AN UNRECURRENT FLOW WITH TRIANGULAR CONDITIONAL DISTRIBUTION OF INTERARRIVAL TIMES

A. Andronov
Transport and Telecommunication Institute
Riga, Latvia
lora@mailbox.riga.lv

An unrecurrent flow is considered when interarrival times form a Markov chain. A conditional distribution of an interarrival time is a triangular one, for that a parameter is a linear function on a previous interarrival time. The stationary one-dimensional and two-dimensional distributions have been derived. It is shown that possible values of a correlation coefficient between the adjacent interarrival times are in the diapason (-1/3, 1/3).

Keywords: Unrecurrent flow, interarrival time, stationary distribution, dependence.

1. INTRODUCTION

Classical models of queueing, reliability, inventory and etc., suppose Poisson or recurrent flows of considered input claims. But numerous statistical data prove the opposite. For example, it has been experimentally stated that interarrival times of Internet flows are the dependent ones [5]. Analogously, flows of insurance claims for damages have dependent structure [4]. In the first case, a dependence between interarrivals is described by the so-called Batch Markovian Arrival Process [7, 8, 9, 10, 11, 13] or Markov-Additive Processes of Arrivals [12]. In the second case, copulas are used usually for a description of the dependence [1, 2, 3, 4, 11].

In our paper we would like to use natural and direct way for a description of the dependence between continue random variables $X_0, X_1, ...$ those are interpreted as interarrival times of flow claims. We suppose that the last ones correspond to a Markov chain with continue state space $\Omega = (0, 1)$. To describe a transit probability for one step we consider a family $\Psi = \{f_\theta(x) : x, \theta \in \Omega\}$ of triangular distributions that is determined by a probability density function

$$f_\theta(x) = \begin{cases} 2\frac{x}{\theta}, & 0 < x < \theta, \\ 2\frac{1-x}{1-\theta}, & \theta \le x < 1. \end{cases} \tag{1}$$

The corresponding cumulative distribution function is calculated as

$$F_\theta(x) = \begin{cases} \frac{1}{\theta}x^2, & 0 < x < \theta, \\ 1 - \frac{(1-x)^2}{1-\theta}, & \theta \le x < 1. \end{cases} \tag{2}$$

Now for Markov chain $X_1, X_2, ..., X_n, ...$ we suppose the following: if $X_n = x$, then X_{n+1} has a distribution from family Ψ with parameter $\theta = \theta(x)$. In the previous paper of the author [2] a case where $\theta = x$ has been considered. It turns out, that a correlation coefficient between X_n and X_{n+1} equals 1/3. A case $\theta = 1 - x$ gives a value -1/3.
Now we consider a case $\theta = \delta(1-x) + x(1-\delta) = \delta + x - 2\delta x$ where δ is a parameter, $0 < \delta < 1$. Below we show that all value between - 1/3 and 1/3 can be received in such a way.
Our paper has the following structure. At first we find marginal stationary cumulative $G(x) = P\{X \leq x\}$ and density $g(x) = \frac{\partial}{\partial x} G(x)$ distribution function for this Markov chain; then two-dimensional stationary density function $g(x,y) = \frac{\partial^2}{\partial x \partial y} P\{X_n \leq x, X_{n+1} \leq y\}$ and its mixed moment. Numerical illustrations and conclusion end the paper.

2. STATIONARY MARGINAL DISTRIBUTION

We would like to present calculating probability density function $g(x) = G(x)'$ for the stationary marginal distribution. Further we consider a case $\delta < 0.5$. Let us consider two consecutive terms of stationary Markov chain. If $X_n = z$ then X_{n+1} has a distribution from family Ψ with parameter $\theta = \delta(1-z) + z(1-\delta) = \delta + z - 2\delta z$. We see that always $\theta \in (\delta, 1-\delta)$. An inequality $X_{n+1} = x < \theta = \delta + z - 2\delta z$ implies $z > (x-\delta)/(1-2\delta)$. Therefore from (1) we get for $\delta < x < 1 - \delta$

$$g(x) = \int_{(x-\delta)/(1-2\delta)}^{1} g(z) \frac{2x}{\delta + z - 2\delta z} dz + \int_{0}^{(x-\delta)/(1-2\delta)} g(z) 2 \frac{1-x}{1-(\delta + z - 2\delta z)} =$$

$$= 2x \int_{(x-\delta)/(1-2\delta)}^{1} g(z) \frac{1}{\delta + z - 2\delta z} dz + 2(1-x) \int_{0}^{(x-\delta)/(1-2\delta)} g(z) \frac{1}{1-(\delta + z - 2\delta z)} dz. \quad (3)$$

Further

$$\frac{\partial}{\partial x} g(x) = 2 \int_{(x-\delta)/(1-2\delta)}^{1} g(z) \frac{1}{\delta + z - 2z\delta} dz - 2xg\Big(\frac{x-\delta}{1-2\delta}\Big) \frac{1}{1-2\delta} \frac{1}{x} -$$

$$- 2 \int_{0}^{(x-\delta)/(1-2\delta)} g(z) \frac{1}{1-(\delta + z - 2z\delta)} dz + 2(1-x) g\Big(\frac{x-\delta}{1-2\delta}\Big) \frac{1}{1-2\delta} \frac{1}{1-x},$$

$$\frac{\partial^2}{\partial x^2} g(x) = -2g\Big(\frac{x-\delta}{1-2\delta}\Big) \frac{1}{1-2\delta} \frac{1}{x} - 2 \frac{1}{1-2\delta} \frac{\partial}{\partial x} g\Big(\frac{x-\delta}{1-2\delta}\Big) -$$

$$- \frac{2}{1-x} g\Big(\frac{x-\delta}{1-2\delta}\Big) \frac{1}{1-2\delta} + 2 \frac{1}{1-2\delta} \frac{\partial}{\partial x} g\Big(\frac{x-\delta}{1-2\delta}\Big). \quad (4)$$

Finally for $\delta < x < 1 - \delta$

$$x(1-x)\frac{\partial^2}{\partial x^2}g(x) = -2g\Big(\frac{x-\delta}{1-2\delta}\Big)\frac{1}{1-2\delta}. \tag{5}$$

For $0 < x < \delta$ we have from (3)

$$g(x) = 2x\int_0^1 g(z)\frac{1}{\delta + z - 2\delta z}dz. \tag{6}$$

If we denote

$$a = 2\int_0^1 g(z)\frac{1}{\delta + z - 2\delta z}dz \tag{7}$$

then

$$g(x) = ax, \quad 0 < x < \delta. \tag{8}$$

For $x > 1-\delta$ we have analogously

$$g(x) = 2(1-x)\int_0^1 g(z)\frac{1}{1-(\delta + z - 2\delta z)}dz. \tag{9}$$

Obviously $g(x)$ is a symmetric function with respect to $1/2 : g(x) = g(1-x)$, $1/2 < x < 1$. Then

$$2\int_0^1 g(z)\frac{1}{1-(\delta + z - 2\delta z)}dz = 2\int_0^1 g(1-z)\frac{1}{\delta + (1-z) - 2\delta(1-z)}dz = a,$$

$$g(x) = a(1-x), \quad 1-\delta \le x \le 1. \tag{10}$$

3. CALCULATING PROCEDURE

The equation (4) is not convenient for a calculation. It is better to use equation (3). When can be calcalated the last integral in (3) for fixed x using $g(z)$ for $z < x$? Obviously when $(x-\delta)/(1-2\delta) < x$, therefore for $x < 1/2$. To have the same possibility for the other integral, we apply equality (6): if $\delta < x$, then

$$\int_{(x-\delta)/(1-2\delta)}^1 g(z)\frac{1}{\delta + z(1-2\delta)}dz = \frac{a}{2} - \int_0^{(x-\delta)/(1-2\delta)} g(z)\frac{1}{\delta + z(1-2\delta)}dz,$$

$$g(x) = ax - 2x\int_0^{(x-\delta)/(1-2\delta)} \frac{g(z)}{\delta + z(1-2\delta)}dz + 2(1-x)\int_0^{(x-\delta)/(1-2\delta)} \frac{g(z)}{1-\delta - z(1-2\delta)}dz. \tag{11}$$

We can use this formula for consequence calculations. Above, we see that on interval $(0, \delta)$ density $g(x)$ has a known view. Let us consider number sequence

$$b_0 = \delta,\ b_i = \delta + b_{i-1}(1-2\delta) = \delta \sum_{j=0}^{i} (1-2\delta)^j = \frac{1}{2}\big(1 - (1-2\delta)^{i+1}\big), \quad i = 1, 2, \dots\ . \tag{12}$$

Equation (10) shows that values of $g(x)$ on interval (b_i, b_{i+1}) were calculated via $g(x)$'s previous values on $(0, b_i)$. A border value $1/2$ is arrived for $i = \infty$, so we must end calculations when b_i is close to $1/2$.
One should note that a numerical calculation by formula (10) for $x < 1/2$ is very simple. During the recurrent calculation we set $a = 1$. On leaving one, constant a is determined from normalization condition for $g(x) :$ a probabilistic mass on interval $(0, 1/2)$ equals $1/2$. For $1/2 < x < 1$ we use equalities $g(x) = g(1-x)$ and $G(x) = 1 - G(1-x)$.
Let us perform analytical calculations for $x \in (b_0, b_1) = (\delta, 2\delta(1-\delta))$. If $z < \delta$ then $g(z) = az$, so

$$g(x) = ax - 2x \int_0^{(x-\delta)/(1-2\delta)} \frac{az}{\delta + z(1-2\delta)} dz + 2(1-x) \int_0^{(x-\delta)/(1-2\delta)} \frac{az}{1-\delta - z(1-2\delta)} dz =$$

$$= ax - \frac{2ax}{1-2\delta}\Big[\frac{x-\delta}{1-2\delta} - \frac{\delta}{1-2\delta} ln\Big(\frac{x}{\delta}\Big)\Big] - \frac{2a(1-x)}{1-2\delta}\Big[\frac{x-\delta}{1-2\delta} + \frac{1-\delta}{1-2\delta} ln\Big(\frac{1-x}{1-\delta}\Big)\Big] =$$

$$= ax - 2a\frac{x-\delta}{(1-2\delta)^2} + 2ax\frac{\delta}{(1-2\delta)^2} ln\Big(\frac{x}{\delta}\Big) - 2a(1-x)\frac{1-\delta}{(1-2\delta)^2} ln\Big(\frac{1-x}{1-\delta}\Big)$$

Therefore for $x \in (b_0, b_1) = (\delta, 2\delta(1-\delta))$

$$g(x) = ax - \frac{2a}{(1-2\delta)^2}\Big\{x - \delta - x\delta ln\Big(\frac{x}{\delta}\Big) + (1-x)(1-\delta) ln\Big(\frac{1-x}{1-\delta}\Big)\Big\}. \tag{13}$$

Corresponding cumulative distribution function for $x \in (b_0, b_1) = (\delta, 2\delta(1-\delta))$

$$G(x) = \frac{1}{2}a\Big(\frac{x-\delta}{1-2\delta}\Big)^2 + \frac{1}{2}ax^2 + (1-2x)\frac{a}{1-2\delta}\frac{x-\delta}{1-2\delta} + $$
$$+ x^2\frac{a\delta}{(1-2\delta)^2} ln\Big(\frac{x}{\delta}\Big) + (1-x)^2\frac{a(1-\delta)}{(1-2\delta)^2} ln\Big(\frac{1-x}{1-\delta}\Big)\Big\}. \tag{14}$$

An extension of two last formulas on interval $x \in (b_0, 0.5) = (\delta, 0.5)$ will be named *the first-order approximation* $g_1(x)$ and $G_1(x)$ correspondingly. As $G_1(0.5) = 0.5$ we determine normalizing constant a from (13) :

$$a^{-1} = \frac{1}{2} - \frac{1}{2}\frac{\delta}{(1-2\delta)^2}\ln(\delta) - \frac{1-\delta}{2(1-2\delta)^2}\ln(1-\delta) - \frac{1}{2(1-2\delta)^2}\ln(2). \tag{15}$$

Our experience shows that for $\delta > 0.1$ the first-order approximation ensures a sufficient precision of the calculation. The subsequent approximations are determined

analogously. As border value 1/2 from (11) is arrived for $i = \infty$, we must always restrict ourselves by some approximation.
Moments $\mu_r = E(X^r)$ are calculated by a usual way throw density function $g(x)$. Obviously $\mu_1 = 0.5$. The other moments $\{\mu_r\}$ for the first-order approximation $g_1(x)$ are given in Appendix.

4. STATIONARY TWO-DIMENSIONAL DISTRIBUTION

Stationary probability density function $g(x, y) = \frac{\partial^2}{\partial x \partial y} P\{X \leq x, Y \leq y\}$ is expressed via the stationary density $g(x)$ and transition probability density function (1). We must consider case $\theta = x + \delta - 2x\delta > y$ and contrary case separately:

$$g(x, y) = g(x) f_{\delta + x(1-2\delta)}(y) = \begin{cases} g(x)\frac{2y}{\delta + x(1-2\delta)}, & y \leq \delta + x(1 - 2\delta), \\ g(x)\frac{2(1-y)}{1-\delta-x(1-2\delta)}, & otherwise. \end{cases} \quad (16)$$

As it is shown in the Appendix, he second mixed moment is calculated as

$$E(XY) = \frac{1}{6}\Big[1 + \delta + 2(1 - 2\delta)\mu_2\Big]. \quad (17)$$

Now correlation coefficient $\rho = (E(XY) - 0.25)/D(X)$ can be calculated.

5. NUMERICAL RESULTS

In this section we give some numerical comparing. Table 1 shows how different approximations of the density function behave oneself for small value of parameter δ. Three functions are presented here for $\delta = 0.01$: the first-order approximation $g1(x)$, the second-order approximation $g2(x)$, and a curve of function $g0(x)$, that gives precise values for $\delta = 0$ (see [3]). We see, there exists an appreciable difference between these real values and the values received by the first-order approximation. For $\delta = 0.3$ the first-order and the second-order approximations fully coincide, therefore they are precise.

Table 1 **Comparing of different approximations for small parameter** δ

x	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
$g0(x)$	0.285	0.540	0.765	0.960	1.125	1.260	1.365	1.440	1.485
$g2(x)$	0.277	0.530	0.757	0.954	1.122	1.260	1.369	1.488	1.499
$g1(x)$	0.292	0.560	0.797	1.000	1.166	1.293	1.377	1.414	1.402

The Table 2 contains values of stationary density functions for a sum of two adjacent interarrivals for $\delta = 0.005$. Row $fSum(x)$ corresponds to density (15). Row $fInd(x)$ corresponds to independent interarrival times with density $g2(x, 0.01)$. One can see that positive correlation is appreciable.
The correlation coefficient is a linear function of $\delta : \rho(\delta) = -\frac{2}{3}(\delta - 0.5)$, $0 \leq \delta \leq 1$. So,

for $0 \leq \delta \leq 1/2$ we have a positive correlation dependence, for $1/2 \leq \delta \leq 1$ - a negative one.
We see that the considered dependence has some appreciable influence on indices of interest.

Table 2 **Comparing of density functions for a sum**

x	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
$fSum(x)$	0.028	0.108	0.226	0.367	0.515	0.655	0.781	0.882	0.950	0.975
$fInd(x)$	0.006	0.043	0.129	0.264	0.435	0.619	0.808	0.986	1.122	1.178

6. CONCLUSION

We have considered an approach for appreciating dependence between interarrival times. It is supposed that interarrival times correspond to Markov chain with triangular distribution of transition probabilities. Expressions for stationary one-dimensional and two-dimensional distributions of interarrival times and corresponding numerical characteristics have been derived. The following investigations will be devoted to a modification of the used form of a triangular distribution to expand a possible diapason of correlation coefficient values.

7. APPENDIX

Firstly, we derive the expressions for moments $\mu_r = E(X^r)$, $r = 2, 3, ...$, for the first-order approximation $g_1(x)$. Let us introduce the following notations:

$$\Delta(k,c,d) = \int_c^d x^k ln(x)dx = \frac{1}{k+1}d^{k+1}ln(d) - \frac{1}{k+1}c^{k+1}ln(c) - \frac{1}{k+1}\Big(d^{k+1} - c^{k+1}\Big),$$

$$v_r = \int_0^{1/2} x^r g_1(x)dx.$$

Then

$$v_r = \int_0^{1/2} x^r g_1(x)dx = \int_0^{1/2} x^r axdx + \frac{2a}{(1-2\delta)^2}\int_\delta^{1/2} [\delta + (1-\delta)ln(1-\delta)]x^r dx -$$

$$-\frac{2a}{(1-2\delta)^2}\int_\delta^{1/2} [1 + \delta ln(\delta) + (1-\delta)ln(1-\delta)]x^{r+1}dx +$$

$$+\frac{2a\delta}{(1-2\delta)^2}\int_\delta^{1/2} x_{r+1}ln(x)dx - \frac{2a(1-\delta)}{(1-2\delta)}\int_\delta^{1/2} (1-x)ln(1-x)x^r dx =$$

$$= a\frac{1}{r+2}\Big(\frac{1}{2}\Big)^{r+2} + \frac{2a}{(1-2\delta)^2}[\delta + (1-\delta)ln(1-\delta)]\frac{1}{r+1}\Big(\Big(\frac{1}{2}\Big)^{r+1} - \delta^{r+1}\Big) -$$

$$-\frac{2a}{(1-2\delta)^2}[1+\delta ln(\delta)+(1-\delta)ln(1-\delta)]\frac{1}{r+2}\Big(\Big(\frac{1}{2}\Big)^{r+2}-\delta^{r+2}\Big)+$$

$$+\frac{2a\delta}{(1-2\delta)^2}\Delta\Big(r+1,\delta,\frac{1}{2}\Big)+\frac{2a(1-\delta)}{(1-2\delta)^2}\sum_{i=0}^{r}\binom{r}{i}(-1)^{r-i}\Delta\Big(r-i+1,1-\delta,\frac{1}{2}\Big).$$

Further

$$\mu_r=\int_0^1 x^r g_1(x)dx=\int_0^{1/2} x^r g_1(x)dx+\int_{1/2}^1 x^r g_1(1-x)dx=$$

$$=\int_0^{1/2} x^r g_1(x)dx-\int_{1/2}^0 (1-y)^r g_1(y)dy=\int_0^{1/2} x^r g_1(x)dx+\sum_{i=0}^{r}(-1)^i\binom{r}{i}\int_0^{1/2} y^i g_1(y)dy=$$

$$=v_r+\sum_{i=0}^{r}(-1)^i\binom{r}{i}v_i.$$

$$\mu_1=v_0=\frac{1}{2},\quad \mu_2=2v_2-2v_1+\frac{1}{2}.$$

Secondly, we derive formula (16) for a mixed moment:

$$E(XY)=\int_0^1\int_0^{\delta+x(1-2\delta)}\frac{g(x)2y}{\delta+x(1-2\delta)}xydydx+\int_0^1\int_{\delta+x(1-2\delta)}^1\frac{g(x)2(1-y)}{1-[\delta+x(1-2\delta)]}xydydx=$$

$$=\frac{2}{3}\int_0^1 g(x)(\delta+x(1-2\delta))^2xdx+$$

$$+\int_0^1\frac{2x}{1-[\delta+x(1-2\delta)]}g(x)\frac{1}{6}(3y^2-2y^3)dx\Big|_{y=\delta+x(1-2\delta)}^{y=1}=$$

$$=\frac{2}{3}\int_0^1 g(x)(\delta+x(1-2\delta))^2xdx+\frac{1}{3}\int_0^1\frac{x}{1-[\delta+x(1-2\delta)]}g(x)\Big\{1-(\delta+x(1-2\delta))^2-$$

$$-2[\delta+x(1-2\delta)]^2[1-[\delta+x(1-2\delta)]]\Big\}dx=\frac{2}{3}\int_0^1 g(x)(\delta+x(1-2\delta))^2xdx+$$

$$+\frac{1}{3}\int_0^1 x[1+\delta+x(1-2\delta)]g(x)dx-\frac{2}{3}\int_0^1 xg(x)[\delta+x(1-2\delta)]^2dx=$$

$$=\frac{1}{3}[(1+\delta)\mu_1+(1-2\delta)\mu_2]=\frac{1}{3}\Big[(1+\delta)\frac{1}{2}+(1-2\delta)\mu_2\Big].$$

REFERENCES

1. *Abegaz F., Naik-Nimbalkar U. V.* Modelling statistical dependence of Markov chains via copula models // Journal of statistical planning and inference. 2008. № 138(4). P. 1131–1146.
2. *Andronov A.* A triangular symmetric copula // The Third International Conference Problems of Cybernetics and Informatics . 2010. V. II. P. 36–40.
3. *Andronov A* On a copula for failure times of system elements // Mathematical and statistical models and methods in reliability. V.V Rykov et al. (eds.), Springer Science + Business Media, LLC. 2011. P. 39–51.
4. *Embrechts P. , Lindskog F., McNeil A.* Modelling Dependence with Copulas and Applications to Risk Management (Chapter 8) // Handbook of Heavy Tailed Distributions in Finance, Amsterdam, Elsevier. 2003. P. 329–384.
5. *Hernandez-Campos F., Jeffay K. F., Park C., Marron J. S., Resnick S. I.* Extremal dependence: Internet traffic applications // Stochastic Models. 2005. V. 22. № 1. P. 1–35.
6. *KHeyman D. P., Lucantoni D.* Modeling multiple IP traffic streams with rate limits. // IEE/ACM Transactions on Networking (TON). 2003. V. 11. № 6. P. 948–958.
7. *Kim C. S., Klimenok V., Taramin O.* A tandem retrial queueing system with two Markovian flows and reservation of channels // Computers and Operations Research. 2010. V. 37. № 7. P. 1238–1246.
8. *Kim C. S., Klimenok V., Mushko V., Dudin A.* The BMAP/PH/N retrial queueing system operating in Markovian random environment // Computers and Operations Research. 2010. V. 37. № 7. P. 1228–1237.
9. *Kim C. S., Park S. H., Dudin A., Klimenok V., Tsarenkov G.* Investigaton of the BMAP/G/1,/PH/1/M tandem queue with retrials and losses // Applied Mathematical Modelling. 2010. V. 34. № 10. P. 2929–2940.
10. *Kim C. S., Dudin A., Klimenok V., Khramova V.* Erlang loss queueing system with batch arrivals operating in a random environment // Computers and Operations Research. 2009. V. 36. № 3. P. 674–697.
11. *Nelsen B. R.* An Introduction to Copulas. Second Edition New York: Springer. 2006.
12. *Pacheco A., Tang L. C., Prabhu U. N.* Markov-Modulated processes and Semiregenerative Phenomena New Jersey - London - Singapore: World Scientific. 2009.
13. *Scott S. L., Smyth P.* The Markov Modulated Poisson Process and Markov Poisson Cascade with Applications to Web Traffic Modeling // Bayesian Statistics. Oxford University Press. 2003. V. 7. P. 1–10.