the standard Normal variate. Regarding the size of λ, different authors have mentioned different conditions (Dinov 2019, AP Statistics 2014, Calcworkshop 2021, Chaudhary 2017) for the Normal approximation to hold good for the Poisson distribution. In my previous study, it was shown that the Normal approximation to Poisson distribution holds good even when $\lambda \leq 5$ (Takiar 2021). This was in much deviation from the conditions mentioned in the literature. In the present study, an attempt is made further to check the role of sample size, if any, besides the λ value, in achieving the validity of the Normal approximation to the Poisson distribution?

METHODOLOGY

For the study purposes, the four values of λ are selected that is 2, 3, 5 and 10. For each λ value, two random samples of Poisson distribution are generated with the sample size of 50, 100 and 200. Accordingly, six random samples in total are generated for each λ value, using the Excel function, described below.

GENERATION OF RANDOM SAMPLES OF POISSON DISTRIBUTION

In Excel, under the Data Analysis, there are few function programs which are available. Among them, one program is that of Random Number Generation (Microsoft Corporation 2019). On selection of the program, a prompt appears seeking the information on some variables. On providing the relevant information, the random sample is generated. To generate a random sample of size 50 of Poisson distribution with $\lambda = 2$, proceed as follows:

- Number of Variables: 1
- Number of Random Numbers: 50
- Distribution: Poisson (you have option for other distributions, also)
- Parameter (Lambada): 2
- Output Range: C3 (This indicates the cell location in Excel Spread sheet where from the random sample is generated).

THE NORMAL APPROXIMATION TO THE POISSON DISTRIBUTION

If X follows the Poisson distribution with mean = λ, then Z = $\frac{x-\lambda}{\sqrt{\lambda}}$ is assumed to follows N (0,1). As the Poisson distribution is a discrete distribution and the Normal distribution is a continuous distribution, an application of appropriate continuity correction is advocated (Takiar 2021). For the study purposes, for given X values, the Excel functions are utilized to calculate the Poisson and the Normal probabilities (Microsoft Corporation 2019) as shown below:

POISSON.DIST(x, mean, cumulative)

Example: When λ = 3, X = 5, to find P (X = 5), use the Excel function, as follows:

POISSON.DIST(5, 3, False) will give you 0.1008 which is the required probability. For exact probability calculation, against "cumulative" enter "FALSE". In case of cumulative probability that is to find P (X ≤ 5), enter POISSON.DIST(5, 3, TRUE).

NORM.DIST(x, mean, standard_dev, cumulative)

For X = 5, on applying the continuity correction, we get, P (X = 5) = P (4.5 < X < 5.5)

NORM.DIST(4.5, 3, 1.7321, TRUE) = 0.8067

NORM.DIST(5.5, 3, 1.7321, TRUE) = 0.9255

So, P (X=5) = P (4.5 < X < 5.5) = P (X < 5.5) – P (X < 4.5) = 0.9255 – 0.8067 = 0.1188

VALIDATION OF NORMAL APPROXIMATION USING THE χ^2 APPROACH

For each of the random sample generated for the Poisson distribution with given λ and the sample size, the corresponding Normal sample is also generated. To assess the validation of the Normal approximation, the χ^2 approach is utilized. To do so, the relevant sequence of steps that are followed, are given below:

1. From the random sample generated for the Poisson distribution, obtain the frequency distribution by different X values.
2. For each X value, find the corresponding Normal Probability applying the appropriate continuity correction.
3. Thus, obtain the Normal probability distribution.
4. Now, multiplying the above probability distribution, with the selected sample size, the Normal sample is generated.
5. Proceeding this way, corresponding to each random sample generated for given $\lambda\,(2,3,5, and\ 10)$, and the given sample size (50,100, and 200), a Normal sample is also generated.
6. The χ^2 test is used to test for the validity of the Normal approximation.
7. A Non-Significant χ^2 value obtained is taken to confirm the validity of the Normal approximation to the Poisson distribution.
8. On the contrary, a Significant χ^2 value obtained is taken to confirm that the Normal approximation to the Poisson distribution is not valid.

RESULTS

The random samples of Poisson distribution, generated for (λ = 2) and the sample size of 50, 100 and 200 are shown in Table 1. The results of goodness of fit of Normal distribution to the random samples of Poisson distribution are also provided in the Table 1.

For each λ value and the sample size, the Normal fit is found to be valid for the Poisson distribution.

Table 1: The Comparison between the Random Samples of the Poisson distribution and the Fitted Normal distribution by the Selected Sample sizes - (λ = 2)

X	Sample size = 50			Sample size = 100			Sample size = 200		
	Sample 1	Sample 2	Normal fit	Sample 1	Sample 2	Normal fit	Sample 1	Sample 2	Normal fit
0	7	6	7	9	14	14	21	21	29
1	9	12	11	28	35	22	44	46	44
2	19	14	14	30	20	28	57	52	55
3	10	11	11	18	19	22	38	43	43
4	2	4	5	13	7	10	26	25	21
≥ 5	3	3	2	2	5	4	14	13	8
χ^2 value	2.66	0.25	-	5.15	8.84	-	7.31	5.87	-
Degrees of freedom	4	4	-	4	4	-	5	5	-
Significance at 5%	NS	NS	-	NS	NS	-	NS	NS	-

Table 2: The Comparison between the Random Samples of the Poisson distribution and the Fitted Normal distribution by the Selected Sample sizes - (λ = 3)

X	Sample size = 50			Sample size = 100			Sample size = 200		
	Sample 1	Sample 2	Normal fit	Sample 1	Sample 2	Normal fit	Sample 1	Sample 2	Normal fit
0	5	1	4	4	5	7	12	11	15
1	10	7	6	16	14	12	36	28	24
2	7	12	10	19	29	19	39	44	39
3	13	13	11	29	16	23	46	49	45
4	9	7	10	14	17	19	30	31	39
5	3	5	6	6	10	12	25	17	24
6	3	5	3	12	9	8	12	20	14
χ^2 value	4.87	2.53	-	10.41	7.26	-	7.85	9.67	-
Degrees of freedom	4	4	-	5	5	-	6	6	-
Significance at 5%	NS	NS	-	NS	NS	-	NS	NS	-

The random samples of Poisson distribution, generated for (λ = 3) and the sample size of 50, 100 and 200 are shown in Table 2. The results of goodness of fit of Normal distribution to the random samples of Poisson distribution are also provided in the Table 2.

Again, for each λ value and the sample size, the Normal fit is found to be valid for the Poisson distribution.

The random samples of Poisson distribution, generated for (λ = 5) and the sample size of 50, 100 and 200 are shown in Table 3. The results of goodness of fit of Normal distribution to the random samples of Poisson distribution are also shown in Table 3.

Table 3: The Comparison between the Random Samples of the Poisson distribution and the Fitted Normal distribution by the Selected Sample sizes - (λ = 5)

X	Sample size = 50			Sample size = 100			Sample size = 200		
	Sample 1	Sample 2	Normal fit	Sample 1	Sample 2	Normal fit	Sample 1	Sample 2	Normal fit
1	1	4	3	6	10	6	10	6	11
2	8	3	4	12	7	7	11	15	15
3	8	10	6	13	12	12	33	22	24
4	9	7	8	19	14	16	32	36	32
5	6	11	9	14	23	18	38	33	35
6	6	8	8	19	15	16	32	33	32
7	3	4	6	9	5	12	19	24	24
8	4	2	4	3	9	7	12	17	15
9	3	1	2	0	3	4	8	10	7
10	2	0	0	5	2	2	5	4	5
χ^2 value	3.57	5.68	-	8.38	13.48	-	6.44	5.46	-
Degrees of freedom	6	5	-	7	8	-	9	8	-
Significance at 5%	NS	NS	-	NS	NS	-	NS	NS	-

As in previous cases, for each λ value and the sample size, the Normal fit is found to be valid for the Poisson distribution.

The random samples of Poisson distribution, generated for (λ = 10) and the sample size of 50, 100 and 200 are shown in Table 4. The results of goodness of fit of Normal distribution to the random samples of Poisson distribution are also shown in Table 4.

As in previous cases, for each λ value and the sample size, the Normal fit is found to be valid for the Poisson distribution.

CONCLUSION

The results of the present study suggest that the Normal approximation to the Poisson distribution is valid even for the lower values of λ as shown in the range of $2 \leq \lambda \leq 10$ and when the sample size is in the range of $50 \leq Sample\ size \leq 200$.

Table 4: The Comparison between the Random Samples of the Poisson distribution and the Fitted Normal distribution by the Selected Sample sizes - (λ = 10)

X	Sample size = 50			Sample size = 100			Sample size = 200		
	Sample 1	Sample 2	Normal fit	Sample 1	Sample 2	Normal fit	Sample 1	Sample 2	Normal fit
< 5	1	-	2	4	6	4	8	7	8
5	1	5	2	3	5	4	9	5	7
6	3	5	3	8	6	6	11	15	11
7	2	8	4	8	9	8	21	16	16
8	5	2	5	14	8	10	16	22	21
9	6	8	6	14	12	12	27	24	24
10	9	6	6	10	22	13	28	24	25
11	3	3	6	10	10	12	21	23	24
12	6	4	5	12	5	10	24	20	21
13	8	4	4	6	7	8	7	19	16
14	2	1	3	3	5	6	13	12	12
15	1	2	2	4	1	4	7	3	7
> 15	3	2	2	4	4	3	8	10	8
χ^2 value	3.46	4.23	-	4.74	11.55	-	16.3	2.97	-
Degrees of freedom	5	5	-	9	9	-	12	11	-
Significance at 5%	NS	NS	-	NS	NS	-	NS	NS	-

REFERENCES

[1]. AP Statistics Curriculum 2007 Limits Norm2Poisson (2014): Normal Approximation to Poisson Distribution,
http://wiki.stat.ucla.edu/socr/index.php/AP_Statistics_Curriculum_2007_Limits_Norm2Poisson Accessed on 26/06/2021.

[2]. CalcWorkshop (2001): Normal Approximation – w/5 Step-by-Step Examples -
https://calcworkshop.com/continuous-probability-distribution/normal-approximation/Accessed on 26/09/2021

[3]. Chaudhary CR (2017): Normal Approximation to Poisson(λ) Distribution
https://vrcacademy.com/tutorials/normal-approximation-poisson-distribution/Accessed on 26/06/2021

[4]. Dinov ID (2019): Normal Approximation to Poisson(λ) Distribution
http://www.socr.ucla.edu/Applets.dir/NormalApprox2PoissonApplet.html Accessed on 26/06/2021

[5]. Gupta SC (2012): Fundamental of Statistics, 7th Edition, Himalaya Publishing House, 14.38.

[6]. Gupta SC and Kapoor VK (2001): Fundamentals of Mathematical Statistics, Tenth revised Edition – Reprint 2001, Sultan Chand & Sons, 7.57

[7]. Microsoft Corporation, 2019. Microsoft Excel, Available at: https://office.microsoft.com/excel.

[8]. Takiar R (2021): The Normal Approximation to the Poisson Distribution – How Large λ Should Be? – Bulletin of Mathematics and Statistics Research – 9(3)-1.