Overview — Input modelling for non-stationary discrete-event simulation

In order to make principled operational decisions, there is a great interest in many areas of industry in understanding systems that exhibit some form of randomness or uncertainty. Such systems are generally modelled mathematically, with understanding stemming from *performance measures* extrapolated from the model, usually numerical summaries of some characteristic of interest. An illustrative example we will continually refer to is a hospital accidents and emergency (A&E) department model, where patients arrive at uncertain and irregular times and queue to be tended to by staff. Adequate capacity and staffing can be inferred by asking questions such as "How many beds do we require in order to have leftover capacity 95% of the time?" and encoding them into performance measures.

Often, unfortunately even for unsophisticated mathematical models, "the desired performance measures are intractable or there is no numerical approximation whose error is bounded"[1]. A solution is then to simulate the system. For systems that exhibit uncertainty, the field of *stochastic simulation* has developed around this requirement.

A simulation has three main components. The first is the *input*, which represents samples from an approximation of the random processes governing the system. The second is the *model*, specifying rules governing the interactions among the inputs. These produce the third component, the *output*, which represents the performance measure we seek. A simulation may have several inputs and outputs. For the A&E example, the inputs could be representations of how patients arrive at the hospital and the length of time they are treated for, the model could specify how the patients queue before being tended to, and the outputs could be the maximum number of beds in use at a particular run of the simulation.

Since the inputs come from approximations to these random processes, they carry an inherent level of error which propagates to the outputs. In applications such as the aforementioned A&E example, there is a critical need to obtain very representative inputs. If they are not, and demand is underestimated, then a decision might be made that leaves the department understaffed or with too few beds, putting patients' lives at risk. At the same time, we cannot overestimate demand as it means drawing resources away from other wards. Therefore, capturing the behaviour of the system inputs accurately is extremely important.

This report focuses on the problem of *input modelling*, that is modelling and estimating the sources of randomness that underlie the system. Approaches that attempt to capture the structure of a commonly used model for discrete arrivals such as ones to the A&E queue, the *Poisson process*, are reviewed. While this is a problem with a rich history in statistics, where the only goal is getting a good estimate and quantifying errors, the added challenge of needing to generate new arrivals efficiently for simulation purposes makes this an active area of research in stochastic simulation.

[1] Page 1 of "Foundations and Methods of Stochastic Simulation" by Barry L. Nelson (Springer, 2013).

Input modelling for non-stationary discrete-event simulation

Tamás Papp, supervised by Dr Lucy E. Morgan

1 Introduction

Simulation is commonly used to replicate the behaviour of mathematically modelled systems that have analytically intractable performance measures. In industrial applications such as healthcare or manufacturing, systems that involve discrete events occurring over time are typically simulated. Examples of such systems include call centers or accident and emergency departments, which are usually formulated using queuing models. Here, individuals arrive at random times, queue to be served by a set of servers, then leave the service station after a random service time.

Most queuing models assume that arrivals follow a Poisson process with some (possibly time-varying) rate function. If we would like to simulate the queuing model to gain knowledge about the system, this rate function will have to be representative of the real behaviour of individuals' arrivals to the system. This is particularly important in critical applications such as healthcare, where patients' lives may be at risk if misinformed decisions are made based on a simulation.

This report focuses on input modelling in simulations where the inputs come from a non-homogeneous Poisson process. Motivation for work in this area comes from successful applications such as Pritsker et al. (1995), who use Poisson processes to simulate the United States organ donation system. While Poisson process rate function estimation is well-established in statistics, with eminent statistician David Cox working in the area as early as the 1950s, the added challenge of needing to simulate from the estimated Poisson process means this is an active area of research in simulation.

The report is organised as follows. Section 2 introduces Poisson processes, with properties that are relevant for simulation and methods for generating arrivals. Section 3 focuses on approaches of modelling Poisson processes. Section 4 contains practical remarks and offers suggestions for future work, and section 5 summarises.

2 The Poisson process

A commonly used mathematical model for discrete event simulation assumes arrivals to the system follow a *homogeneous Poisson process* (HPP) with constant arrival rate $\lambda \geq 0$. In this model, all inter-arrival times are independent and identically distributed exponential random variables with rate λ and the total number of arrivals $N(t)$ by time t is distributed as

$$N(t) \sim \text{Poisson}\left(\lambda t\right).$$

for $t \in [0, T]$. This assumption of stationarity is, however, unsuitable in most practical applications of interest. An A&E department, for instance, will have very different patient arrival rates depending on the time of day, with more frequent arrivals during the day and evening and sparser ones in the early morning. To convince ourselves that the assumption of stationarity is wrong, we can collect data on the system and run a statistical test for the stationarity hypothesis.

For the simulation to be realistic, it is therefore necessary to capture the non-stationarity of the arrivals, so the next best thing is to instead assume a *Poisson process* (NHPP) with non-negative rate function $\lambda(\cdot)$ generates them. The total arrivals by time t are distributed as

$$N(t) \sim \text{Poisson}\left(\Lambda(t)\right)$$

for $t \in [0, T]$, where $\Lambda(t) := \int_0^t \lambda(y)\mathrm{d}y$ is the integrated rate function. Note that the integrated rate must be a non-negative and non-decreasing function of time.

2.1 Properties of Poisson processes

There are several properties of Poisson processes that make them attractive for simulation, chief among which is their mathematical simplicity: they are completely determined by their rate function $\lambda(\cdot)$ or, equivalently, the integrated rate function $\Lambda(\cdot)$. Another is their superposition property: if N_1 and N_2 are independent Poisson processes with rate functions $\lambda_1(\cdot)$ and $\lambda_2(\cdot)$, then the superposed Poisson process $N_1 + N_2$ has rate function $\lambda(\cdot) = \lambda_1(\cdot) + \lambda_2(\cdot)$. Conversely, a decomposition property holds: if N is Poisson with rate $\lambda(\cdot) = \lambda_1(\cdot) + \lambda_2(\cdot)$ where $\lambda_{1,2}(\cdot)$ are non-negative functions, then N can be decomposed into independent Poisson processes N_1, N_2 with corresponding rate functions $\lambda_1(\cdot), \lambda_2(\cdot)$.

There is a shortcoming to using Poisson processes for arrival modelling, even if we allow them to be non-stationary. This is the issue of over- and underdispersion: a Poisson process strongly assumes that the total number of arrivals in any time interval has the same mean and variance. This is plainly wrong in practice, with the variance almost always being higher or lower than the mean, so care must be taken that the application does not deviate too much from the Poisson assumption before constructing a simulation. Chapter 6 of Law (2014) suggests methods to test whether a Poisson model is reasonable at all.

2.2 Generating arrivals

Exact samples from a Poisson Process can be drawn by *inversion* or *thinning*. Inversion is the analogue of inversion sampling for univariate distributions, and works as follows. Generate arrivals $t_1 < t_2 < \ldots$ from a Poisson process with rate constant unit rate. Then $\Lambda^{-1}(t_1) < \Lambda^{-1}(t_2) < \ldots$ are arrivals from a Poisson process with integrated rate function $\Lambda(\cdot)$. A formal treatment of inversion can be found in Çinlar (1975). Integrated rate functions that are linear or exponential can be inverted straightforwardly:

$$\begin{aligned} \Lambda(t) &= at + b &&\implies \Lambda^{-1}(t) = \frac{t-b}{a}, \\ \Lambda(t) &= \exp(at+b) &&\implies \Lambda^{-1}(t) = \frac{\log t - b}{a}. \end{aligned}$$

This naturally extends to rate functions that are piecewise linear or exponential, which makes sampling from processes with such rates convenient. In addition, Klein and Roberts (1984) propose an efficient method of generating arrivals by inversion from a Poisson process with a piecewise linear rate function $\lambda(\cdot)$, equivalent to a piecewise quadratic integrated rate.

Thinning (Lewis and Shedler, 1979) is the analogue of rejection sampling for univariate distributions, uses the decomposition property of Poisson processes, and works as follows. Let $\tilde{\lambda}(\cdot)$ be the rate function of a Poisson process such that $\tilde{\lambda}(t) \geq \lambda(t)$ for all t, that is it *majorises* the rate function $\lambda(\cdot)$. Generate arrivals $t_1 < t_2 < \ldots$ from the Poisson process with rate function $\tilde{\lambda}(\cdot)$ and accept each t_i with probability $\lambda(t_i)/\tilde{\lambda}(t_i)$. The resulting "thinned" set of arrivals come from a Poisson process with rate function $\lambda(\cdot)$. Thinning is particularly effective if the following two criteria are met. One is the acceptance ratio being as close to 1 as possible over the domain of $\lambda(\cdot)$, as rejecting often will slow down arrival generation severely. The other is the ease of generating arrivals from the majorising process — this is met if for instance $\tilde{\lambda}(\cdot)$ is piecewise constant, linear or exponential.

3 Modelling rate functions

If the inputs to our discrete-event simulation come from a non-homogeneous Poisson process, our goal is to estimate either $\lambda(\cdot)$ or $\Lambda(\cdot)$ as accurately as possible, as errors in the inputs propagate to the outputs. At the same time, we are also interested in efficiently generating arrivals from the fitted Poisson processes, which greatly limits the type of estimates that can be considered in practice.

We assume data is collected from m repeats of the process over $[0, T]$, which could be data over several days of a queuing system with a period of one day. The rates will be modelled based on this data, which is either recorded as a sequence of individual arrivals or a sequence of counts over intervals. The statistical approach to this problem is to assume a (parametric) family of functions $\lambda(\cdot) \in \{\lambda_\theta(\cdot) : \theta \in \Theta\}$ and then estimate by maximum likelihood. We henceforth drop subscripts in estimates of the rate function to avoid cluttering the notation.

In the interval count case, $(0, T]$ is divided into M windows $(t_{i-1}, t_i]$, each containing n_i arrivals for $i \in \{1, \ldots, M\}$, where $t_0 = 0$ and $t_M = T$. We assume independence between windows, an assumption that is justified if they are chosen independently to the arrival process (e.g. a hospital logging arrivals every two hours). The log-likelihood is, up to a constant,

$$l(\lambda) = \sum_{i=1}^{M} n_i \log \left(\int_{t_{i-1}}^{t_i} \lambda(u) \mathrm{d}u \right) - m \int_0^T \lambda(u) \mathrm{d}u.$$

If we additionally assume independent parameters across windows $(t_{i-1}, t_i]$, maximum likelihood estimates are any functions that satisfy

$$\int_{t_{i-1}}^{t_i} \lambda(u) \mathrm{d}u = \frac{n_i}{m} \tag{1}$$

for all $i \in \{1, \ldots, M\}$. Identity (1) forms the basis of many estimates for arrival counts, such as the basic piecewise constant estimate $\lambda(t) = n_i / [m(t_i - t_{i-1})]$ for $t \in (t_{i-1}, t_i]$. Methods that improve piecewise constant estimates (Chen and Schemiser 2013, 2017) are appropriate for such a setting and are discussed in more detail in §3.2. Nicol and Leemis (2014) fit a continuous piecewise linear function for interval counts while maintaining constraint (1), formulating and solving a quadratic programming problem. We dwell on this no further, focusing instead on the setting of arrival time data as it allows for more precise modelling.

For the remainder of this report, arrival times are assumed to recorded individually, unless stated otherwise. If arrival times $t_1 < \ldots < t_M$ in m repeats of $[0, T]$ are observed, the log-likelihood is, up to a constant,

$$l(\lambda) = \sum_{i=1}^{M} \log \lambda(t_i) - m \int_0^T \lambda(u) \mathrm{d}u. \tag{2}$$

If we would like to encourage our function to have a certain structure, such as a degree of smoothness, we can penalise this log-likelihood, as in Morgan et al. (2019).

We now begin to consider methods for the arrival time case from the simulation literature, which can be broadly classified into three categories: exponential models for the rate, piecewise models for the rate and models for the integrated rate.

3.1 Exponential models for $\lambda(\cdot)$

The NHPP rate function $\lambda(\cdot)$ must be non-negative, and by introducing this constraint the likelihood optimisation step is made more challenging. One way around this issue is to model $\lambda(\cdot) = \exp(g(\cdot))$, which enforces positivity, but also makes this approach unsuitable if the rate function should be zero over some sub-intervals (such as in a call centre with a lunch break). The papers of Lewis and Shedler (1976), Lee et al. (1991) and Kuhl et al. (1997) model the rate function in this way, considering increasingly complex models. These are, in order:

$$\lambda(t) = \exp\left(\sum_{m=0}^{r} \alpha_m t^m \right),$$

$$\lambda(t) = \exp\left(\sum_{m=0}^{r} \alpha_m t^m + \sin(\omega t + \phi)\right),$$
$$\lambda(t) = \exp\left(\sum_{m=0}^{r} \alpha_m t^m + \sum_{k=1}^{p} \gamma_k \sin(\omega_k t + \phi_k)\right).$$

The polynomial terms are justified by the arbitrarily good approximation that can be obtained by truncating a Taylor series, while the sine terms attempt to capture any periodicities in the rate. For fixed values of r and p, parameters are estimated by maximum likelihood, with the frequencies possibly being known a priori.

Choosing the degree r of the polynomial and the number of trigonometric terms p represents a statistical model selection problem. Lewis and Shedler (1976) suggest selecting r in a step-wise fashion, at each step testing the significance of adding in an extra term by a likelihood ratio test and stopping when the expanded model is rejected twice consecutively. Lee et al. (1991) and Kuhl et al. (1997) suggest stopping at the first rejection in the above, with the latter authors suggesting fixing p a priori by e.g. a spectral analysis.

While capable of achieving good approximations, all these approaches are plagued by computational complexity at the optimisation step, as pointed out by Morgan et al. (2019). When obtaining parameter estimates via maximum likelihood the objective is highly non-convex, so while Newton's method is applicable, it will not converge to the true optimum unless started from a neighbourhood of it. Finding a suitable neighbourhood requires multiple good starting points, a non-trivial task in and of itself, and for each point an optimisation routine must be run separately.

Sampling from the fitted process by inversion, as proposed by Kuhl et al. (1997), is impractical since $\Lambda^{-1}(\cdot)$ cannot be computed in closed form. For any arrival time t from the unit rate Poisson process to be inverted, $\Lambda^{-1}(t)$ is therefore computed by a bisection search, slowing down arrival generation severely. While we might justify an increase in computational cost when fitting our estimate, we will likely want to simulate many arrivals from the fitted process, so this step should be made as efficient as possible. Lee et al. (1991) propose efficiently sampling from the second representation by thinning and devise an automated procedure to construct a tight-fitting piecewise linear cover, making their approach computationally more tractable.

3.2 Piecewise models for $\lambda(\cdot)$

Similar to estimates for interval count data, approaches that model $\lambda(\cdot)$ directly in the arrival data case often represent it in a piecewise manner. The most widely used such estimate is the piecewise constant one, which divides $(0, T]$ into M windows $(t_{i-1}, t_i]$ for $i \in \{1, \ldots, M\}$. Keeping the same notation as for the interval count case, the estimate is $\lambda(t) = n_i/[m(t_i - t_{i-1})]$ for $t \in (t_{i-1}, t_i]$. Analogous to a histogram, this estimate is sensitive to the window locations and widths and is highly discontinuous, so it does not capture the rate adequately near edges of the windows.

Attempting to mend piecewise constant representations, Chen and Schmeiser (2013) make any given piecewise constant rate smoother by halving each window and shifting the value of each half. They propose an iterative procedure that repeats this smoothing step while maintaining constraints such as non-negativity, constant interval means as in (1), and periodicity, showing that this procedure converges to a continuous function in the limit. The paper of Chen and Schmeiser (2017) follows a similar idea, transforming a given piecewise constant representation into a continuous piecewise quadratic one. In the first stage, they formulate the estimate as the solution of an optimisation problem constrained by interval means (1), continuity and first-derivative continuity. If the rate estimate is negative at any point, an additional procedure is run that transforms this estimate into a non-negative one which maintains interval means and continuity. Both approaches of Chen and Schmeiser work purely on the assumption that a smoother rate function is more representative, and so should be used with care as they are not shown to improve the statistical properties of the original piecewise constant estimate. Sampling from estimates provided by these methods can be done readily by inversion.

Zheng and Glynn (2017) take a more statistically principled approach and consider fitting a continuous piecewise linear approximation by maximum likelihood, assuming that the intervals are known. In the interval count case, they consider both maximum likelihood and ordinary least squares estimates. Sampling can be done quickly by inversion. Morgan et al. (2019) point out there is an issue with the assumption of known intervals in practice, and that often a subjective choice must be made by the modeller when using this procedure in practice.

A very recent contribution by Morgan et al. (2019) considers modelling the rate function as a cubic spline (a piecewise polynomial, with pieces defined in between knots where it is twice continuously differentiable), which is then encouraged to be globally smooth by subtracting a smoothing penalty

$$\theta \int_0^T \lambda''(u)\mathrm{d}u$$

with $\theta \geq 0$ from log-likelihood (2). They use equidistant knots (in which case the splines are *cardinal*) which are fixed a priori and the basis spline decomposition property of cardinal splines,

$$\lambda(t) = \sum_{k=1}^{N} c_k B_k(t),$$

forcing coefficients $c_k \geq 0$ in order to obtain a non-negative rate estimate. Once the knots are fixed, these basis splines are completely determined. For a fixed penalty value θ, the coefficients are determined by a trust region algorithm using a quadratic Taylor series approximation, which is claimed to perform well in practice although it only has good local convergence properties. The joint values $(c_1, \ldots, c_N, \theta)$ are determined by optimising a *regularisation information criterion*, which accounts for the penalty term when determining the most parsimonious model.

An advantage of using cardinal basis splines is that arrival generation can be performed efficiently. By using the decomposition property of the Poisson process, arrivals generated from each individual basis component can be pooled together. In addition, all basis splines are simply translated versions of each other, so only arrivals from the first spline component with full support in $[0, T]$ ($B_4(\cdot)$ in this case) need to be generated, as they can then be translated and scaled to construct arrivals from the entire process. Therefore, arrival generation is reduced to sampling from an NHPP with rate $B_4(\cdot)$, which the authors suggest performing by thinning.

Unaddressed issues for this method include ways of selecting the number of knots, constructing a tight-fitting cover to $B_4(\cdot)$ that can be efficiently inverted, and ensuring the optimisation step converges to the global maximum. The authors offer some guidance for the first point, noting that empirical results indicate more basis splines can offer a more accurate result (a marginal improvement was observed when doubling the number of basis splines from 23 to 46).

3.3 Models for $\Lambda(\cdot)$

A full description of the NHPP can also be obtained by directly modelling the integrated rate function $\Lambda(\cdot)$ instead. Leemis (1991) surveyed non-parametric methods for doing this, including piecewise linear and quadratic constructions. Their linear estimator, which simply interpolates points $(0, 0)$, $(t_i, iM/[(M+1)m])$ for all i, and $(T, M/m)$ is statistically consistent and easy to sample from by inversion. Note this is completely equivalent to a piecewise constant estimator for the rate function $\lambda(\cdot)$.

The problem of estimating the integrated rate $\Lambda(\cdot)$ has received comparatively little attention in recent years. One reason for this might be that only estimates that are readily invertible are worth considering, as sampling from a fitted process by thinning requires modelling $\lambda(\cdot)$ as well. Another, that estimates of the rate function $\lambda(\cdot)$ are more interpretable, allowing for a qualitative investigation of the behaviour of the process over time. Additionally, it is more difficult to adequately capture unknown periodic structure when modeling the integrated rate. Finally, this is a more constrained and therefore difficult problem, as $\Lambda(\cdot)$ is required to be both non-negative and increasing.

4 Remarks and extensions

Having reviewed a number of different approaches for modeling the rate of Poisson arrivals, one might ask how a practitioner would choose among these. If data is particularly plentiful, then even the simple piecewise constant approach would provide a good enough estimate. Otherwise, if particularly great accuracy is desired, then the spline approach of Morgan et al. (2019) or the exponential approach of Lee et al. (1991) would be suitable candidates, at the cost of slower arrival generation. A good middle ground between accuracy and ease of sampling would be the piecewise linear approach of Zheng and Glynn (2017) or the quadratic smoothed one of Chen and Schmeiser (2017).

As far as further work is concerned, this report could grow in several directions. One would be a simulation study, as in Morgan et al. (2019), where a selection of methods could be compared and contrasted on estimating known rate functions with different structure, such as smooth rates with sharp peaks and rates with periods that are unknown a priori. Another would be to put these methods in the greater context of *input uncertainty* in simulation, testing the validity of pointwise confidence intervals and seeing how model misspecification impacts simulation outputs.

More broadly in simulation research, one direction for further work would be to attempt to address some of the issues in the aforementioned methods. For the multi-period exponential model of Kuhl et al. (1997), an automated procedure to construct a piecewise linear cover would allow for more efficient sampling. For the piecewise linear method of Zheng and Glynn (2017), an automated procedure for determining the intervals would be useful for practitioners. For Morgan et al. (2019), a way of constructing a piecewise linear cover for $B_4(\cdot)$ and selecting the number of knots in a principled way are open questions.

In addition, the problem of estimating non-homogeneous *non-Poisson* processes could be investigated for use in simulation, following on from Gerhardt and Nelson (2009). These processes allow arrivals to be over- or under-dispersed compared to the mean, making them more realistic for problems that strongly deviate from the Poisson assumption. Note that any of the aforementioned methods that do not obtain estimates by maximum likelihood can be readily applied in this context.

Finally, an interesting avenue of research could be to investigate deeper connections between thinning and rejection sampling. Adaptive rejection sampling (Wild and Gilks, 1993) constructs an increasingly tighter piecewise exponential cover for efficiently sampling any log-concave distribution. This idea, together with extensions such as the work of Görür and Teh (2011), could be used to automatically construct majorising functions for efficient thinning, enabling a wider class of Poisson rate estimates to be used for simulation.

5 Summary

This report surveyed a number of methods for estimating a nonhomogeneous Poisson process for use in input generation for stochastic simulation. The first class of methods represents the rate function as an exponential, the more recent second class represents the rate in a piecewise manner, and the third class models the integrated rate nonparametrically. Finally, suggestions for future work in this area of simulation were offered, building upon these methods and exploiting connections with rejection sampling in statistics.

References

Chen, H. and B. Schmeiser. (2013). "I-SMOOTH: Interatively Smoothing Mean-Constrained and Nonnegative Piecewise-Constant Functions". *INFORMS Journal on Computing*, 25(3): 432–445.

Chen, H. and B. Schmeiser. (2017). "MNO–PQRS: Max Nonnegativity Ordering—Piecewise-Quadratic Rate Smoothing". *ACM Transactions on Modeling and Computer Simulation*, 27 (3):18:1–18:19.

Çinlar, E. (1975). *Introduction to Stochastic Processes.* Prentice-Hall.

Gerhardt, I. and B.L. Nelson. (2009). Transforming renewal processes for simulation of nonstationary arrival processes. *INFORMS Journal on Computing*, 21(4):630–640.

Görür, D. and Y.W. Teh. (2011). "Concave-Convex Adaptive Rejection Sampling". *Journal of Computational and Graphical Statistics*, 20(3):670–691.

Klein, R.W. and S.D. Roberts. (1984). "A Time-Varying Poisson Arrival Process Generator". *Simulation*, 43(4):193–195.

Kuhl, M.E. et al. (1997). "Estimating and simulating Poisson processes having trends or multiple periodicities". *IIE Transactions*, 29(3):201–211.

Law, A.M. (2014). *Simulation Modeling and Analysis.* McGraw-Hill Education, 5th edition.

Lee, S. et al. (1991). "Modeling and Simulation of a Nonhomogeneous Poisson Process Having Cyclic Behavior". *Communications in Statistsics - Simulation and Computation*, 20(2-3): 777–809.

Leemis, L.M. (1991). "Nonparametric Estimation of The Cumulative Intensity Function for a Nonhomogeneous Poisson Process". *Management Science*, 37(7):886–900.

Lewis, P.A.W. and G.S. Shedler. (1976). "Statistical Analysis of Non-stationary Series of Events in a Data Base System". *IBM Journal of Research and Development*, 20(5):465–482.

Lewis, P.A.W. and G.S. Shedler. (1979). "Simulation of Nonhomogeneous Poisson Processes by Thinning". *Naval Research Logistics Quarterly*, 26:403–413.

Morgan, L.E. et al. (2019). "A Spline-based Method for Modelling and Generating a Nonhomogeneous Poisson Process". In *Proceedings of the 2019 Winter Simulation Conference.*

Nicol, D. M. and L.M. Leemis. (2014). "A Continuous Piecewise-Linear NHPP Intensity Function Estimator". In *Proceedings of the 2014 Winter Simulation Conference*, pages 498–509.

Pritsker, A.A.B. et al. (1995). "Organ Transplantation Policy Evaluation". In *Proceedings of the 1995 Winter Simulation Conference*, pages 1314–1323.

Wild, P. and W.R. Gilks. (1993). "Algorithm AS 287: Adaptive Rejection Sampling from Log-Concave Density Functions". *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 42(4):701–709.

Zheng, Z. and P.W. Glynn. (2017). "Fitting Continuous Piecewise Linear Poisson Intensities via Maximum Likelihood and Least Squares". In *Proceedings of the 2017 Winter Simulation Conference*, pages 1740–1749.