Pólya-Aeppli process of order k of second kind with an application

Meglena Lazarova

Faculty of Mathematics and Informatics
Technical University of Sofia
e-mail: meglena.laz@tu-sofia.bg

In collaboration with
Stafanka Chukova and Leda Minkova

1. **Introduction**

Our motivation is based on the risk process,

$$X(t) = c\,t - \sum_{i=1}^{N(t)} Z_i. \tag{1}$$

We assume that the individual claim amount has a continuous distribution with distribution function F, $F(0) = 0$, and mean value $\mu = EZ_1 < \infty$. In the classical risk model the process $N(t)$ is assumed to be a homogeneous Poisson process.

Let us consider the following stochastic process $N(t) = X_1 + \ldots + X_{N_1(t)}$, where $X_1, X_2, \ldots$ are mutually independent random variables and also independent of the process $N_1(t)$.

It is well known that if the compounding random variable X has a discrete distribution, truncated at 0 and from the right away from $k+1$, the random variable $N(t)$ has a distribution of order k, see for example

Aki S., Kuboku H. and Hirano K. (1984). On discrete distributions of order k, *Ann. Inst. Statist. Math.* **36**, Part A, 431–440.

and

Balakrishnan N. and Koutras M.V. (2002). *Runs and Scans with Applications*, Wiley Series in Probability and Statistics.

Pólya-Aeppli distribution of order k was introduced by

Minkova L.D. (2010). Pólya-Aeppli Distribution of Order k, *Commun. Statist.-Theory and Methods,* **39**, 408–415,

and applied as a counting distribution in the risk model considered in

Chukova S. and Minkova L.D. (2015). Pólya-Aeppli of order k Risk Model, *Commun. Statist.-Simulation and Computation,* **44**, 551–564.

There, the random variable $N_1(t)$ is Poisson distributed with parameter λ and X_i are truncated geometrically distributed with probability mass function (PMF) and probability generating function (PGF) given by

$$P(X = i) = \frac{1-\rho}{1-\rho^k}\rho^{i-1}, \quad i = 1, 2, \ldots k \tag{2}$$

and

$$\psi_X(s) = \frac{(1-\rho)s}{1-\rho^k}\frac{1-\rho^k s^k}{1-\rho s},$$

where $k \geq 1$ is fixed integer number. As a result, the above process $N(t)$ is called Pólya-Aeppli process of order k, denoted by $PA_k(\lambda, \rho)$.

In this talk we introduce another Pólya-Aeppli process of order k and call it Pólya-Aeppli process of order k of second kind, and denote it by $PA_{kII}(\lambda, \rho)$. The difference is in the construction of the compounding distribution. In the truncated geometric distribution in (2) the mass from $k+1$ to infinity is uniformly distributed over the points $1, 2, \ldots, k$. Here, we consider the case when the mass from $k+1$ to infinity is clumped at point k.

2. Pólya-Aeppli process of order k of second kind

The distribution of the compounding random variables X_i is given by the following PMF, which clumps the right tail of the distribution at point k:

$$P(X = i) = \begin{cases} (1-\rho)\rho^{i-1}, & i = 1, 2, \dots k-1 \\ \\ \rho^{i-1}, & i = k. \end{cases} \tag{3}$$

The PGF is given by

$$\psi_X(s) = \frac{(1-\rho)s + (1-s)(\rho s)^k}{1-\rho s}. \tag{4}$$

Definition 1 *The distribution defined by (3) and (4) is called a clumped geometric distribution with parameters k and $1-\rho$, and it is denoted by $CGe(k, 1-\rho)$.*

In this case, the PGF of the $N(t)$ is given by

$$\Psi_{N(t)}(s) = e^{-\lambda t\left(1 - \frac{(1-\rho)s + (1-s)(\rho s)^k}{1-\rho s}\right)}. \tag{5}$$

Definition 2 *The process defined by the PGF in (5) is called a Pólya-Aeppli process of order* k *of second kind with parameters* $\lambda > 0$ *and* $\rho \in [0,1)$*. We denote this process by* $PA_{kII}(\lambda, \rho)$*.*

If $k \to \infty$, the clumped geometric distribution approaches the usual geometric distribution with parameter $1 - \rho$.

Remark 1 *If* $k \to \infty$*, the Pólya-Aeppli process of order* k *of second kind, approaches the usual Pólya-Aeppli process, see [5] and [2]. If* $\rho = 0$*, it is the usual homogeneous Poisson process.*

Remark 2 *The mean and the variance functions of the* $PA_{kII}(\lambda, \rho)$ *are given by*

$$EN(t) = \lambda t \frac{1-\rho^k}{1-\rho}$$

and

$$Var(N(t)) = \frac{\lambda t}{(1-\rho)^2}[1 + \rho - (2k+1)\rho^k + (2k-1)\rho^{k+1}].$$

For the Fisher index, we obtain

$$FI(N(t)) = \frac{Var(N(t))}{E(N(t))} = \frac{1+\rho}{1-\rho} - 2k\frac{\rho^k}{1-\rho^k}.$$

The Fisher index of the distribution of the Pólya-Aeppli process is equal to $\frac{1+\rho}{1-\rho}$*, see [2]. Hence, the distribution of the counting process* $PA_{kII}(\lambda, \rho)$ *is underdispersed with respect to the distribution of the Pólya-Aeppli process.*

Let us denote by $P_n(t) = P(N(t) = n),\ n = 0, 1, \ldots$ The following proposition gives an extension of the Panjer recursion formulas, see [6].

Proposition 1 *The PMF of the $N(t) \sim PA_{kII}(\lambda, \rho)$ satisfies the following recursion formulae:*

$$P_1(t) = \lambda t(1-\rho)P_0(t),$$

$$P_n(t) = (2\rho + \tfrac{\lambda t(1-\rho)-2\rho}{n})P_{n-1}(t) - (1-\tfrac{2}{n})\rho^2 P_{n-2}(t), \quad n = 2, 3, \ldots k-1$$

$$P_n(t) = (2\rho + \tfrac{\lambda t(1-\rho)-2\rho}{n})P_{n-1}(t) - (1-\tfrac{2}{n})\rho^2 P_{n-2}(t) + \lambda t\rho^k \tfrac{k}{n} P_{n-k}(t)$$

$$-\lambda t\rho^k[\tfrac{k+1}{n} + \tfrac{k-1}{n}\rho]P_{n-k-1}(t) + \lambda t\rho^{k+1}\tfrac{k}{n}P_{n-k-2}(t), \quad n = k, k+1, k+$$

and $P_{-1}(t) = P_{-2}(t) = 0.$

3. Pólya-Aeppli process of order k of second kind as a birth process

Suppose that $N(t) \sim PA_{kII}(\lambda, \rho)$. The properties of this process are specified by the following assumptions: For any small $h > 0$

$$P(N(t+h) = n \mid N(t) = m) = \begin{cases} 1 - \lambda h + o(h), & n = m, \\ (1-\rho)\rho^{i-1}\lambda h + o(h), & n = m + i, \\ & i = 1, 2, \dots, k-1, \\ \rho^{k-1}\lambda h + o(h), & n = m + k, \end{cases} \tag{6}$$

for every $m = 0, 1, \dots,$ where $o(h) \to 0$ as $h \to 0$. Note that the assumptions imply that for $i = k+1, k+2, \dots,$ $P(N(t+h) = m + i \mid N(t) = m) = o(h)$.

The above assumptions yield the following Kolmogorov forward equations:

$$P_0'(t) = -\lambda P_0(t),$$

$$P_n'(t) = -\lambda P_n(t) + (1-\rho)\lambda \textstyle\sum_{j=1}^{n} \rho^{j-1} P_{n-j}(t), \quad n = 1, 2, \ldots, k-1,$$

$$P_n'(t) = -\lambda P_n(t) + (1-\rho)\lambda \textstyle\sum_{j=1}^{k-1} \rho^{j-1} P_{n-j}(t) + \lambda \rho^{k-1} P_{n-k}(t), \quad n = k, k+ \tag{7}$$

with the conditions

$$P_0(0) = 1 \quad \text{and} \quad P_n(0) = 0, \quad n = 1, 2, \ldots. \tag{8}$$

Multiplying the nth equation of (7) by s^n and summing for all $n = 0, 1, 2, \ldots$ we get the following differential equation

$$\frac{\partial \Psi_{N(t)}(s)}{\partial t} = -\lambda[1 - \psi_X(s)]\Psi_{N(t)}(s). \tag{9}$$

The solution of (9) with the initial condition

$$\Psi_{N(1)}(s) = 1$$

is given by (5), which is the PGF of the distribution of $PA_{kII}(\lambda, \rho)$. This leads to the following, equivalent to Definition 2, definition for the Pólya-Aeppli process of order k of second kind, namely:

Definition 3 *The process defined by (7) and (8) is the Pólya-Aeppli process of order k of second kind.*

4. Application to risk model

We consider the risk model (1), where $N(t) \sim PA_{kII}(\lambda, \rho)$. We call this model a Pólya-Aeppli of order k of second kind risk model. In this case the relative safety loading θ is defined by

$$\theta = \frac{EX(t)}{E\sum_{i=1}^{N(t)} Z_i} = \frac{c(1-\rho)}{\lambda\mu(1-\rho^k)} - 1.$$

To ensure that $\theta > 0$, the premium income per unit time c should satisfy the following inequality

$$c > \frac{\lambda\mu(1-\rho^k)}{1-\rho}.$$

Denote by $\tau = \inf\{t : X(t) < -u\}$ the time to ruin of an insurance company having initial capital $u \geq 0$, and by

$$\Psi(u) = P(\tau < \infty) \tag{10}$$

the related ruin probability.

Let $G(u, y)$ be the probability of the following event:
$\{$ruin occurs with initial capital u and deficit, immediately after ruin occurs, is
with $u \geq 0$ and $y \geq 0$. Hence

$$G(u, y) = P(\tau < \infty, D \leq y), \tag{11}$$

where $D = |u + X(\tau)|$ is the deficit immediately after ruin occurs. Therefore

$$\lim_{y \longrightarrow \infty} G(u, y) = \Psi(u). \tag{12}$$

Let us denote by

$$H(x) = (1 - \rho) \sum_{i=1}^{k-1} \rho^{i-1} F^{*i}(x) + \rho^{k-1} F^{*k}(x) \tag{13}$$

the non defective probability distribution function of the claims with

$$H(0) = 0, \qquad H(\infty) = 1.$$

Then, using the assumptions in (6), we obtain the following differential equation

$$\frac{\partial G(u,y)}{\partial u} = \frac{\lambda}{c}\left[G(u,y) - \int_0^u G(u-x,y)dH(x) - [H(u+y) - H(u)]\right]. \tag{14}$$

4.1. Ruin probability

Theorem 1 *The probability of ruin* $\Psi(u)$ *satisfies the equation*

$$\frac{d\Psi(u)}{du} = \frac{\lambda}{c}\left[\Psi(u) - \int_0^u \Psi(u-x)dH(x) - [1 - H(u)]\right], \quad u \geq 0. \tag{15}$$

Integrating in (15), we obtain the function $G(0, y)$ given by

$$G(0, y) = \frac{\lambda}{c}\int_0^y [1 - H(u)]du, \tag{16}$$

and for the ruin probability with no initial capital we obtain

$$\Psi(0) = \frac{\lambda\mu}{(1-\rho)c}(1-\rho^k). \tag{17}$$

4.2. Exponentially distributed claims

Let us consider the case of exponentially distributed claim sizes with mean μ, i.e. $F(x) = 1 - e^{-\frac{x}{\mu}}, \quad x \geq 0, \quad \mu > 0$. In this case, the function

$$F^{*i}(x) = 1 - \sum_{j=0}^{i-1} \frac{\left(\frac{x}{\mu}\right)^j}{j!} e^{-\frac{x}{\mu}},\ x \geq 0$$

is an Erlang distribution function. Then, the distribution function $H(x)$ in (13) is given by

$$H(x) = 1 - \sum_{i=0}^{k-1} \frac{\left(\frac{\rho x}{\mu}\right)^i}{i!} e^{-\frac{x}{\mu}}.$$

The density function $h(x)$ has the form

$$h(x) = \frac{1}{\mu} \left[(1-\rho) \sum_{i=0}^{k-2} \frac{\left(\frac{\rho x}{\mu}\right)^i}{i!} + \frac{\left(\frac{\rho x}{\mu}\right)^{k-1}}{(k-1)!} \right] e^{-\frac{x}{\mu}}.$$

So, the initial condition (16) in the case of exponential distribution is

$$G(0, y) = \frac{\lambda\mu}{c} \sum_{i=0}^{k-1} \frac{\rho^i}{i!} \gamma(i + 1, y/\mu),$$

where $\gamma(\alpha, x) = \int_0^x t^{\alpha-1} e^{-t} dt$ is the incomplete Gamma function.

5. Simulation

In what follows, we apply the simulation approach for calculating the probability of ruin suggested in [3] for the case of exponentially distributed claims with initial capital $u = 0$. We confirm the validity of our simulation results by matching them with the value of the ruin probability computed analytically using (17). Then, using our simulator, we provide results for the case of non-zero initial capital not only for exponentially distributed claims but also for claims with gamma and Weibull distributions.

Next, we provide some results regarding the probability of ruin for different scenarious of the claim distribution as well as the value of the initial capital.

5.1. Results

We consider the case of exponentially distributed claims and no initial capital $u = 0$. We verify the correctness of our simulator by comparing the results for the probability of ruin for fixed model parameters, produced in two different ways : (i) by the simulator, given in column “simulated“, and (ii) computed using (17) given in column “analytical“. These are given in Table 1.

λ	k	ρ	simulated Exp(1)	analytical Exp(1)
1.0	15	0.6	0.208117	0.208235
1.5	4	0.8	0.316531	0.316286
2.0	10	0.4	0.256365	0.256383
2.5	3	0.9	0.423526	0.423437
3.0	6	0.2	0.288426	0.288443

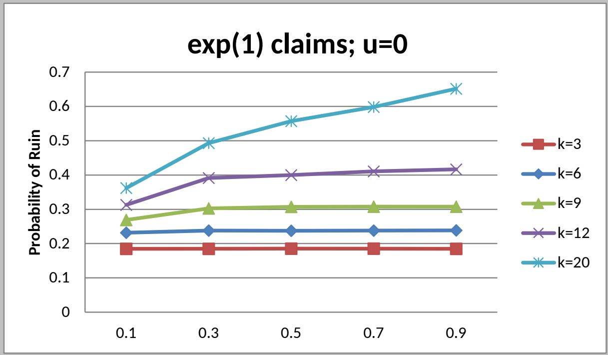
Table 1: Simulated and analytical Exp(1)

As it is easy to see, the “analytical“ and “simulated“ results are very close. So, we use our simulator, written in MATHEMATICA, to compute a reasonable approximation of the probability of ruin for non-exponentially distributed claims and non-zero initial capital $(u \neq 0)$ and a summary of our results is given in subsection .

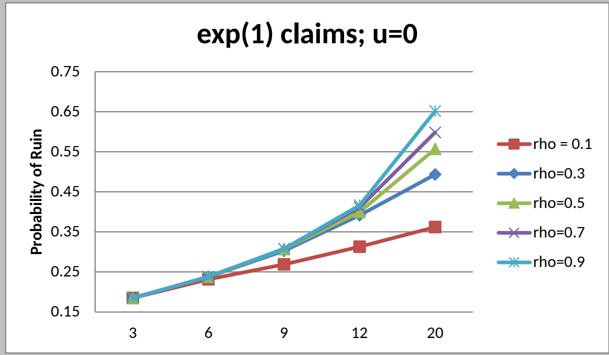
5.1.1. Case 1: Exponentially distributed claims

Here, we present some simulation results for the case of exponentially distributed claims with non-zero initial capital.

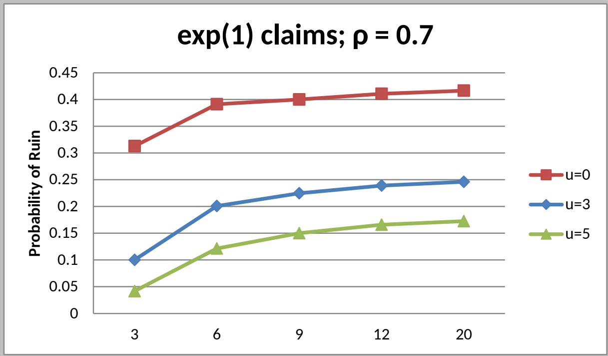
Comparing part(b) and part(d) of Figure 1, it is easy to see that the probability of ruin is shifted downwards as the initial capital increases. If the initial capital is $u = 0$, the smallest values for the probability of ruin is just above 0.35 for $\rho = 0.1$, whereas the analogous value for $u = 5$ is just below 0.1. The depicted overall dependence on ρ, regardless of the value of the initial capital, is as expected, the probability of ruin increases as ρ increases. The overall trends depicted in part(a) and part(c) of Figure 1 also agree with our intuition. Namely, for a fixed value of ρ, the probability of ruin is higher for low values of the initial capital and it increases on k. It is worth to point out the sharp increase of the probability of ruin for large values of ρ and large k, as shown in part(a) of Figure 1.



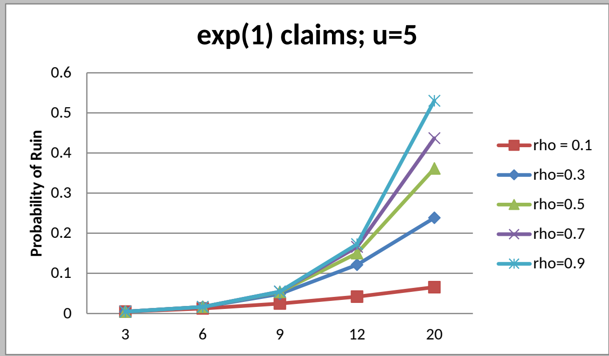
(a) $u = 0$, dependence on k and ρ



(b) $u = 0$, dependence on ρ and k



(c) $\rho = 0.7$, dependence on u and k



(d) $u = 5$, dependence on ρ and k

Figure 1: Probability of ruin: exponentially distributed claims

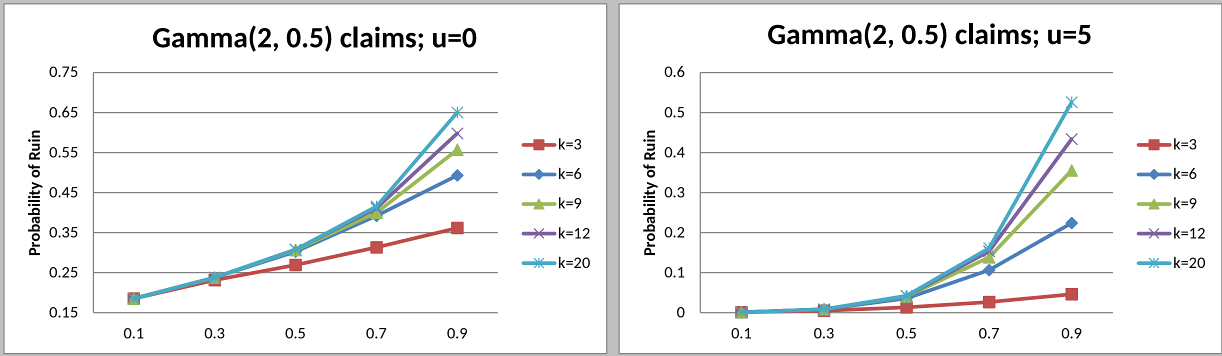
5.1.2. Case 2: Gamma distributed claims

Next, we consider gamma distributed claims with parameters α and β, i.e., the density function of the claim sizes is

$$f(x) = \frac{x^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)} e^{-\frac{x}{\beta}}, \quad x \geq 0,$$

where $\Gamma(\alpha)$ is the Gamma function. Suppose that $\alpha = 2$ and $\beta = 0.5$. In this case the mean values of the claims are $EZ_i = \alpha\beta = 1$. We present results for different values of the model parameters u, k and ρ.

The trends observed for the gamma distributed claims are similar to the one we have presented and discussed for the case of exponentially distributed claims in subsection . Here, in Figure 2, we depict the dependence of the probability of ruin from u, for similar ρ and k. Overall the probability of ruin for lower value of the capital u is higher, similar to what we have observed in the exponential case. In addition we see that for high values of u, and ρ, k have a strong impact on the probability of ruin, e.g., see for $\rho = 0.9$, and $u = 0$ the range of the probability of ruin is approximately $(0.35, 0.65)$, whereas for $u = 5$ this range is much larger,



(a) $u = 0$, dependence on k and ρ

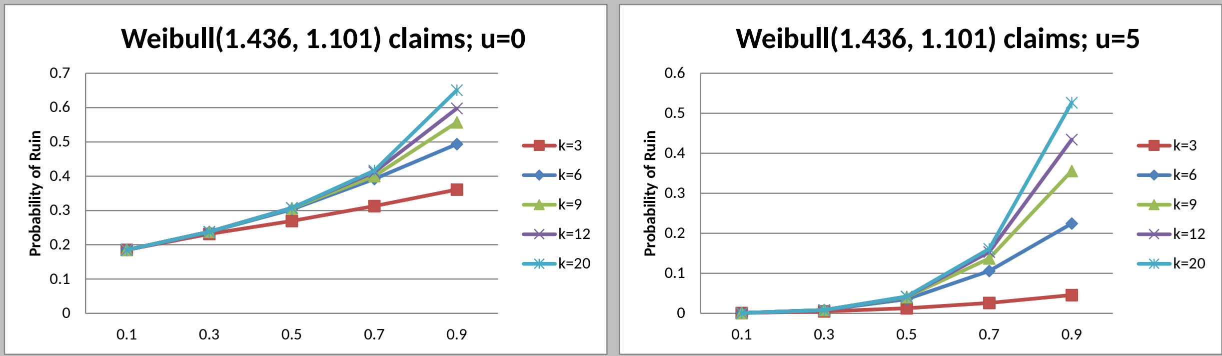
(b) $u = 5$, dependence on k and ρ

Figure 2: Probability of ruin: gamma distributed claims

approximately $(0.05, 0.53)$.

5.1.3. Case 3: Weibull distributed claims

Next, we consider the Weibull distribution with parameters $\alpha = 1.43552259$ and $\beta = 1.1013206$ distributed claims. Here α is the shape parameter and β is the scale parameter. The parameters of the Weibull and gamma distributions were selected so that the three claim size distributions considered in sections , and have the same expectation $\mu = 1$ and the Weibull and gamma claim sizes have the same variances.



(a) $u = 0$, dependence on k and ρ

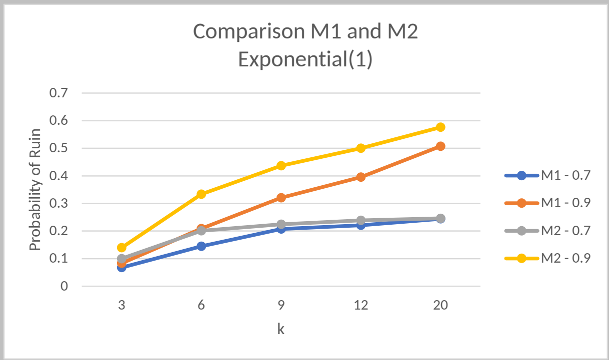
(b) $u = 5$, dependence on k and ρ

Figure 3: Probability of ruin: Weibull distributed claims

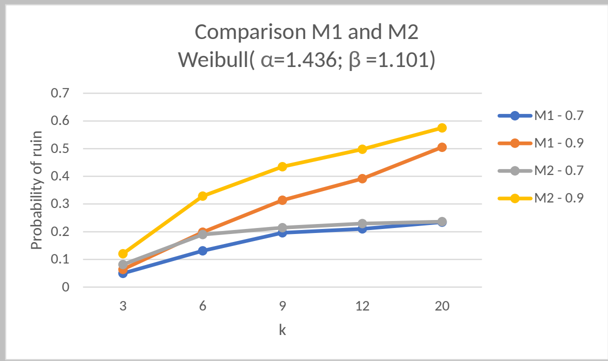
We were quite surprised to see that the behavior of the probability of ruin under Weibull distributed claims, part(a) and part(b) in Figure 3, mimics quite closely the behavior of this probability for gamma distributed claims. So, then the natural question is: under a risk model based on the Pólya-Aeppli process of order k, are the mean value and the variance of the claim distribution what determines the probability of ruin, i.e., the actual form of the claim size distribution does not have an effect on the probability of ruin. Interestingly, similar observations were made in [**?**]. Again, observing these results is a good motivation for future research because at this point we are not able to answer this question.

5.2. Comparison between M1 and M2

For brevity we will refer to the current model as M2 and to $PA_k(\lambda, \rho)$ as M1. Here we provide a brief comparison between the probabilities of ruin for the two models.



(a) $\rho = 0.7, 0.9,\ u = 3$



(b) $\rho = 0.7, 0.9,\ u = 3$

Figure 4: Probability of ruin: comparison between M1 and M2

In part(a) and part(b) in Figure 4, we fix the value of the parameter $u = 3$, and illustrate the dependence of the probability of ruin for M1 and M2 for two different values of $\rho = 0.7, 0.9$. Again, the probability of ruin for M1 and M2 is similar for the selected exponential and Weibull claim size distributions. The probability of ruin is an increasing function of k and its value is shifted upwards for higher values of parameter ρ. As expected, the probability of ruin for M2 is higher than for M1 and this is exactly what we expect to observe as an outcome for the insurance company at the time of severe natural disaster. Having model $PA_{kII}(\lambda, \rho)$ in place provides a reasonable theoretical background for the company to plan accordingly for natural calamities. From the observations above a natural question arises: are there any condition on the mean and the variance of the claim size distribution that will guaranty the satisfaction of some inequalities on the related ruin probabilities. These inequalities will be very useful in the sense that, even at the time of calamity, the probability of ruin would not exceed a known value. Again, further numerical and theoretical studies are needed to gain some insight on this question.

References

[1] Charalambides Ch. (1986). On discrete distributions of order k, *Ann. Inst. Statist. Math.*, **38**, Part A, 557–568.

[2] Chukova S. and Minkova L.D. (2013). Characterization of the Pólya-Aeppli process, *Stochastic Analysis and Applications*, **31**, 590–599.

[3] Dufresne F. and Gerber H. U.(1989). Three methods to calculate the probability of ruin, *ASTIN Bull.*, **19(1)**, 71–90.

[4] Hirano K. (1986). Some properties of the distributions of order k, in *Fibonacci Numbers and Their Applications, A. N. Philippou et al.(eds)*, 43–53.

[5] Minkova L.D. (2004). The Pólya-Aeppli process and ruin problems, *J. Appl. Math. Stoch. Analysis*, **3**, 221–234.

[6] Panjer H. (1981). Recursive evaluation of a family of compound discrete distributions, *ASTIN Bull.*, **12**, 22–26.