On the Computation of Poisson Probabilities - Summary

S. D. Chagaboina, J. A. Carrasco, V. Suñé

June 14, 2019

1 Introduction

The Poisson distribution is a distribution commonly used in statistics and in operations research [1, 2, 3] and plays a central role in the analysis of the transient behavior of continuous-time Markov chains (see, e.g., [4]). Let $\lambda > 0$. A random variable X_λ is said to have a Poisson distribution with parameter λ if

$$\Pr[X_\lambda = n] = \frac{\lambda^n}{n!}\,\mathrm{e}^{-\lambda}, n = 0, 1, \ldots\,. \tag{1}$$

Broadly speaking, the methods that have been published for the computation of the probability $\Pr[X_\lambda = n]$ fall into two categories: Methods aimed at the computation of a whole set of probabilities and methods aimed at the computation of a single probability. Here, we target the computation of a single probability.

2 State of the Art

To the best of the authors' knowledge, the methods described in [5, 6, 3, 7] constitute the state of the art for the computation of a single probability $\Pr[X_\lambda = n]$.

The method proposed in [3] consists in using eq. (1) with n replaced by either a Stirling approximation or a polynomial approximation. The method can suffer from numerical overflow when n is not very close to λ and can suffer from numerical underflow when λ is large.

Eq. (1) can be rewritten in several ways. Among them, we find

$$\begin{aligned}\Pr[X_\lambda = n] &= \mathrm{e}^{-\lambda + n\log\lambda - \log\Gamma(n+1)} &(2)\\ &= \mathrm{e}^{n-\lambda+n\log\frac{\lambda}{n} - \log \mathrm{G}(n)} &(3)\\ &= \mathrm{e}^{n-\lambda-n\,\mathrm{lns}(\lambda,n)-\log\mathrm{G}(n)},\ n \geq 1\,, &(4)\end{aligned}$$

where Γ denotes the gamma function, $\mathrm{G}(n) = \frac{\Gamma(n+1)}{\left(\frac{n}{\mathrm{e}}\right)^n}$, $n \geq 1$ is a scaled gamma function, and, denoting $\operatorname{log1p} x = \log x$,

$$\mathrm{lns}(\lambda, n) = \begin{cases} -\operatorname{log1p}\left(\frac{\lambda-n}{n}\right) & \text{if } n < \lambda \\ \operatorname{log1p}\left(\frac{n-\lambda}{\lambda}\right) & \text{if } n \geq \lambda \end{cases}.$$

The method proposed in the well-known book [7] consists in using Eq. (2) together with an accurate approximation for $\log\Gamma(n+1)$. The method is inaccurate when λ and n are large.

The method proposed in [5] consists in using Eq. (3) together with an accurate approximation for $\log \mathrm{G}(n)$. The method is inaccurate when λ and n are large and $n \approx \lambda$.

The method proposed in [6] consists in using Eq. (4) together with the same approximation for $\log \mathrm{G}(n)$. The method is inaccurate when λ is small and when λ is large and n is very close to λ but different from it.

3 Our Contribution

We develop a new method for the computation of a single probability $\Pr[X_\lambda = n]$ that is accurate and fast for a wide range of (λ, n) pairs. In the method, the probability is approximated differently depending on the value of λ and n at hand.

For the set of (λ, n) pairs satisfying $2^{-43} \leq \lambda \leq 2^9$, $n \leq 22$, the probability is computed using eq. (1) straightforwardly.

For the set of (λ, n) pairs satisfying $0 < \lambda < 2^{43}$ or $\lambda > 2^9$, and $0 < n \leq 22$, the probability is computed using

$$\Pr[X_\lambda = n] = \frac{1}{\sqrt{2\pi n}} \mathrm{e}^{(n-\lambda) - n\,\mathrm{lns}(\lambda,n) + \log\frac{\mathrm{G}(n)}{\sqrt{2\pi n}}}\,. \tag{5}$$

The correctly rounded values for $\log\frac{\mathrm{G}(n)}{\sqrt{2\pi n}}$, $0 < n \leq 22$ are computed beforehand using the MPFI multiprecision interval arithmetic library [8] and stored.

For the set of (λ, n) pairs satisfying $\lambda > 0$ and, $n > \max\{22, 1.5\lambda\}$, or $n = \lambda > 22$, or $22 < n < \frac{\lambda}{1.5}$, the probability is approximated using eq. (5) together with an accurate approximation for $\log\frac{\mathrm{G}(n)}{\sqrt{2\pi n}}$. The approximation is such that, assuming exact arithmetic, the incurred relative error is $\leq 2^{-53}$.

Let

$$\mathrm{f}(x) = -x + (x+1)\,\mathrm{log1p}\,x\,.$$

For the set of (λ, n) pairs satisfying $\max\{\lambda, 22\} < n \leq 1.5\lambda$, the probability is approximated using

$$\Pr[X_\lambda = n] = \frac{1}{\sqrt{2\pi n}} \mathrm{e}^{-\left(\lambda\,\mathrm{f}\left(\frac{n-\lambda}{\lambda}\right) + \log\frac{\mathrm{G}(n)}{\sqrt{2\pi n}}\right)}$$

together with accurate approximations for both $\mathrm{f}\left(\frac{n-\lambda}{\lambda}\right)$ and $\log\frac{\mathrm{G}(n)}{\sqrt{2\pi n}}$. The approximations are such that, assuming exact arithmetic, the incurred relative error is $\leq 2^{-53}$.

Let

$$\mathrm{g}(x) = \frac{x - \mathrm{log1p}\,x}{1+x}\,.$$

For the set of (λ, n) pairs satisfying (λ, n) pairs satisfying $\max\{\frac{\lambda}{1.5}, 22\} \leq n < \lambda$, the probability is approximated using

$$\Pr[X_\lambda = n] = \frac{1}{\sqrt{2\pi n}} \mathrm{e}^{-\left(\lambda \, \mathrm{g}\left(\frac{\lambda - n}{n}\right) + \log \frac{\mathrm{G}(n)}{\sqrt{2\pi n}}\right)}$$

together with accurate approximations for both $\mathrm{g}\left(\frac{\lambda-n}{n}\right)$ and $\log \frac{\mathrm{G}(n)}{\sqrt{2\pi n}}$. The approximations are such that, assuming exact arithmetic, the incurred relative error is $\leq 2^{-53}$.

To assess the accuracy of the new method, we performed the following numerical experiment. First, we devised a procedure that using the MPFI library, allows us to compute tight rigorous bounds for $\Pr[X_\lambda = n]$. Next, we chose a representative set of values for the λ parameter, namely $10^{-1}, 10^{0}, \ldots, 10^{15}$, and for each value of λ we obtained, using the procedure, the set of n values, $\mathcal{N}(\lambda)$, for which $\Pr[X_\lambda = n]$ does not underflow.[1] Finally, for each set $\mathcal{N}(\lambda)$, we chose 100 values of n uniformly distributed between the endpoints of the set and for each such n computed, using the procedure referred to previously, tight rigorous bounds for $\Pr[X_\lambda = n]$ and, using them and the approximation yielded the method, $\widetilde{P}_n(\lambda)$, computed as a measure of accuracy the index

$$d = -\log_{10} \left| \frac{\Pr[X_\lambda = n] - \widetilde{P}_n(\lambda)}{\Pr[X_\lambda = n]} \right|. \tag{6}$$

Te experiment was also performed for each of the methods described in [7, 5, 6]. The results are reported in table 1. Also shown in the table are the average CPU times required by each method to estimate a single probability $\Pr[X_\lambda = n]$ as a function of λ. As we can see, the method seems to be slightly faster than the remaining methods and more accurate, with reductions in the actual relative error by more than ten orders of magnitude in some cases. This improvement in accuracy is still more noteworthy if we take into account that in the experiment, we are dealing with λ up to 10^{15} and are estimating probabilities that can be close to $2.2 \cdot 10^{-308}$.

References

[1] F. A. Haight, *Handbook of the Poisson distribution.* John Wiley & Sons, 1967.

[2] K. Krishnamoorthy, *Handbook of statistical distributions with applications.* Chapman and Hall/CRC, 2016.

[3] N. L. Johnson, A. W. Kemp, and S. Kotz, *Univariate discrete distributions*, vol. 444. John Wiley & Sons, 3rd ed., 2005.

[4] K. S. Trivedi, *Probability & Statistics with Reliability, Queuing and Computer Science Applications.* PHI Learning Pvt. Limited, 2011.

[1] We consider that a floating-point number x underflows if it is smaller than $2^{-1\,022} \approx 2.2 \cdot 10^{-308}$, which is the smallest normal number that can be encoded using the `binary64` floating-point format defined by the IEEE 754 floating-point arithmetic standard [9].

	meth. [7]			meth. [5]			meth. [6]			prop.		
λ	d_m	d_M	t	d_m	d_M	t	d_m	d_M	t	d_m	d_M	t
10^{-1}	12.9	15.7	0.27	4.42	6.69	0.18	3.49	6.69	0.17	13.0	17.7	0.12
10^{0}	13.0	16.5	0.25	4.42	7.02	0.18	2.30	7.02	0.17	12.9	18.1	0.11
10^{1}	12.5	15.7	0.26	4.42	7.49	0.18	4.06	7.49	0.18	13.0	17.0	0.12
10^{2}	12.1	14.9	0.27	4.42	8.21	0.19	2.30	8.21	0.18	13.0	16.5	0.12
10^{3}	11.4	14.5	0.27	6.43	9.32	0.19	6.43	9.32	0.18	12.6	17.4	0.13
10^{4}	10.5	12.9	0.27	10.2	10.9	0.19	10.2	10.8	0.17	12.3	16.5	0.18
10^{5}	9.57	12.2	0.27	11.0	13.2	0.19	11.8	13.9	0.17	12.8	16.6	0.15
10^{6}	8.57	10.8	0.27	9.98	14.6	0.18	11.1	14.7	0.17	12.8	16.5	0.14
10^{7}	7.32	10.6	0.27	8.97	15.4	0.18	10.7	15.4	0.17	12.8	17.3	0.13
10^{8}	6.26	9.71	0.27	7.96	15.3	0.18	10.2	15.3	0.17	12.6	16.6	0.13
10^{9}	5.18	8.06	0.27	6.96	14.7	0.18	9.68	14.7	0.17	12.6	16.4	0.12
10^{10}	4.32	7.15	0.26	5.96	15.1	0.19	9.37	15.1	0.17	12.7	16.7	0.12
10^{11}	3.09	6.24	0.30	4.97	16.4	0.18	8.74	16.4	0.17	12.6	17.1	0.12
10^{12}	2.23	4.97	0.28	3.96	15.0	0.19	8.28	15.0	0.17	12.7	17.0	0.12
10^{13}	1.11	4.11	0.27	2.95	15.9	0.19	7.69	15.9	0.17	12.8	16.9	0.12
10^{14}	0.00	2.90	0.27	1.96	14.6	0.19	7.25	14.6	0.17	12.8	17.2	0.11
10^{15}	0.00	1.24	0.27	0.94	15.2	0.18	6.63	15.2	0.17	12.7	16.5	0.11

Table 1: Comparison of the methods described in [5, 6, 7] with the method proposed in the paper. For each method and for each value of λ we report, from left to right, the minimum value of the accuracy index (6) d obtained for 100 values of n distributed uniformly between the endpoints of the in sets $\mathcal{N}(\lambda) = \{n \geq 0 \, : \, \Pr[X_\lambda = n] \geq 2^{-1\,022}\}$, the maximum value of d over the same n values, and the average CPU time obtained as a function of λ.

[5] J. R. Whittlesey, “Incomplete gamma functions for evaluating Erlang process probabilities,” *Mathematics of Computation*, pp. 11–17, 1963.

[6] L. Knüsel, “Computation of the chi-square and Poisson distribution,” *SIAM Journal on Scientific and Statistical Computing*, vol. 7, no. 3, pp. 1022–1036, 1986.

[7] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press, 2007.

[8] N. Revol and F. Rouillier, “Motivations for an arbitrary precision interval arithmetic and the MPFI library,” *Reliable Computing*, vol. 11, pp. 275–290, Aug 2005.

[9] IEEE, *Standard IEEE Std-754 for floating-point arithmetic*, 2008.