International Conference on Advances in Business and Law
Proceedings Series, ICABL, 2019, Vol. 3, ISSN: 2523-6547
Available at https://publications.ud.ac.ae/index.php/ICABML-CP

3rd International Conference on Advances in Business and Law (ICABL-2019)
23-24 November 2019, Dubai, UAE

Business process simulation and process quality assessment of Czech online bookshop

*Petra Tomanová[a]**

[a]*University of Economics, Prague, Czech Republic, petra.tomanova@vse.cz*
**Corresponding author.*

ABSTRACT

Accurate estimation of times between arrivals is a key aspect of process simulations and thus efficient decision making. Times between arrivals typically exhibit strong autocorrelation structure which is commonly ignored in the queueing theory. In order to capture the time dependence accurately, we utilize a Generalized Autoregressive Score (GAS) model based on the generalized gamma distribution. Once the process of arrivals is estimated, a process assessment can be performed using process simulations. The results from an empirical study of an online bookshop in Prague, Czech Republic pointed out insufficient resources allocated for the pre-processing and the final stage.

Keywords: *arrival process; duration model; generalized autoregressive score model; process simulation*.

JEL codes: *C15, C41*.

1. INTRODUCTION

A common assumption in the queueing theory is that times between arrivals follow the exponential distribution with a constant rate. We relax this assumption by considering a much broader class of generalized gamma distribution. We model times between arrivals by the Generalized Autoregressive score model which captures the clustering of arrivals properly (Tomanová, 2019). Then, the business process simulation and process quality assessment of a particular online bookshop from Prague, Czech Republic are performed. The results from an empirical study pointed out insufficient resources allocated for the pre-processing and the final stage.

2. METHODOLOGY

The key aspect of process simulation is an estimation of times between arrivals. For this purpose, we utilize the Generalized Autoregressive Score (GAS) models of Creal et al. (2013), also known as Dynamic Conditional Score models of Harvey (2013).

GAS models capture dynamics of time-varying parameters $f_i = \left(f_{i,1}, \dots, f_{i,k}\right)'$of a given distribution by the autoregressive term and the scaled score of the conditional observation density (or the conditional observation probability mass function in the case of discrete distribution). The time-varying parameters f_i follow the updating scheme

$$f_{i+1} = C + Bf_i + AS(f_i)\nabla(x_i, f_i), \tag{1}$$

where $C = (c_1, \dots, c_k)'$ are the constant parameters, $B = \mathrm{diag}(b_1, \dots, b_k)'$ are the autoregressive parameters, $A = \mathrm{diag}(a_1, \dots, a_k)'$ are the score parameters and $S(f_i)$ is the scaling function for the score. The following the scaling functions are commonly adopted:

- inverse of the Fisher information, $S(f_i) = \mathcal{J}(f_i)^{-1}$,
- square root of inverse of the Fisher information, $S(f_i) = \mathcal{J}(f_i)^{-1/2}$,
- or simply unite scaling, $S(f_i) = I$,

http://dx.doi.org/10.30585/icabl-cp.v3i1.403

http://dx.doi.org/10.30585/icabl-cp.v3i1.403

where x_i are observed durations, $g = (g_1, \ldots, g_l)$ are static parameters and

$$\mathcal{J}(f_i) = \mathrm{E}[\nabla(x_i, f_i)\nabla(x_i f_i)|f_i, g], \qquad \nabla(x_i, f_i) = \frac{\partial \ln \mathrm{P}[X_i = x_i|f_i, g]}{\partial f_i}.$$

The conditional durations (times between arrivals) x_i are modelled as

$$x_i = \mu_i \epsilon_i,$$

where different specifications of the conditional mean duration $\mu_i = f_{i,1}$ and the distribution of the error term ϵ_i result in different models. In this paper, we consider generalized gamma distribution which is a generalization of the two-parameter gamma distribution. The generalized gamma has three parameters $a > 0, d > 0, p > 0$. The probability density function of the generalized gamma is

$$f(x;\ a, d, p) = \frac{(p/a^d)x^{d-1}\exp\{-(x/a)^p\}}{\Gamma(d/p)},$$

where $\Gamma(\cdot)$ denotes the gamma function. We allow the parameters to be time-varying. When $d = p$ generalized gamma distribution degenerates into the Weibull distribution; and when $p = 1$ it degenerates into the gamma distribution.

Tomanová (2019) showed how the GAS model utilizing generalized gamma distribution is useful in the estimation of arrivals and process simulations. The author argued that the GAS model captures the time dependence in arrivals properly and showed that the proposed approach leads to a more faithful representation of the mean and extreme values of the arrival process than the common approach utilizing exponential distribution with constant rate since the exponential distribution underestimates the probability of extreme values.

3. SIMULATION STUDY

The simulation study of the Czech online bookshop is based on the findings documented in the paper of Tomanová (2019). Once the parameters of the selected GAS model are estimated (Table 1) based on 5,364 observations (times between arrivals) spanning from June 8 to December 28, 2018, the time-varying parameters can be simulated using GAS recursion stated in Equation (1). Times between arrivals for the simulation study are computed based on generalized gamma distribution with estimated time-varying parameters.

Table 1. GAS with generalized gamma distribution

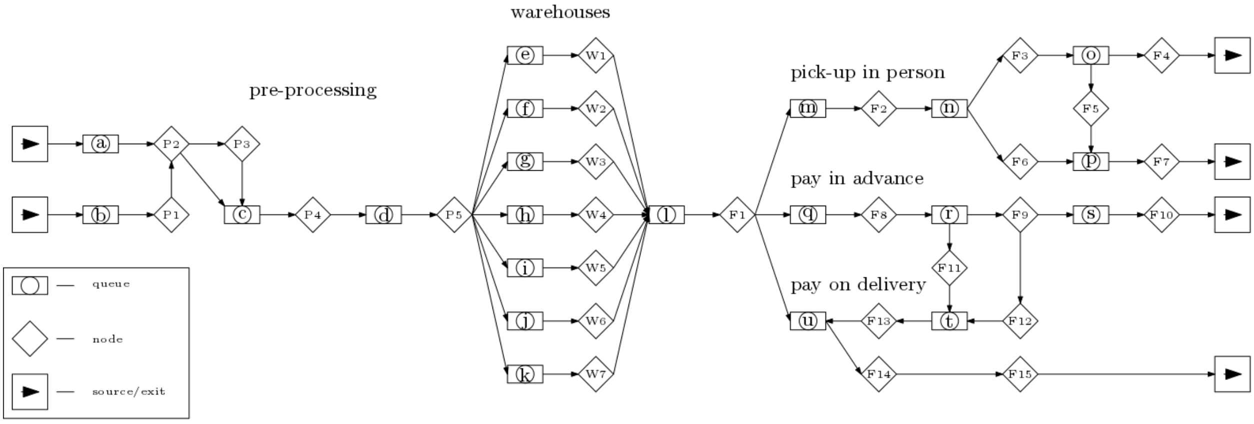
Parameter	c1	c2	c3	b1	b2	a1	a2
Estimate	0.0375	0.3847	-0.3631	0.9848	0.3944	0.0255	0.0275
Std. Error	0.0263	0.1149	0.0567	0.0105	0.1560	0.0080	0.0061
P-value	0.0772	0.0004	0.0000	0.0000	0.0057	0.0007	0.0000

Note: The table presents the estimated parameters of the GAS model with generalized gamma distribution. Source: Tomanová (2019).

Once the process of arrivals is estimated, process mapping has to be performed. The map of processes of Czech online bookshop is depicted in Figure 1. Orders arrive from two sources: website and e-mail. First, the pre-processing is performed (P1 – P4 nodes). Second, books have to be found and bring from one of the seven warehouses (W1 – W7 nodes). Third, the order is finalized and prepared for a customer (F1 – F14 nodes). The corresponding queues are denoted by letters: $\boldsymbol{a} - \boldsymbol{u}$. The number of simulated arrivals is equal to 100,000 which corresponds to 6.5 years approximately.

http://dx.doi.org/10.30585/icabl-cp.v3i1.403

Figure 1. Process mapping



Note. This figure reports the process mapping of the Czech antiquarian bookshop in Prague.

4. ANALYSIS AND RESULTS

In this section, the main findings from the simulation study are discussed and divided into three defined stages:

<u>*Pre-processing*</u>

- The average number of orders waiting in the queue a (which come from the online ordering system) and b (ordered through e-mail) is equal to 26.42 and 1.28 respectively.
- The maximum number of orders waiting in the queue a and b is equal to 155 and 7 respectively. The relatively high numbers are caused by weekend since resources (workers) are available only during working days.
- The 79 % of orders wait less than 7 hours in queue c, i.e. 21 % of orders are not satisfied within a day once they are received.
- Then, 79 % of orders wait up to 15 minutes for the finalization of pre-processing stage and they are prepared to be removed from the storage.

<u>*Warehouses*</u>

- The maximum numbers of orders waiting in the queue e – k are 12, 97, 157, 112, 7, 16, 8 – corresponding to warehouses W1 – W7 respectively.
- The orders are removed from the storages serially (not parallelly) starting with warehouse W1 and ending with W7. The average queueing time in e (warehouse W1) is 1.4 minutes and the maximum time is 16.93 minutes. The highest average queueing time is 135.68 minutes and the maximum time is 323.11 minutes.
- Thus, it takes 2.25 hours on average to remove all pre-processed orders from storage and it takes 5.38 hours in the worst case.

<u>*Final stage*</u>

- The average and the maximum queueing time in the queue l are 149.98 and 209.73 minutes respectively. It means that it takes 2.5 hours on average to sort the orders out and it takes no longer than 3.5 hours.
- The orders are sorted according to the type of delivery into a pick-up in person (queue m), pay in advance (queue q) and pay on delivery (queue u). The finalization of the three types of deliveries is satisfied serially starting with "pick-up in person" and ending with "pay on delivery".
  - The results show that the average and maximum queueing time in the queue m is 62.09 and 1,237.22 minutes respectively. Thus, the "pick-up in person" orders are satisfied approximately within 1 hour after sorting out, however, some orders are not satisfied that day – but it happens only on rare occasions (less than 1% of cases).

- The average and maximum queueing time in the queue q is 189.79 and 1,298.18 minutes respectively. Thus, the “pay in advance” orders are satisfied approximately within 3 hours on average after sorting out, however, some orders are not satisfied that day. This happens in 8% of cases.
- The average and maximum queueing time in the queue u is 279.81 and 2,887.86 minutes respectively. Thus, the “pay on delivery” orders are satisfied approximately within 4.5 hours on average after sorting out, however, some orders are not satisfied that day. This happens in 10% of cases.

- On average, the 171.89 orders are waiting in the queue n to be pick-up by the customer. The maximum number of orders is 275. The average number of waiting orders after the reminder is 11.72 (queue o).
- On average, the 48.33 orders are waiting in the queue r to be paid by the customer. The maximum number of orders is 79. The average number of waiting orders after reminder is 4.93 (queue s).

4. CONCLUSIONS

In this paper, we capture the time dependence in arrivals by the Generalized Autoregressive Score model since it was shown that this approach leads to a more faithful representation of the mean and extreme values of the arrival process (Tomanová, 2019). Using simulation, we analysed the business processes of a particular Czech online bookshop. The main findings pointed out that 21% of orders are not satisfied within a working day due to insufficient allocated resources for the pre-processing stage. Moreover, in 8% and 10% of cases, it happens that the “pay in advance” and “pay on delivery” orders are not satisfied within a day after sorting out respectively. This is caused by insufficient resources allocated for the final stage. The results are useful for managers of the bookshop and will be used for business process optimization.

FUNDING

The study is the part of the project F4/53/2019 funded by the IGA of University of Economics, Prague.

REFERENCES

Creal, D., Koopman, S.J. & Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28(5), 777-795.

Harvey, A.C. (2013). *Dynamic models for volatility and heavy tails: with applications to financial and economic time series* (Vol. 52). Cambridge University Press.

Tomanová, P. (2019). Clustering of Arrivals and Its Impact on Process Simulation. *The 15th International Symposium on Operational Research SOR'19, Proceedings. Ljubljana: Slovenian Society Informatika, Section for Operational Research*, 314-319.