**American Journal of Theoretical and Applied Statistics**
2013; 2(5): 110-121
Published online August 30, 2013 (http://www.sciencepublishinggroup.com/j/ajtas)
doi: 10.11648/j.ajtas.20130205.11

# On nonnegative integer-valued Lévy processes and applications in probabilistic number theory and inventory policies

**Huiming Zhang*, Jiao He, Hanlin Huang**

Dept. of Mathematics and Statistics, Central China Normal University, Wuhan, China

**Email address:**
at352693585@gmail.com(Zhang H.)

**To cite this article:**
Huiming Zhang, Jiao He, Hanlin Huang. On Nonnegative Integer-Valued Lévy Processes and Applications in Probabilistic Number Theory and Inventory Policies. *American Journal of Theoretical and Applied Statistics.* Vol. 2, No. 5, 2013, pp. 110-121.
doi: 10.11648/j.ajtas.20130205.11

**Abstract:** Discrete compound Poisson processes (namely nonnegative integer-valued Lévy processes) have the property that more than one event occurs in a small enough time interval. These stochastic processes produce the discrete compound Poisson distributions. In this article, we introduce ten approaches to prove the probability mass function of discrete compound Poisson distributions, and we obtain seven approaches to prove the probability mass function of Poisson distributions. Finally, we discuss the connection between additive functions in probabilistic number theory and discrete compound Poisson distributions and give a numerical example. Stuttering Poisson distributions (a special case of discrete compound Poisson distributions) are applied to numerical solution of optimal (s, S) inventory policies by using continuous approximation method.

**Keywords:** Probability Mass Function, Nonnegative Integer-Valued Lévy Processes, Probabilistic Number Theory, Discrete Compound Poisson Distribution, (S, S) Inventory Policies

## 1. Introduction

Poisson distribution is a famous distribution in discrete distribution family, and it has important applications in social and economic sciences, physics, biology and other fields. For example, the number of passengers came to a bus stop, the number of particles emitted by radioactive substances, the number of microorganisms in a region under the microscope and so on. Poisson (1837) [25] tried to use the binomial distribution of several experiments to derive distribution function of the Poisson distribution in his representative work *Recherches sur la probabilité des jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabilités*

$$P = (1+\omega+\frac{\omega^2}{1\cdot 2}+\frac{\omega^3}{1\cdot 2\cdot 3}+\cdots+\frac{\omega^n}{1\cdot 2\cdot 3\cdot\cdots\cdot n})e^{-\omega}$$

After about 100 years, de Finetti (1929) [6] initiated that the probability of an event in the interval $[t_0, \Delta t + t_0)$ of Poisson Processes is $\lambda\Delta t + o(\Delta t), (\lambda > 0)$ and the probability more than one events is $o(\Delta t)$. Khintchine [21] summarized the equivalent conditions of Poisson distribution: zero initial conditions, stationary increments, independent increments, and orderliness (impossible of two or more events occurring in the same moment of time). In actual life, Poisson processes (distribution) is not an adequate model for the observed data that the possibility of two or more events occur at a given instant. Khintchine [21] also generalized the Poisson processes by giving more restrict of orderliness. As for a situation: during period of $t$, the number of cars that arrive at the terminal station can be regard to Poisson distribution, every car take $i$ passengers with probability $\alpha_i$, thus the number of passengers who arrive at the terminal station can not satisfy Poisson processes. So we assume that there are many events occurring in a small segment of length $t$. Hight [14] considered another situation: somebody might throw some letters (more than one) into a postbox at the same time.

## 2. Discrete Compound Poisson Model

### *2.1. Nonnegative Integer-Valued Lévy Processes*

We employ the following definition due to [14].

Definition 1. (Discrete compound Poisson processes)

Nonnegative integer-valued stochastic processes $\{X(t)\}_{t\geq 0}$ satisfy the following four conditions:

(i) Initial condition: $X(0)=0$ ;

(ii) Stationary increments: the events occur in $[t_0, t+t_0)$ only depends on $t$, and is not relevant to $t_0$ ;

(iii) Independent increments: the events occur in $[t_0, t+t_0)$ is independent with the events which happen before $t_0$ ;

(iv) Superimposition: the probability of $i$ events taking place between $[t_0, t+t_0)$ is:

$$P_i(\Delta t)=\lambda\alpha_i\Delta t+o(\Delta t),(i\geq 1,\lambda>0,\alpha_i\geq 0,\sum_{i=1}^{\infty}\alpha_i=1)$$

Especially, when it applies to the inventory management theory, the number of consumers coming in a period of $t$ can be seen to subject to the stuttering Poisson distribution (SPD). When the interval is small enough and the number has the property of geometric distribution, that is

$$a_i=(1-\alpha)\alpha^i,(i=1,2,\cdots)$$

Galliher [12] applied SPD to the inventory management firstly, and named it as stuttering Poisson distribution. We will prove the explicit expression of probability mass function (pmf) of SPD in the following part. Definition 1 was put forward by Khintchine [21], where $\alpha_i$ be probability of nonnegative discrete distribution. Hight [14] still named stuttering Poisson distribution instead of discrete compound Poisson distribution existing in the broad sense of inventory management. There are some other names:

Pollaczek-Geiringer distribution,
Generalized Poisson distribution,
Composed Poisson distribution,
Poisson power series distribution,
Poisson par grappes distribution,
Poisson-stopped sum distribution,
Multiple Poisson distribution,
Infinite divisible distribution on the nonnegative integer.

Readers can find more general information of Poisson distribution and discrete compound Poisson distribution in [18] and [14]. The wide range of applications with discrete compound Poisson, see [21], [16], [34], [12] and [2].

Definition 2. (Discrete compound Poisson distribution) In particular, we say the discrete random variable $X$ satisfying the conditions above has a discrete compound Poisson (DCP) distribution with parameters

$$(\alpha_1\lambda,\alpha_2\lambda,\cdots)\in \mathrm{R}^{\infty}(\textstyle\sum_{i=1}^{\infty}\alpha_i=1,\alpha_i\geq 0,\lambda>0)$$

We denote it as

$$X\sim CP(\alpha_1\lambda,\alpha_2\lambda,\cdots)$$

If $X\sim CP(\alpha_1\lambda,\cdots,\alpha_r\lambda)$ , we say $X$ has a DCP distribution of order $r$.

Remark 1. In the Section 3.8, we will show $X(t)$ have compound Poisson sum properties.

Theorem 1. Discrete compound Poisson process $X(t)$ is a Markov chain with stationary independent increments. Let $P_i(t)=P\{X(t)=i|X(0)=0\}$ . The pmf of DCP processes is

$$P_n(t)=\sum_{\Sigma_{i=1}^{n}ik_i=n,k_i\in \mathrm{N}}\frac{\alpha_1^{k_1}\alpha_2^{k_2}\cdots\alpha_n^{k_n}}{k_1!k_2!\cdots k_n!}(\lambda t)^{k_1+k_2+\cdots+k_n}e^{-\lambda t} \quad (1)$$

where $\mathrm{N}=\{0,1,2,\cdots\}$.

By the definition of Bell polynomial [4]:

$$\exp[\sum_{i=1}^{\infty}\frac{a_i}{i!}x^i]=\sum_{i=0}^{\infty}\frac{B_i(a_1,\cdots,a_i)}{i!}x^i$$

hence $P_i(t)$ satisfies

$$B_i(1!\alpha_1\lambda t,\cdots,i!\alpha_i\lambda t)=i!P_i(t).$$

So $P_n(t)$ can be expressed by Bell polynomial. The indeterminate equation $\Sigma_{i=1}^{n}ik_i=n,(k_i\in\mathrm{N})$ is called Diophant's equations. The depth theoretical properties and statistical applications of Bell polynomial with Diophant's equations $\Sigma_{i=1}^{n}ik_i=n$, see [30].

Theorem 2. If $X(t)\sim CP(\alpha_1\lambda t,\alpha_2\lambda t,\cdots)$ , then the probability generating function (pgf) of $X(t)$ is

$$P(s)=e^{\lambda t\sum_{i=1}^{\infty}\alpha_i(s^i-1)},(|s|\leq 1). \quad (2)$$

Remark 2. We will give ten approaches to proof the pmf of DCP distribution and two to approaches proof the pgf of DCP distribution in Section 3.

Lévy processes $\{X(t)\}_{t\geq 0}$ satisfy the following four conditions as follows:

(i) $P\{X(0)=0\}=1,a.s.$ ;

(ii) $X(t)$ has independent increment;

(iii) $X(t)$ has stationary increment;

(iv) $X(t)$ is almost surely right continuous with left limits.

Lévy processes can derive Lévy–Khinchine formula (details can be seen in Bertoin(1996)). The characteristic function $\mathrm{E}[e^{i\theta X(t)}]$ of $X(t)$ is

$$\exp\left(ait\theta-\frac{1}{2}\sigma^2t\theta^2+t\int_{R\setminus\{0\}}(e^{i\theta x}-1-i\theta x\mathrm{I}_{|X|<1})w(dx)\right),$$

where $a\in R,\sigma\geq 0$ , $\mathrm{I}_A$ is an indicator function. And $w(dx)$ is a measure which is called Lévy measure. It satisfies

$$\int_{R\setminus\{0\}}\min\{x^2,1\}w(dx)<\infty$$

Lévy–Khinchine formula is firstly obtained in the generalization of de Finetti [6] and Kolmogorov [22] results by Lévy [23] and Khinchine [20]. This formula can be expressed as the union of Brown motion, constant drift, compound Poisson process, and a pure jump martingale. By the definition of Lévy process, suppose

$$w(dx)=\alpha_x\lambda d\delta_x, \qquad (3)$$

where $\delta_x$ is Dirac measures (or point measure) with $\int fd\delta_x=f(x),(x\in \mathrm{N})$ and $\int_{R\setminus\{0\}}\min\{x^2,1)w(dx)<\infty$.

Because $X(t)$ is a integer value as well as the definition of $w(dx)$, we have

$$ait\theta-\frac{1}{2}\sigma^2 t\theta^2=t\int_{R\setminus\{0\}}(i\theta x \mathrm{I}_{|X|<1})w(dx)=0$$

Hence, the characteristic function of $X(t)$ is

$$\mathrm{E}[e^{i\theta X(t)}]=\exp[\lambda t\sum_{j=1}^{\infty}\alpha_j(e^{ji\theta}-1)],$$

Theorem 2 shows that the Lévy measure (3) we defined before is reasonable.

Discrete compound Poisson process is a continuous-time non-negative integer value Lévy process. Barndorff [2] put forward the integer-valued Lévy process, and he use it in latency financial econometrics.

### 2.2. Special Case

#### 2.2.1. Hermite Distribution

When $r=2$ in definition 2, it called Hermite distribution. The pmf is

$$P_n(t)=\sum_{i=0}^{[\frac{n}{2}]}\frac{\alpha_1^{n-2i}\alpha_2^i(\lambda t)^{n-i}}{(n-2i)!i!}e^{-\lambda t}.$$

The pmf above is gotten by the pgf of Hermite distribution expanded in terms of Hermite polynomial. More details can be seen in Kemp [19].

#### 2.1.2. Stuttering Poisson Distribution

In inventory system, Galliher [12] use the DCP distribution with parameter $X(t)\sim CP(\alpha_1\lambda t,\alpha_2\lambda t,\cdots)$ to describe the demands, where $\alpha_n\lambda=(1-\alpha)\alpha^{n-1}\lambda$. He called it stuttering Poisson distribution.

The following parts show that the pmf of SPD can be expressed by Laguerre polynomial. We need a lemma. The power series solution of second-order linear differential equation

$$xy''+(1-x)y'+ny=0$$

is $y=\frac{e^x}{n!}\frac{d^n(x^ne^{-x})}{dx^n}\triangleq L_n(x)$, which is called Laguerre polynomial.

We begin by proving a property of Laguerre polynomial. First, we need Lemma 1. It is the connection between Laguerre polynomial and pgf of SPD.

Lemma 1: The Taylor's formula with respect to $t$ of $P(t,x)=\frac{1}{1-t}e^{-\frac{xt}{1-t}}$ is $\sum_{n=0}^{\infty}L_n(x)t^n$ in $|t|\le r<1$.

Proof: It is easy to prove that $\frac{e^x}{n!}\cdot\frac{d^n(x^ne^{-x})}{dx^n}$ satisfies ODE: $xy''+(1-x)y'+ny=0$. Suppose $t$ is a complex variable and assume that $P(t,x)=(1-t)^{-1}e^{-(1-t)^{-1}xt}$ is analytic in $|t|\le r<1$. Define the Taylor expansion of $P(t,x)$ in $|t|\le r<1$ is $\sum_{n=0}^{\infty}a_n(x)t^n$. By the Cauchy formula of higher order derivative of analytic function, the coefficient $a_n(x)$ of power series is

$$\frac{1}{2\pi i}\int_{|\xi|=r}\frac{1}{\xi^{n+1}}\cdot\frac{1}{(1-\xi)}e^{-\frac{x\xi}{1-\xi}}d\xi=\frac{e^x}{2\pi i}\int_{\left|1-\frac{x}{z}\right|=r}\frac{z^ne^{-z}}{(z-x)^{n+1}}dz$$
$$=\frac{e^x}{n!}\cdot\frac{d^n(x^ne^{-x})}{dx^n}$$

Theorem 3：The pmf of SPD is

$$P_n(t)=\alpha^n[L_n(\frac{\alpha-1}{\alpha}\lambda t)-L_{n-1}(\frac{\alpha-1}{\alpha}\lambda t)]e^{-\lambda t}.$$

Proof: The pgf of SPD can be rewritten as follow by the Lemma 1 above:

$$P(s,t)=e^{\sum_{i=1}^{\infty}(1-\alpha)\alpha^{i-1}\lambda t(s^i-1)}=e^{-\lambda t}e^{\frac{1-\alpha}{\alpha}\lambda t\cdot\frac{\alpha s}{1-\alpha s}}$$
$$=e^{-\lambda t}\sum_{n=0}^{\infty}L_n(-\frac{1-\alpha}{\alpha}\lambda t)(as)^n(1-as)$$
$$=e^{-\lambda t}[\sum_{n=0}^{\infty}L_n\left(-\frac{1-\alpha}{\alpha}\lambda t\right)(as)^n-\sum_{n=0}^{\infty}L_n\left(-\frac{1-\alpha}{\alpha}\lambda t\right)(as)^{n+1}]$$

$P_n(t)$ is the coefficient of $s^n$ in the expanded formula of $P(s,t)$, hence

$$P_n(t)=\left.\frac{\partial^n P(s,t)}{\partial t^n}\right|_{s=0}=\alpha^n[L_n(\frac{\alpha-1}{\alpha}\lambda t)-L_{n-1}(\frac{\alpha-1}{\alpha}\lambda t)]e^{-\lambda t}.$$

Let $\alpha=0.2、0.5,\lambda t=10$, plot the graphic in SPD and PD of $y=P_n(t)$ by Maple 16.

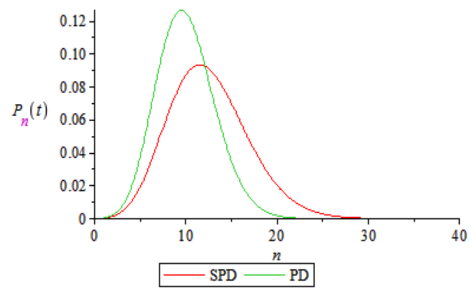



***Figure 1.*** $\alpha=0.2$

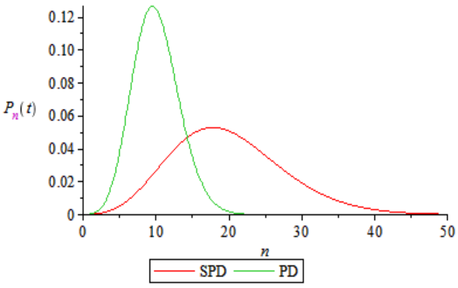



***Figure 1.*** $\alpha = 0.5$

The figure of SPD is more dwarfed than it is in PD from Fig.1 and Fig. 2 This is because the SPD have superimposed events occurred within a sufficiently small time interval, while the PD is not allowed. Number of incidents in a unit time is relatively more in SPD. More discrete compound Poisson distribution examples can be seen in Wimmer [32], it lists more than 100 special cases.

## 3. Ten Approaches to Prove the PMF of Discrete Compound Poisson Distribution

### *3.1. Lemma*

Lemma 2. (Cauchy functional equation) Suppose $f$ is continuous on R , so $f(x+y)=f(x)+f(y)$ for all $x,y\in R$ , Then $f(x)=ax(a\in \mathrm{R})$ . The proposition of Cauchy functional equation: Suppose $f$ is continuous on R , so $f(x+y)=f(x)f(y)$ for all $x,y\in R$ , Then $f(x)=e^{ax}(a\in \mathrm{R})$ .

Lemma 3. (Euler's method of linear differential equation with constant coefficients) Suppose a $p$-dimensional column vector function $\boldsymbol{x}$ , $\boldsymbol{P}$ is an $n$-order coefficient matrix of differential equation $\frac{d\boldsymbol{x}}{dt}=\boldsymbol{Px}$ .

Let $\boldsymbol{F}(\lambda)=\boldsymbol{P}-\lambda\boldsymbol{E}$ , and $\det \boldsymbol{F}(\lambda)=0$ is the characteristic equation of $\boldsymbol{P}$ . $\det \boldsymbol{F}(\lambda)=0$ has different characteristic roots in the complex field. Suppose they are $\lambda_1,\lambda_2,\cdots,\lambda_m (m\le n)$ which have algebraic multiplicities

$$r_1,r_2,\cdots,r_m \ (r_1+r_2+\cdots+r_m=n) .$$

Then the particular solution of the equation is

$$\boldsymbol{x}_i=[\boldsymbol{E}+\boldsymbol{F}(\lambda_i)t+\frac{1}{2!}\boldsymbol{F}(\lambda_i)t+\cdots + \frac{1}{(r_i-1)!}\boldsymbol{F}^{r_i-1}(\lambda_i)t^{r_i-1}]\boldsymbol{A}_i e^{\lambda_i t},(1\le i\le m) ,$$

where the vector $\boldsymbol{A}_i$ satisfies equation $\boldsymbol{F}^{r_i}(\lambda_i)\boldsymbol{A}_i=0$ .

Hence, the general solution of equation $\frac{d\boldsymbol{x}}{dt}=\boldsymbol{Px}$ is

$$\boldsymbol{x}=c_1\boldsymbol{x}_1+c_2\boldsymbol{x}_2+\cdots+c_m\boldsymbol{x}_m, \ (c_1,c_2,\cdots,c_m \text{ are constants }) \quad (4)$$

Lemma 4. (Polynomial of $n$-th power)

$$(\sum_{i=1}^{\infty}\alpha_i x^i)^m=\alpha_1^m x^m+\cdots + (\sum_{\substack{k_1+k_2+\cdots k_n=m,k_i\in \mathrm{N}\\ k_1+\cdots+ik_i+\cdots+nk_n=l}}\frac{\alpha_1^{k_1}\alpha_2^{k_2}\cdots\alpha_n^{k_n}}{k_1!k_2!\cdots k_n!}m!)x^l+\cdots \quad (5)$$

Another expression can be seen by Theorem 1 in [31]. In order to simplify the symbols, set the coefficient of $x^l$ as

$$N_m^l=\sum_{\substack{k_1+k_2+\cdots k_n=m,k_i\in \mathrm{N}\\ k_1+\cdots+ik_i+\cdots+nk_n=l}}\frac{\alpha_1^{k_1}\alpha_2^{k_2}\cdots\alpha_n^{k_n}}{k_1!k_2!\cdots k_n!}m! .$$

Lemma 5. (Nilpotent matrix) $\boldsymbol{I}_n$ is a $n$-dimension identity matrix, let shift matrix be $\boldsymbol{N}\triangleq\begin{pmatrix}0 & \boldsymbol{I}_n\\ 0 & 0\end{pmatrix}$ , if $i\ge n+1$ , then $\boldsymbol{N}^i=0$ .

Lemma 6. (Faà di Bruno formula, [10]) If $g$ and $f$ are functions with a sufficient number of derivatives, then we have

$$\frac{d^n g[f(t)]}{dt^n}=\sum_{i=1}^{n}[\sum_{\substack{k_1+k_2+\cdots+k_n=i,\ k_j\ge 0\\ k_1+2k_2+\cdots+nk_n=n}}\frac{n!}{k_1!\cdots k_n!}g^{(i)}(f(t)) \cdot(\frac{f'(t)}{1!})^{k_1}(\frac{f''(t)}{2!})^{k_2}\cdots(\frac{f^{(n)}(t)}{n!})^{k_n}]$$

It is easy to verify Lemma 2 to Lemma 5, and the proofs can be found in many text books.

### *3.2. Univariate Multinomial Distribution Approximation*

Now we discuss the generalization of $n$ times of Bernoulli trials, and consider it under the situation: $n\to\infty$ . This case is similar to univariate multinomial distribution (see [18] p522): Suppose that there is a sequence of $n$ independent trials during segment length of $t$, where each trial has infinite possible outcomes $A_0,A_1,A_2,\cdots$ that are mutually exclusive and the probability of each outcomes can be written as :

$$p_0=1-\lambda\Delta t+o(\Delta t),(\sum_{i=0}^{\infty}p_i=1,\sum_{i=1}^{\infty}\alpha_i=1),$$

$$p_i=p(A_i)=\lambda\alpha_i\Delta t + o(\Delta t),$$

respectively. Let the occurrence of $A_i (i=1,2,\cdots)$ be deemed to be equivalent to $i$ successes and the occurrence of $A_0$ be deemed to be a failure. The number of successes $X(t)$ achieved in the $n$ trials. When $n\to+\infty$ , the reader can see in Fig.3, we divide $[t_0,t_0+t)$ in to $N$ pieces evenly, and every interval is $\Delta t$ , that is $\lim_{N\to\infty}N\Delta t=t$ . Set them

$$[t_0,t_0+\Delta t),[t_0+\Delta t,t_0+2\Delta t),\cdots,[t_0+(N-1)\Delta t,t_0+N\Delta t).$$

There is only one events choosing from $A_0, A_1, A_2, \cdots$ in every $[t_0+h\Delta t, t_0+(h+1)\Delta t), (h=1,2\cdots)$ . The number of events $A_1, A_2, \cdots$ happened in all small intervals are limited, while the number of $A_0$ happened during $[t_0, t_0+t)$ is infinite. To understand the process easier, we imagine there is a dice with infinite surfaces and each surface write one number $0,1,2,\cdots$ . During time interval $[t_0, t_0+t)$ , after $N=\frac{t}{\Delta t}$ (approximate to an integer) times of tossing, figure out the probability that total number is $n$.


***Figure 2.*** *Divide t into infinity many interval of* $\Delta t$ *uniformly*

When $N\to+\infty$ , DCP distribution is equal to a conditional multinomial distribution that was produced by a generalized infinite times independent repetition trial, the pgf in one trial is $G(s)=\sum_{i=0}^{+\infty}p_i s^i, (|s|\le 1)$ . Then pgf $P(s)=[G(s)]^N$ of DCP can be written as:

$$\lim_{N\to\infty}[p_0{}^N+\cdots+(\sum_{\substack{k_0+k_1+\cdots k_n=N, k_u\in\mathrm{N}\\ 0\cdot k_0+k_1+\cdots+nk_n=n}} N!\frac{p_0^{k_0}p_1^{k_2}\cdots p_n^{k_n}}{k_0!k_1!\cdots k_n!})s^n+\cdots]$$

by Lemma 5. According to the pgf, we know $P_n(t)$ is the coefficient of $s^n$ in the expansion formula of $P(s)$ , since

$$p_i=\frac{\alpha_i\lambda t}{N},(i=1,2,\cdots), p_0=1-\frac{\lambda t}{N},\ N=\sum_{i=1}^{\infty}\alpha_i\lambda t$$

where $k_1,\ldots,k_n$ are fixed nonnegative integers and the constraint is $\sum_{u=0}^{n}k_u=N$ . Then,

$$\begin{aligned}P_n(t)&=\lim_{N\to\infty}\sum_{\substack{k_0+k_1+\cdots k_r=N,k_u\in\mathrm{N}\\ 0\cdot k_0+k_1+\cdots+nk_n=n}}\frac{N!}{k_0!k_1!\cdots k_n!}p_0^{k_0}p_1^{k_1}\cdots p_n^{k_n}\\ &=\lim_{N\to\infty}[1-\frac{(\alpha_1\lambda+\cdots+\alpha_n\lambda)t}{N}]^{N-(k_1+\cdots+k_n)}\\ &\cdot\lim_{N\to\infty}\sum_{\substack{k_0+k_1+\cdots k_n=N,k_u\in\mathrm{N}\\ 0\cdot k_0+k_1+\cdots+nk_n=n}}\frac{(\alpha_1\lambda)^{k_1}\cdots(\alpha_n\lambda)^{k_n}}{k_1!\cdots k_n!}t^{k_1+\cdots+k_n}\\ &\cdot\lim_{N\to\infty}\frac{N(N-1)\cdots[N-(k_1+\cdots+k_n)+1]}{N^{k_1+\cdots+k_n}}\\ &=\sum_{k_1+2k_2+\cdots+nk_n=n,k_u\in\mathrm{N}}\frac{(\alpha_1\lambda)^{k_1}\cdots(\alpha_n\lambda)^{k_n}}{k_1!\cdots k_n!}t^{k_1+\cdots+k_n}e^{-\lambda t}\end{aligned}$$

### *3.3. System of Differential Equation*

New we consider $X(t)\sim CP(\alpha_1\lambda t,\alpha_2\lambda t,\cdots)$.

If $X(t)\sim CP(\alpha_1\lambda t,\cdots,\alpha_r\lambda t)$ , when $i>r$ , we have $\alpha_i=0$.

If $i=0$ , we have $P_0(t+\Delta t)=P_0(t)P_0(\Delta t)$. By Chapman-Kolmogorov equation (it also called the total probability formula), solved the equation above by Lemma 2, then

$$P_0(\Delta t)=\mathrm{e}^{-\lambda\Delta t}=1-\lambda\Delta t+o(\Delta t),(\lambda>0)$$

By the independent increment, Chapman-Kolmogorov equation and superimposition, it is easy to see that

$$\begin{aligned}P_i(t+\Delta t)&=[1-\lambda\Delta t+o(\Delta t)]P_i(t)+[\lambda\alpha_1\Delta t+o(\Delta t)]P_{i-1}(t)\\ &+\cdots+[\lambda\alpha_i\Delta t+o(\Delta t)]P_0(t)\end{aligned}$$

$$\begin{aligned}\frac{P_i(t+\Delta t)-P_i(t)}{\Delta t}&=\lambda[-P_i(t)+\alpha_1P_{i-1}(t)+\cdots+\alpha_{i-1}P_1(t)\\ &+\alpha_iP_0(t)]+o(\Delta t)\end{aligned}$$

Let $\Delta t\to 0$ , we obtain a difference-differential equation with initial conditions:

$$P_i'(t)=\lambda[-P_i(t)+\alpha_1P_{i-1}(t)+\cdots+\alpha_{i-1}P_1(t)+\alpha_iP_0(t)], P_0(t)=\mathrm{e}^{-\lambda t} \quad (6)$$

Equation (6) can also be represented by matrix, as follow:

$$\begin{pmatrix}P_n(t)\\ \vdots\\ P_1(t)\\ P_0(t)\end{pmatrix}'=\begin{pmatrix}-\lambda & \lambda\alpha_1 & \cdots & \lambda\alpha_n\\ & \ddots & \ddots & \vdots\\ & & \ddots & \lambda\alpha_1\\ & & & -\lambda\end{pmatrix}\begin{pmatrix}P_n(t)\\ \vdots\\ P_1(t)\\ P_0(t)\end{pmatrix}$$

We can obtain the equation $\boldsymbol{P}_n'(t)=\boldsymbol{Q}\boldsymbol{P}_n(t)$ , and the characteristic equation of the coefficient matrix $\boldsymbol{P}$ is $\det(\boldsymbol{P}-\lambda'\boldsymbol{E})=(-\lambda-\lambda')^{n+1}=0$ , so $-\lambda$ is a characteristic root of $\boldsymbol{P}$ . The solution of $\boldsymbol{F}^{n+1}(-\lambda)\boldsymbol{A}=0$ is an arbitrary constant vector. Furthermore, $c\boldsymbol{A}$ is a certain vector by the uniqueness of pmf and (4), and it won't changed with the change of $\alpha_1,\alpha_2,\cdots$.

Considering an extreme case: if $\alpha_1=1$ , then we have

$$\boldsymbol{P}_n^*(t)=(\frac{(\lambda t)^n e^{-\lambda t}}{n!},\cdots,\lambda te^{-\lambda t},e^{-\lambda t})^T,$$

where $\boldsymbol{P}_n^*(t)$ can be expressed as

$$\boldsymbol{P}_n^*(t)=[E_{n+1}+\begin{pmatrix}0 & \boldsymbol{I}_n\\ 0 & 0\end{pmatrix}\lambda t+\begin{pmatrix}0 & 0 & \boldsymbol{I}_{n-1}\\ & 0 & 0\\ & & 0\end{pmatrix}\frac{(\lambda t)^2}{2!}+\cdots+$$

$$\begin{pmatrix}0 & \cdots & 1\\ & \ddots & \vdots\\ & & 0\end{pmatrix}\frac{(\lambda t)^n}{n!}]c\boldsymbol{A}e^{-\lambda t}$$

By Lemma 3 and $\boldsymbol{I}_n$ is an $n$-dimension identity matrix. Thus we obtain

$$\begin{pmatrix} 1 & \lambda t & \cdots & \frac{(\lambda t)^n}{n!} \\ & \ddots & \ddots & \vdots \\ & & \ddots & \lambda t \\ & & & 1 \end{pmatrix} cAe^{-\lambda t} = \begin{pmatrix} \frac{(\lambda t)^n}{n!}e^{-\lambda t} \\ \vdots \\ \lambda t e^{-\lambda t} \\ e^{-\lambda t} \end{pmatrix}$$

Solve this system of equation, we obtain

$$c\boldsymbol{A} = (0,...,0,1)^T \cdot$$

Using the expression of particular solution to solve $\boldsymbol{F}(-\lambda)c\boldsymbol{A},\cdots,\boldsymbol{F}^n(-\lambda)c\boldsymbol{A}$, we can get

$$\boldsymbol{F}(-\lambda)c\boldsymbol{A} = \lambda \begin{pmatrix} 0 & \alpha_1 & \cdots & \alpha_{n-1} & \alpha_n \\ & 0 & \alpha_1 & \ddots & \alpha_{n-1} \\ & & 0 & \alpha_1 & \vdots \\ & & & 0 & \alpha_1 \\ & & & & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

$$= \lambda(\alpha_1\boldsymbol{N} + \alpha_2\boldsymbol{N}^2 + \cdots + \alpha_n\boldsymbol{N}^n)^1 c\boldsymbol{A} = \lambda(N_1^n,...,N_1^1,0)^T$$

$$\boldsymbol{F}^2(-\lambda)c\boldsymbol{A} = \lambda^2 \begin{pmatrix} 0 & 0 & N_2^1 & \cdots & N_2^{n-1} \\ & 0 & 0 & \ddots & \vdots \\ & & 0 & 0 & N_2^1 \\ & & & 0 & 0 \\ & & & & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

$$= \lambda^2(\alpha_1\boldsymbol{N} + \alpha_2\boldsymbol{N}^2 + \cdots + \alpha_n\boldsymbol{N}^n)^2 c\boldsymbol{A} = \lambda^2(N_2^n,...,N_2^2,0,0)^T$$

$$, \ldots\ldots,$$

$$\boldsymbol{F}^i(-\lambda)c\boldsymbol{A} = \lambda^i(\alpha_1\boldsymbol{N} + \alpha_2\boldsymbol{N}^2 + \cdots + \alpha_n\boldsymbol{N}^n)^i c\boldsymbol{A}$$
$$= \lambda^i(N_i^n,...,N_i^i,0,\cdots,0)^T,\cdots,$$
$$\boldsymbol{F}^n(-\lambda)c\boldsymbol{A} = \lambda^n(\alpha_1\boldsymbol{N} + \alpha_2\boldsymbol{N}^2 + \cdots + \alpha_n\boldsymbol{N}^n)^n c\boldsymbol{A}$$
$$= \lambda^n(N_n^n,0\cdots,0)^T$$

According to Lemma 3, it will be noticed that the unique solution of (6) is as follow:

$$(P_n(t),\cdots,P_1(t),P_0(t))^T$$
$$= [(0,...,0,1)^T + (N_1^n,...,N_1^1,0)^T\lambda t + (N_2^n,...,N_2^2,0,0)^T\frac{(\lambda t)^2}{2!}$$
$$+\cdots+(N_i^n,...,N_i^i,0,\cdots,0)^T\frac{(\lambda t)^i}{i!} + \cdots + (N_n^n,0,...,0)^T\frac{(\lambda t)^n}{n!}]e^{-\lambda t}$$

By the definition of $N_i^n$, to show (1), we only consider the first row of $\boldsymbol{P}_n(t)$.

### *3.4. Matrix Differential Equation*

According to the proof in system of differential equation method, we have the identity in matrix form

$$\boldsymbol{P}_n'(t) = \boldsymbol{Q}\boldsymbol{P}_n(t),$$

thus

$$\boldsymbol{P}_n(t) = e^{\boldsymbol{Q}t}\boldsymbol{C} = \sum_{m=0}^{\infty}\frac{(-\boldsymbol{I} + \alpha_1\boldsymbol{N} + \alpha_2\boldsymbol{N}^2 + \cdots + \alpha_n\boldsymbol{N}^n)^m}{m!}(\lambda t)^m\boldsymbol{C}$$

Let $t \to 0$, hence we have the initial condition:

$$(0,\cdots,0,1)^T = \lim_{t\to 0}\boldsymbol{P}_n(t) = \lim_{t\to 0}e^{\boldsymbol{Q}t}\boldsymbol{C} = \boldsymbol{C}$$

By Lemma 4, we have

$$\frac{(-\boldsymbol{I} + \sum_{i=1}^{n}\alpha_i\boldsymbol{N}^i)^m}{m!} = \frac{(-\boldsymbol{I})^m}{m!} + \cdots +$$
$$(\sum_{\substack{k_0+k_1+\cdots+k_n=m,k_i\in\mathrm{N} \\ k_1+\cdots+ik_i+\cdots+nk_n=l}}\frac{(-\boldsymbol{E})^{k_0}\alpha_1^{k_1}\cdots\alpha_n^{k_n}}{k_0!k_1!\cdots k_n!})\boldsymbol{N}^l + \cdots + \frac{\alpha_n^m\boldsymbol{N}^{nm}}{m!},$$

$$\boldsymbol{P}_n(t) = \sum_{k_0=0}^{\infty}\sum_{l=1}^{nm}[\sum_{\substack{k_0+k_1+\cdots+k_n=m,k_i\in\mathrm{N} \\ k_1+\cdots+ik_i+\cdots+nk_n=l}}\frac{(-1)^{k_0}(\lambda t)^{k_0}}{k_0!}$$
$$\cdot\frac{\alpha_1^{k_1}\alpha_2^{k_2}\cdots\alpha_n^{k_n}}{k_2!\cdots k_n!}(\lambda t)^{k_1+\cdots+k_n}]\boldsymbol{N}^l\boldsymbol{C}$$
$$= \sum_{l=1}^{n}\sum_{m=1}^{n}\sum_{\substack{k_1+k_2+\cdots k_n=m,k_i\in\mathrm{N} \\ k_1+\cdots+ik_i+\cdots+nk_n=l}}\frac{\alpha_1^{k_1}\alpha_2^{k_2}\cdots\alpha_n^{k_n}}{k_1!k_2!\cdots k_n!}(\lambda t)^{k_1+\cdots+k_n}e^{-\lambda t}\boldsymbol{N}^l\boldsymbol{C}.$$

Because of $\boldsymbol{N}^l\boldsymbol{C} = (1,0,\cdots,0)^T$, if we choose the first row of $\boldsymbol{P}_n(t)$, (1) is easy to be prove by the statement above.

The following two methods is the derivation of the pmf by pgf.

### *3.5. Faà di Bruno Formula*

Firstly, we give two approaches to prove Theorem 2. Similar to the binomial distribution approximate Poisson distribution, as well as the building method in the Section 3.4, it is easily to obtain the parameter limited conditions of multinomial distribution approximate to DCP by pgf:

$$P(s) = \lim_{\substack{Np_k=\alpha_i\lambda t \\ N\to\infty}}[(1-\sum_{i=1}^{\infty}p_i) + \sum_{i=1}^{\infty}p_is^i]^N \;.$$
$$= \lim_{N\to\infty}[1 + \frac{\lambda t}{N}(\sum_{i=1}^{\infty}\alpha_is^i - \sum_{i=1}^{\infty}\alpha_i)]^N = e^{\lambda t\sum_{i=1}^{\infty}\alpha_i(s^i-1)}$$

Or by defining the pgf as $P(t,s)=\sum_{i=0}^{\infty}P_i(t)s^i$, substitute to (5) we can get the expression as an ODE

$$\frac{\partial Q(t,s)}{\partial t} = [\lambda t\sum_{I=1}^{\infty}\alpha_i(s^i-1)]Q(t,s) \;.$$

Thus we have

$$Q(t,s) = ce^{\lambda t\sum_{i=1}^{\infty}\alpha_i(s^i-1)} \;.$$

Let $s = 1$, by the definition of pgf, we know that (2) is the solution of the ODE above.

Then, according to the inversion formula of generating function, as well as the generalized form of Leibniz formula—Faà di Bruno formula, it is immediately to get the pmf of DCP. By the pgf of DCP, we have

$$\{(j!)^{-1}\sum_{i=1}^{\infty}[\alpha_i\lambda t(s^i-1)]^{(j)}\Big|_{s=0}\}^{k_j}=(\alpha_i\lambda t)^{k_j},$$

$$[e^{\lambda t\sum_{i=1}^{\infty}\alpha_i(s^i-1)}]^{(i)}\Big|_{s=0}=e^{-\lambda t}.$$

$$\begin{aligned}\rightarrow P_n(t)&=[\frac{1}{n!}\frac{d^n}{ds^n}e^{\lambda t\sum_{i=1}^{\infty}\alpha_i(s^i-1)}]\Big|_{s=0}\\&=\sum_{i=1}^{n}[\sum_{\substack{k_1+k_2+\cdots+k_n=i,k_u\in\mathrm{N}\\k_1+2k_2+\cdots+nk_n=n}}\frac{\alpha_1^{k_1}\alpha_2^{k_2}\cdots\alpha_n^{k_n}}{k_1!k_2!\cdots k_n!}(\lambda t)^i]e^{-\lambda t}\end{aligned}$$

Let $i$ running from 1 to $n$, and the Theorem 1 is proved.

### 3.6. Cauchy'S Integral Formula

Utilizing the relationship of Cauchy formula of higher order derivative and power series as well as formula (4), we have the pmf

$$\begin{aligned}P_n(t)&=\frac{1}{2\pi i}\int_C\frac{1}{z^{n+1}}e^{-\lambda t+\sum_{i=1}^{\infty}\alpha_i\lambda tz^i}dz\\&=\frac{1}{2\pi i}\int_C\frac{1}{z^{n+1}}\sum_{m=0}^{\infty}\frac{(-1+\alpha_1z+\alpha_2z^2+\cdots)^m}{m!}(\lambda t)^m dz\\&=\frac{1}{2\pi i}\int_C\frac{1}{z^{n+1}}\sum_{k_0=0}^{\infty}\sum_{l=1}^{\infty}[\sum_{\substack{k_0+k_1+\cdots+k_n=m,k_i\in\mathrm{N}\\k_1+\cdots+ik_i+\cdots+nk_n=l}}\frac{(-1)^{k_0}(\lambda t)^{k_0}}{k_0!}\\&\cdot\frac{\alpha_1^{k_1}\cdots\alpha_n^{k_n}}{k_1!\cdots k_n!}(\lambda t)^{k_1+\cdots+k_n}]z^l dz\\&=\frac{1}{2\pi i}\int_C\frac{1}{z^{n+1}}\sum_{l=1}^{\infty}(\sum_{\substack{k_0+k_1+\cdots+k_n=m,k_i\in\mathrm{N}\\k_1+\cdots+ik_i+\cdots+nk_n=l}}\frac{\alpha_1^{k_1}\cdots\alpha_n^{k_n}(\lambda t)^{k_1+\cdots+k_n}e^{-\lambda t}}{k_1!\cdots k_n!})z^l dz\end{aligned}$$

By the results of complex integration, $a$ is an internal point in $C$ :

$$\frac{1}{2\pi i}\int_C\frac{1}{(z-a)^n}dz=\begin{cases}1, & n=1\\0, & n\neq 1,n\in\mathrm{Z}\end{cases}.$$

When $l=n$ , (2) is clearly.

### 3.7. Functional Equations

When $i=0$ , the Chapman-Kolmogorov equation makes it obviously that $P_0(t+\Delta t)=P_0(t)P_0(\Delta t)$. By Lemma 1, then we have

$$P_0(\Delta t)=\mathrm{e}^{-\lambda\Delta t}=1-\lambda\Delta t+o(\Delta t),(\lambda>0) \quad (7)$$

When $i=1$, similarly, we have:

$$P_1(t+\Delta t)=P_0(t)P_1(\Delta t)+P_1(t)P_0(\Delta t). \quad (8)$$

By Lemma 1, $P_1(\Delta t)=\alpha_1te^{-\lambda t}$ $(\alpha_1\geq 0)$ is clearly to see by (7) and (8).

Continually, we have formulas below:

$$P_2(t)=[\alpha_2\lambda t+\frac{1}{2!}\alpha_1^2(\lambda t)^2]e^{-\lambda t}\ (\alpha_2\geq 0),\cdots$$

$$P_n(t)=\sum_{\Sigma_{i=1}^{n}ik_i=n,k_i\in\mathrm{N}}\frac{\alpha_1^{k_1}\alpha_2^{k_2}\cdots\alpha_n^{k_n}}{k_1!k_2!\cdots k_n!}(\lambda t)^{k_1+k_2+\cdots+k_n}e^{-\lambda t}$$

Then we use mathematical induction method to prove that (2) is true.

Moreover, we have

$$P_i(\Delta t)=\lambda\alpha_i\Delta t+o(\Delta t),(i=1,2,\cdots),$$

and it is easy to see the parameters of DCP process. $\alpha_i$ must be nonnegative constant due to $P_i(\Delta t)\geq 0$ .

### 3.8. Compound Poisson Sum

DCP process $X(t)$ can be decomposed as

$$X(t)=Y_1+Y_2+\cdots+Y_{N(t)},$$

where $Y_i$ is i.i.d nonnegative integer-valued random variables with $P\{Y_1=j\}=\alpha_j$ , and $N(t)\sim P(\lambda t)$ .

We first verify that the pgf is formula (2), by the conditional expectation and independence:

$$\begin{aligned}\mathrm{E}s^{X(t)}&=\sum_{n=0}^{\infty}\mathrm{E}(s^{X(t)}\mid N(t)=n)P(N(t)=n)\\&=\sum_{n=0}^{\infty}(\mathrm{E}s^{Y_1})^n\frac{(\lambda t)^ne^{-\lambda t}}{n!}=e^{\lambda t\sum_{i=1}^{\infty}\alpha_i(s^i-1)}\end{aligned}$$

Hence, by independence, we can figure that

$$\begin{aligned}P_n(t)&=P\{\sum_{i=1}^{N(t)}Y_i=n\}\\&=P\{k_i\sim\frac{(\alpha_i\lambda t)^{k_i}e^{-\alpha_i\lambda t}}{k_i!},i=1,2,\cdots,n\Big|\Sigma_{i=1}^{n}ik_i=n,k_i\in N\}\\&=\sum_{\Sigma_{i=1}^{n}ik_i=n,k_i\in N}\frac{\alpha_1^{k_1}\alpha_2^{k_2}\cdots\alpha_n^{k_n}}{k_1!k_2!\cdots k_n!}(\lambda t)^{k_1+k_2+\cdots+k_n}e^{-\lambda t}\end{aligned}$$

### 3.9. Sum of Weighted Poisson

DCP process $X(t)$ can be decomposed as

$$X(t)=Z_1(t)+2Z_2(t)+\cdots+iZ_i(t)+\cdots,$$

where $Z_i(t)$ is independent of each other, and

$$Z_i(t)\sim P(\lambda t\alpha_i).$$

We first verify that the pgf of $X(t)$ is formula (3), by the conditional expectation and independence:

$$\begin{aligned}\mathrm{E}s^{X(t)}&=\mathrm{E}s^{\sum_{i=1}^{\infty}iZ_i(t)}=\prod_{i=0}^{\infty}\mathrm{E}(s^i)^{Z_i(t)}\\&=\prod_{i=0}^{\infty}e^{\lambda t\alpha_i(s^i-1)}=e^{\lambda t\sum_{i=1}^{\infty}\alpha_i(s^i-1)}.\end{aligned}$$

To verify its pmf, by independence, we have:

$$P_n(t)=P\{\sum_{i=1}^{n}iZ_i(t)=n\}$$
$$=P\{k_i\sim\frac{(\alpha_i\lambda t)^{k_i}e^{-\alpha_i\lambda t}}{k_i!},i=1,2,\cdots,n\Big|\Sigma_{i=1}^{n}ik_i=n,k_i\in N\}$$
$$=\sum_{\Sigma_{i=1}^{n}ik_i=n,k_i\in N}\frac{\alpha_1^{k_1}\alpha_2^{k_2}\cdots\alpha_n^{k_n}}{k_1!k_2!\cdots k_n!}(\lambda t)^{k_1+k_2+\cdots+k_n}e^{-\lambda t}.$$

Remark 3. The expectation of DCP process $X(t)$ is

$$\mathrm{E}X(t)=\mathrm{E}[\Sigma_{i=1}^{\infty}iZ_i]=\Sigma_{i=1}^{\infty}i\alpha_i\lambda t\cdot$$

$\mathrm{E}X(t)$ is finite, iff $\Sigma_{i=1}^{\infty}i\alpha_i<\infty$. By independent increments property,

$$\mathrm{E}\{X(t)-X(s)+X(s)\big|X(\tau),\tau\le s\}=0+\mathrm{E}\{X(s)\big|X(\tau),\tau\le s\}$$
$$=X(s).$$

$X(t)-\Sigma_{i=1}^{\infty}i\alpha_i\lambda t$ satisfies the definition of continuous-time martingale,

$$\mathrm{E}(|X(t)|)<\infty;\ \mathrm{E}(X(t)\big|\{X(t),\tau\le s\})=X(s),(\forall s\le t).$$

Hence, $X(t)-\Sigma_{i=1}^{\infty}i\alpha_i\lambda t$ is called discrete compound Poisson martingale.

The variance of discrete compound Poisson process $X(t)$ is

$$\mathrm{D}X(t)=\mathrm{D}[\Sigma_{i=1}^{\infty}iZ_i]=\Sigma_{i=1}^{\infty}i^2\alpha_i\lambda t\cdot$$

$\mathrm{D}X(t)$ is finite, iff $\Sigma_{i=1}^{\infty}i^2\alpha_i<\infty$. Hence, $X(t)$ is a square-integrable martingale.

Definition 3. (Square-integrable martingale,[27] p59) A square-integrable martingale $\{M(t)\}_{t\ge 0}$ such that

$$\mathrm{E}\left([M(t)\ -M(s)]^2\ |\ \{M(\tau),\tau\le s\}\right)=t-s,\ (0\le s\le t)$$

is called a normal martingale.

The compensated and rescaled process

$$M(t)\triangleq\frac{X(t)-\Sigma_{i=1}^{\infty}i\alpha_i\lambda t}{\sqrt{\Sigma_{i=1}^{\infty}i^2\alpha_i\lambda t}}$$

is a normal martingale (See Chapter 2 of [27]).

### 3.10. Recursive Formula

We begin by defining the indicator function

$$g_j(x)=\begin{cases}0 & ,x\ne j\ne 0\\ j^{-1} & ,x=j\ne 0\end{cases}\cdot$$

Then, $\forall k\in R,n\ge 1$, by the compound Poisson sum in Section 3.8, it follows that

$$P_n(t)=P\{X(t)=n\}=\sum_{i=0}^{+\infty}ig_n(i)P(X(t)=i)=\mathrm{E}[X(t)g_n(X(t))]$$
$$=\mathrm{E}[(\sum_{i=1}^{N(t)}Y_i)g_n(X(t))]=\sum_{k=1}^{\infty}k\mathrm{E}[Y_1g_n(X(t))]\frac{(\lambda t)^k e^{-\lambda t}}{k!}$$
$$=\sum_{k=1}^{\infty}\sum_{j=1}^{n}j\alpha_j\lambda t\mathrm{E}g_n(\sum_{i=1}^{k-1}Y_i+j\big|Y_k=j)\frac{(\lambda t)^{k-1}e^{-\lambda t}}{(k-1)!}$$
$$=\sum_{j=1}^{n}j\alpha_j\lambda t\sum_{k=1}^{\infty}\frac{P_{n-j}(t)}{n}\frac{(\lambda t)^{k-1}e^{-\lambda t}}{(k-1)!}=\frac{\lambda t}{n}\sum_{j=1}^{n}j\alpha_jP_{n-j}(t)$$

and

$$P_0(t)=P\{\sum_{i=1}^{N(t)}Y_i=0\}=P\{N(t)=0\}=e^{-\lambda t}\cdot$$

By using the recursive formula, and using mathematical induction method, it is clearly to get equation (6).

### 3.11. Convolution

According multinomial distribution expansion formula, compound Poisson sum property and Lemma 4 we get

$$P\{X(t)=n\}=\sum_{i=0}^{\infty}P\{X(t)=n\big|N(t)=i\}P\{N(t)=i\}$$
$$=\sum_{i=0}^{\infty}P\{\sum_{k=1}^{i}Y_k=n\}\frac{(\lambda t)^i e^{-\lambda t}}{i!}$$
$$=\sum_{i=0}^{\infty}P\{\frac{1}{n!}\frac{\partial^n}{\partial s^n}(p_1s+p_2s^2+\cdots)^i\}\frac{(\lambda t)^i e^{-\lambda t}}{i!}$$
$$=\sum_{i=0}^{\infty}\sum_{\substack{k_1+\cdots+k_n=i\ ,k_u\in\mathbb{N}\\ k_1+\cdots+uk_u++nk_n=n}}\frac{p_1^{k_1}\cdots p_n^{k_n}}{k_1!\cdots k_n!}(\lambda t)^{k_1+\cdots+k_n}e^{-\lambda t},$$

where $P\{\sum_{k=1}^{i}Y_k=n\}$ is pmf of $i$th convolution of $Y_k$.

### 3.11. Remark of the Methods

Arguing from the Chapman-Kolmogorov equation, Whittaker [33] obtained the difference differential equation relating to $P_n(t)$, solving one by one, he gave the first 5 probability expressions of

$$P_n(t)(n=0,1,2,3,4)\cdot$$

Janossy [17] directly solve the Chapman-Kolmogorov equation to get the expression of $P_n(t)$, it shows that the superimposition condition in the definition of DCP process are unnecessary. Luders[24] name it Pollaczek-Geiringer distribution in memory of these two people's finding, and derived it by the method of sum of weight Poisson. Hofman [16] derive the expression of $P_n(t)$ by the use of Faà di Bruno formula, from the pdf of $P_n(t)$. Adelson [1] differentiate pgf of DCP distribution for several times and obtained the recurrence relation of pmf.

Discrete compound Poisson distribution is most interesting discrete distribution, it's hard to find other distributions have such many properties and the multiple approaches of proving the probability mass function of discrete compound Poisson distribution. Among ten approaches, the system of differential equation, convolution, Matrix differential equation, Cauchy's integral formula are original works, the other methods give a detail and vivid proof from other works.

Let $\alpha_1 = 1$ in DCP distribution, except for method of compound Poisson sum, sum of weight Poisson and convolution, we obtain seven approaches to prove the pmf of Poisson distribution.

When it comes to parameter estimation of DCP, see [18] and [34].

# 4. Probabilistic Number Theory

In the probabilistic number theory, prime divisors of integers and the distribution of primes in short intervals have properties of Poisson distribution under some condition. The detail results are in this theoretical research papers: [15], [11], [8], [28] and [13]. Let $f(m)$ be a real-valued arithmetical function whose domain is positive integer $\{1,2,\cdots\}$. We say that strongly additive functions satisfies the following restriction.

(i). $f(m+n) = f(m) + f(n)$ whenever $(m,n) = 1$.

(ii). $f(p^k) = f(p)$ for all prime $p$ and $k \ge 1$.

For example, let $f(m)$ be the number of prime divisors of $m$. And the additive functions just satisfy (i).

Then we have the following results due to Bekelis [3]. In Bekelis's work, he give the necessary and sufficient conditions for the weak convergence of the distribution functions of strongly additive functions to the finite Poisson law convolutions.

Theorem 4: Let $f_x(m),(x \ge 1)$ be a strongly additive function depend on $x$. Defined the pmf of $u$ by $v_x(f_x(m) = u)$

$$v_x(f_x(m) = u) = \frac{1}{[x]}\#\{m \le x : f_x(m) = u\}$$

Let functions $f_x(m)$ on each prime number $p$ is values from the nonnegative integer set

$$\{0, c_1, c_2, \cdots, c_r\}, 0 < c_1 < c_2 < \cdots < c_r.$$

If

$$\lim_{x\to\infty} \max_{p\le x, f_x(p)\ne 0} \frac{1}{p} = 0; \ \lim_{x\to\infty} \sum_{p\le x, f_x(p)\ne 0} \frac{\ln p}{p \ln x} = 0; \quad \lim_{x\to\infty} \sum_{p\le x, f_x(p)=c_i} \frac{1}{p} = \lambda_i, (j = 1, \cdots, r), \tag{9}$$

then the limited pgf of $v_x(f_x(m) = u)$ is

$$\lim_{x\to\infty} \sum_{u=0}^{\infty} v_x(f_x(m) = u)s^u = e^{\sum_{i=1}^{r} \lambda_i (s^{c_i} - 1)}, (|s| \le 1).$$

Hence $u \sim CP(\cdots, \lambda_1, \cdots, \lambda_2, \cdots, \lambda_r)$.

Theorem 5: Let strongly additive functions $f_x(m)$, $(x \ge 1)$ be as described above, and the limited pgf of $v_x(f_x(m) = u)$ is

$$\lim_{x\to\infty} \sum_{u=0}^{\infty} v_x(f_x(m) = u)s^u = e^{\sum_{i=1}^{r} \lambda_i (s^{c_i} - 1)}, (|s| \le 1),$$

if numbers $c_j, (j = 1,2,\cdots,r)$ are linearly independent over the field $\mathbb{Q}$, then conditions (9) are satisfied (i.e., they are necessary and sufficient in this case).

The proof of Theorem 4 and Theorem 5 is in [3], then rewrite the corresponding characteristic function to the pgf of $v_x(f_x(m) = u)$.

Šiaulys [28] prove that if $f$ is a strongly additive function on prime number with $f(p) \in \{0,1\}$, then $f$ is Poisson distributed on the integers under the condition that $r = 1$ in (9). This is a special case of Theorem 4.

Numerical example: Let $f_x(m) = \sum_{p|m} f_x(p)$ be the sum of $f_x(p)$ under condition that $p|m$, where

$$f_x(p) = \begin{cases} 0, & p < \ln\ln x \ or \ p > (\ln\ln x)^4 \\ 1, & \ln\ln x \le p \le (\ln\ln x)^2 \\ 2 & (\ln\ln x)^2 \le p \le (\ln\ln x)^4 \end{cases}$$

Bekelis [3] obtain

$$\lambda_1 = \lim_{x\to\infty} \sum_{\substack{p\le x \\ f_x(p)=1}} \frac{1}{p} = \ln 2, \ \lambda_2 = \lim_{x\to\infty} \sum_{\substack{p\le x \\ f_x(p)=2}} \frac{1}{p} = \ln 2 .$$

Thus $u$ approximate to Hermite distribution with pmf

$$v_x(f_x(m) = u) \approx \sum_{i=0}^{[\frac{u}{2}]} \frac{\lambda_1^{u-2i} \lambda_2^i}{(u-2i)!i!} e^{-(\lambda_1+\lambda_2)} = \sum_{i=0}^{[\frac{u}{2}]} \frac{1}{4} \frac{(\ln 2)^{u-i}}{(u-2i)!i!}$$

***Table 1.*** *Frequent when x=1000, 10000, 20000*

| Frequent number | Theoretical frequent | Frequent (x=1000) | Frequents (x=10000) | Frequents (x=20000) |
|---|---|---|---|---|
| 0 | 0.25 | 0.2407 | 0.3127 | 0.3064 |
| 1 | 0.1733 | 0.282 | 0.2272 | 0.226 |
| 2 | 0.2333 | 0.21 | 0.2165 | 0.2168 |
| 3 | 0.134 | 0.161 | 0.1332 | 0.1344 |
| 4 | 0.1041 | 0.064 | 0.0654 | 0.0678 |
| 5 | 0.0516 | 0.026 | 0.0291 | 0.0316 |
| 6 | 0.03 | 0.007 | 0.0093 | 0.01 |
| 7 | 0.0132 | 0.003 | 0.0052 | 0.0051 |
| 8 | 0.0063 | 0 | 0.0014 | 0.0017 |
| 9 | 0.0025 | 0 | 0.0001 | 0.0003 |

Notice that frequents under a sufficient large are not equality theoretical frequents in Table 1 (just $u = 3$ is accurate). The $\ln\ln 20000 \approx 2.29$ is not sufficient large.

In order to obtain equality, we would have to let $\ln\ln x \to \infty$ in some fashion. The convergence rate is very slow. Since limited conditions in computer, we can't calculate the larger value of $x$ in this paper by Matlab.

Theorem 6. (Erdos-Wintner, 1939): A necessary and sufficient condition for a real additive function $f(m)$ to have a limiting distribution is that the following three series converge simultaneously for at least one value of the positive real number $R$ :

$$\sum_{|f(p)|>R} \frac{1}{p}; \quad \sum_{|f(p)|\le R} \frac{f(p)^2}{p}; \quad \sum_{|f(p)|\le R} \frac{f(p)}{p}$$

When these conditions are satisfied, the characteristic function of the limit law is given by the convergent product

$$\varphi(\tau) = \prod_p [\sum_{v=0}^{\infty} p^{-v}(1-p^{-1})e^{i\tau f(p^v)}]$$

The limit law is necessarily pure. It is continuous if, and only if $\sum_{f(p)\neq 0} p^{-1} = \infty$.

For further explanation for Erdos-Wintner theorem, see [9] and [29]. Using Erdos-Wintner theorem, we have the following theorem

Proposition 1. In the Erdos-Wintner theorem, if $\sum_{f(p)\neq 0} \frac{1}{p} < \infty$ and $f(m)$ takes non-negative integer value, then $\varphi(\tau)$ approximate to a discrete compound Poisson distribution with pgf

$$P(s) = e^{\sum_{f(p)\neq 0}\sum_{v=1}^{\infty} p^{-v}(1-p^{-1})s^{f(p^v)} - \sum_{f(p)\neq 0} p^{-1}}, (|s| \le 1)$$

Proof: Rewrite the characteristic function $\varphi(\tau)$ to the pgf, and notice that $1+x \approx e^x$ for a sufficient small $x = \sum_{v=1}^{\infty} p^{-v}(1-p^{-1})s^{f(p^v)} \ll 1$, then

$$\begin{aligned} P(s) &= \prod_p [(1+\sum_{v=1}^{\infty} p^{-v}(1-p^{-1})s^{f(p^v)} - p^{-1}] \\ &= \prod_p [(1+\sum_{v=1}^{\infty} p^{-v}(1-p^{-1})s^{f(p^v)} - p^{-1}] \\ &\approx \prod_p e^{\sum_{v=1}^{\infty} p^{-v}(1-p^{-1})s^{f(p^v)} - p^{-1}} \\ &= e^{\sum_{f(p)\neq 0}\sum_{v=1}^{\infty} p^{-v}(1-p^{-1})s^{f(p^v)} - \sum_{f(p)\neq 0} p^{-1}}. \end{aligned}$$

# 5. (S,s) Inventory Policies under Stuttering Poisson Demand

Considering the sales of an enterprise, the interval is a week, and the week sales are random. Then management decides whether to order goods to satisfy the needs of next week. One of the most simple inventory strategy is (s,S) inventory policies: the lower bound of s, the upper bound of S .When the inventory at weekend is less than s ,order stocks to reach S, otherwise, don't order.

In fact, we need to take ordering fee, storage fee, out of stock payment and purchase fee into account and then formulate an inventory strategy to minimize the total average cost. Generally speaking, the arrival of demand satisfies the property of discrete stationary independent increments process. But a customer may purchase the same goods more than one once, so we can assume that the random demand satisfies stuttering Poisson distribution.

Suppose that each ordering fee is $c_0$, the bid of each good is $c_1$, the storage fee of each good is $c_2$ and the loss of each good that out of store is $c_3$. In order to facilitate computation, we assume that the weekly sale $r$ is a nonnegative integer-valued random variable, and the pmf is $p(r)$. If the stock in the end of the week is $x$, order quantity is $u$, so the stock at the beginning of the next week is $x+u$, storage capacity in every week is $x+u-r$.

According to (S, s) inventory strategy, if $x \ge s$, order quantity $u=0$ ; if $x<s$ , $u>0$ and the constraint is $x+u=S$. Now we determine (S, s) through figuring out the minimum of the average cost. The storage cost and the expected value of back-order loss in a week is

$$L(x) = c_2 \int_0^x (x-r)p(r)dr + c_3 \int_x^{\infty} (r-x)p(r)dr$$

The average cost in a week is the sum of order cost, bid cost, storage cost and the back-order loss,

$$J(u) = \begin{cases} c_0 + c_1 u + L(x+u), & u>0 \\ L(x), & u=0 \end{cases}$$

Determination of S:

Now we take the derivative and second derivative of the average cost function about

$$\begin{aligned} \frac{dJ(u)}{du} &= c_1 + c_2 \int_0^{x+u} p(r)dr - c_3 \int_{x+u}^{\infty} p(r)dr \\ &\overset{S=x+u}{=\!=\!=} (c_2+c_3)\int_0^S p(r)dr - (c_3 - c_1) \end{aligned}$$

Let $\frac{dJ(u)}{du} = 0$, then we can solve that when S satisfies

$$\int_0^S p(r)dr = \frac{c_3 - c_1}{c_2 + c_3},$$

then, $J(u)$ reaches the minimum.

Determination of s:

Then, determine the order number s according to $x$. If management decide to order some goods, and the number is $u$, since $x+u=S$, the total cost is

$$J_1 = c_0 + c_1(S-x) + L(S)$$

If the management decides no to order any new good, then the total cost would be $J_2 = L(x)$.

It's obvious that the cost has to meet the constraint $J_2 \le J_1$, so we have:

$$c_1 x + L(x) \le c_0 + c_1 S + L(S).$$

Let $F(x) = c_1 x + L(x) - [c_0 + c_1 S + L(S)]$, set s is the minimum positive root that makes $F(x) = 0$ satisfied. That is

$$s = \min\{x > 0 \mid F(x) = 0\}$$

Generally speaking, we could only use graphical method (numerical method) to find the minimum positive root. When $p(r)$ is discrete random variable, the method to determine (S,s) is similar to the consecutive situation, the only change is to replace $\int$ with $\Sigma$. For more information about inventory strategy of (S,s), see Veinott [30] and Porteus [26].

For instance, in a day, assuming that the coming customers in a shoe store are Poisson distributed with parameter $\lambda = 50$. But because the shoes are durable goods, we can suppose that the probability of buying a pair of shoes once for a customer is 0.8, and the probability of buying two pair of the same shoes once is very small, such as $0.8 \times 0.2$. If we suppose that the probability of a client buying $i$ $(i \ge 1)$ pair of shoes is $0.8 \times 0.2^{i-1}$, we conclude that the amount of sold shoes is stuttering Poisson distributed with parameter $\alpha = 0.2$. Assume that each ordering fee $c_0 = 400$, the purchase price of each pair of shoes $c_1 = 60$, the storage fee of each pair of shoes $c_2 = 1$, the out of stock damage of each pair of shoes $c_3 = 100$. Then using Maple16 to plot and solve the equation

$$\int_0^S p(r)dr = \frac{c_3 - c_1}{c_2 + c_3} = \frac{40}{101}.$$

Under the same $\lambda$, it is obviously that when $r$ subjects to stuttering SPD, $S_A = 60$, $s_A = 43$; but if $r$ subjects to Poisson distribution(PD), $S_B = 47$, $s_B = 12$.

It is easily seen that S in stuttering Poisson distribution is bigger than in Poisson distribution under the same $\lambda$. The demands may arrive in superimposition. We need to increase storage to satisfy demands. High-tech and quantificational inventory management is an effective way to reduce enterprise cost and to improve the quality of service. Due to the uncertainty of customer's arrival, in many cases, stuttering Poisson distribution is closer to the real life applying to some more general case. And it can also be used in other stochastic process with the characteristic of superposition, such insurance claims data [34].

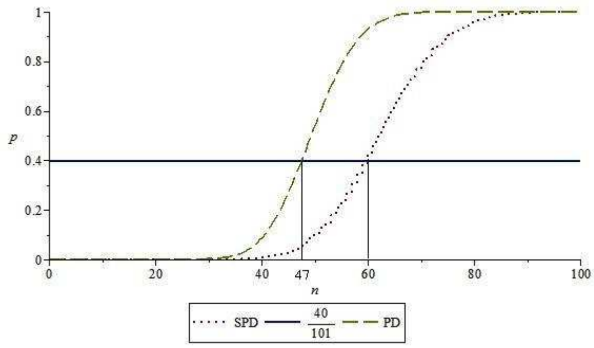
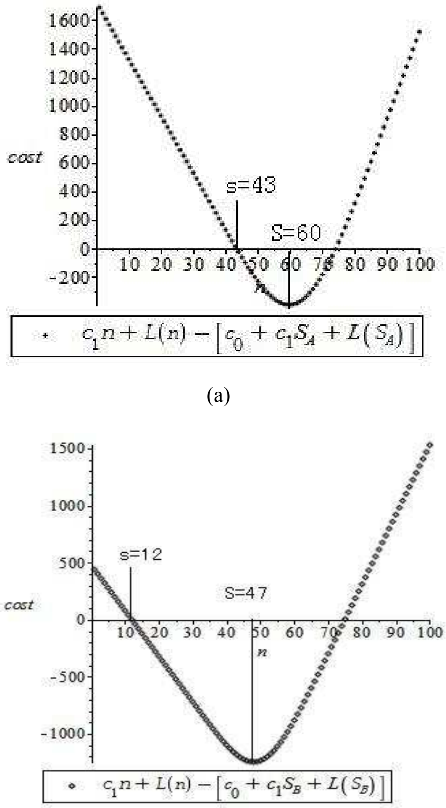



***Figure 4.*** *determine* S *when r subjects to SPD or PD.*




(b)

***Figure 5.*** *The graph of* $F(n)$ *((a):SPD; (b):PD ).*

# References

[1] Adelson R M. Compound poisson distributions. Operational Research.1966,17(1):73-75.

[2] Barndorff-Nielsen O E, Pollard D G, Shephard N. Integer-valued Lévy processes and low latency financial econometrics. Quantitative Finance. 2012;12(4):587-605.

[3] Bekelis D. Convolutions of the Poisson laws in number theory, In Analytic & Probabilistic Methods in Number Theory: Proceedings of the 2nd International Conference in Honour of J. Kubilius, Lithuania, 23-27 September 1996 (Vol. 4, p. 283). Netherlands:VSP BV,1997.

[4] Bell E T. Exponential polynomials. The Annals of Mathematics. 1934,35(2): 258-277.

[5] Bertoin, J. Lévy processes. London: Cambridge university press,1998.

[6] De Finetti, B. Sulla possibilita di valori eccezionali per una legge ad incrementi aleatori. Atti della Reale Accademina Nazionale dei Lincei (Ser VI). 1929,10:325-329.

[7] De Finetti, B. Le funzioni caratteristiche di legge istantanea, Rendicontidella R. Accademia Nazionale dei Lincei (Ser VI). 1930,12:278-282.

[8] De Koninck J M, Galambos J. The intermediate prime divisors of integers. Proceedings of the American Mathematical Society. 1987,101(2):213-216.

[9] Erdös P, Wintner A. Additive arithmetical functions and statistical independence. American Journal of Mathematics. 1939,61(3):713-721.

[10] Faa di Bruno, CF. Sullo sviluppo delle funzioni. Annali di Scienze Matematiche e Fisiche. 1855,6:479-480.

[11] Gallagher PX. On the distribution of primes in short intervals. Mathematika. 1976,23(1): 4-9.

[12] Galliher H P, Morse P M, Simond M. Dynamics of two classes of continuous-review inventory systems. Operations Research. 1959,7(3):362-384.

[13] Granville A. Prime divisors are Poisson distributed. International Journal of Number Theory. 2007,3(01):1-18.

[14] Haight FA. Handbook of the Poisson distribution. Los Angeles: Wiley;1967.

[15] Halász G. On the distribution of additive and the mean-values of multiplicative arithmetic functions. Studia. Sci. Math. Hungarica. 1971;6:211–233.

[16] Hofmann M. Über zusammengesetzte Poisson-Prozesse und ihre Anwendungen in der Unfallversicherung. Diss. Math. ETH Zürich, Nr. 2511, 0000. Ref.: Saxer, W.; Korref.: Nolfi, P., 1955.

[17] Jánossy L, Rényi A, Aczél J. On composed Poisson distributions, I. Acta Mathematica Hungarica. 1950,1(2):209-224.

[18] Johnson N L, Kemp A W, Kotz S. Univariate discrete distributions. New Jersey: Wiley ; 2005.

[19] Kemp C D, Kemp A W. Some properties of the 'Hermite' distribution. Biometrika. 1965,52(3-4):381-394.

[20] Khintchine AY. A new derivation of a formula by P. Lévy. Bulletin of the Moscow State University. 1937,1:1-5. [Russian]

[21] Khintchine AY. Mathematical methods of queuing theory. Proceedings of the Steklov Institute. VA Steklov, 1955,49(0):3-122. [Russian]

[22] Kolmogorov A N. Sulla forma generale di un processo stocastico omogeneo. Atti Accad. Naz. Lincei. 1932,15: 805-808.

[23] Lévy P. Sur les intégrales dont les éléments sont des variables aléatoires indépendantes. Annali della Scuola Normale Superiore di Pisa-Classe di Scienze. 1934,3(3-4):337-366.

[24] Lüders R. Die statistik der seltenen ereignisse. Biometrika. 1934,26(1/2):108-128.

[25] Poisson SD. Recherches sur la probabilité des jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabilités. Paris: Bachelier,1837.

[26] Porteus EL. Foundations of stochastic inventory theory. Stanford: Stanford University Press,2002.

[27] Privault N. Stochastic Analysis in Discrete and Continuous Settings With Normal Martingales. Berlin: Springer, 2009.

[28] Šiaulys J. The von Mises theorem in number theory. New Trends in Probability and Statistics. 1992,2:293-310.

[29] Tenenbaum G. Introduction to analytic and probabilistic number theory. London: Cambridge university press,1995.

[30] Veinott A F, Wagner H M. Computing optimal (s, S) inventory policies. Management Science. 1965,11(5):525-552.

[31] Voinov V G, Nikulin M. On power series, Bell polynomials, Hardy-Ramanujan-Rademacher problem and its statistical applications. Kybernetika. 1994,30(3):343-358.

[32] Wimmer G, Altmann G. The multiple poisson distribution, its characteristics and a variety of forms[J]. Biometrical journal. 1996, 38(8): 995-1011.

[33] Whittaker J M. The shot effect for showers. Mathematical Proceedings of the Cambridge Philosophical Society. Cambridge University Press, 1937,33(04):451-458.

[34] Zhang H, Chu L, Diao Y. Some Properties of the Generalized Stuttering Poisson Distribution and Its Applications. Studies in Mathematical Sciences. 2012,5(1) :11-26.