

# A Study of the Suprenewal Process

by

Chia-Ling Lai

Advisor

Wen-Jang Huang

Institute of Statistics

National University of Kaohsiung

Kaohsiung, Taiwan 811, R.O.C.

June 2006

# Contents

|   |     |
|---|-----|
| 中文摘要.....   | ii  |
| 英文摘要.....   | iii |
| 1. Introduction.....  | 1   |
| 2. Comparisons of $\{Q_t, t \geq 0\}$ and $\{N_t, t \geq 0\}$ ..... | 2   |
| 3. Some fundamental properties of $\{Q_t, t \geq 0\}$ .....         | 4   |
| 4. Limiting and some other related results.....                     | 15  |
| References.....   | 17  |
| 小傳.....   | 19  |

# 超更新過程之研究

指導教授：黃文璋博士

國立高雄大學應用數學系

學生：賴佳鈴

國立高雄大學統計學研究所

## 摘 要

假設消費者每次消費皆可隨機地得到一張票券，古典的票券收集問題，所關心的是收集齊全所有票券共需之消費次數。對大多數的真實情況，消費者不見得每次消費都恰好得到一張票券。由古典的票券收集問題得到靈感，本文中，我們將研究一個所謂的超更新過程。令 $\{X_i, i \geq 1\}$ 為一組獨立且有相同分佈之隨機變數，並令 $S_n = \sum_{i=1}^n X_i, n \geq 1, S_0 = 0$ 。對每一個 $t \geq 0$ ，定義 $Q_t = \inf\{n \mid n \geq 0, S_n \geq t\}$ 。對古典的票券收集問題， $Q_t$ 表消費者到時間 $t$ 為止的總票券數大於或等於 $t$ 的最少消費次數， $t \geq 0$ 。我們首先比較過程 $\{Q_t, t \geq 0\}$ 和由同一組 $\{X_i, i \geq 1\}$ 所生成之更新過程 $\{N_t, t \geq 0\}$ 。接著給出過程 $\{Q_t, t \geq 0\}$ 的一些基本性質。最後也給出過程 $\{Q_t, t \geq 0\}$ 的一些極限及相關的結果。

**關鍵詞：**票券收集問題、幾何分佈、負二項分佈、更新過程、樣本路徑、超更新過程

# A Study of the Suprenewal Process

Advisor: Dr. Wen-Jang Huang  
Department of Applied Mathematics  
National University of Kaohsiung

Student: Chia-Ling Lai  
Institute of Statistics  
National University of Kaohsiung

## ABSTRACT

The classical coupon collector's problem is concerned with the number of purchases in order to have a complete collection, assuming that on each purchase a consumer can obtain a randomly chosen coupon. For most real situations, a consumer may not just get exactly one coupon on each purchase. Motivated by the classical coupon collector's problem, in this work, we will study the so-called suprenewal process. Let  $\{X_i, i \geq 1\}$  be a sequence of independent and identically distributed random variables,  $S_n = \sum_{i=1}^n X_i$ ,  $n \geq 1$ ,  $S_0 = 0$ . For every  $t \geq 0$ , define  $Q_t = \inf\{n \mid n \geq 0, S_n \geq t\}$ . For the classical coupon collector's problem,  $Q_t$  denotes the minimal number of purchases, such that the total number of coupons that the consumer has owned until time  $t$  is greater than or equal to  $t$ ,  $t \geq 0$ . First the process  $\{Q_t, t \geq 0\}$  and the renewal process  $\{N_t, t \geq 0\}$  generated by the same sequence  $\{X_i, i \geq 1\}$  are compared. Next some fundamental properties of  $\{Q_t, t \geq 0\}$  are provided. Finally limiting and some other related results are obtained for the process  $\{Q_t, t \geq 0\}$ .

**Keywords:** Coupon collector's problem, geometric distribution, negative binomial distribution, renewal process, sample path, suprenewal process.

## 1. Introduction

Starting from the end of April 2005, collecting Hello Kitty magnets became an immensely popular hobby in Taiwan. President Chain Stores Corp., which runs Taiwanese largest convenience store chain, 7-Eleven, was giving away one of a series of commemorative Hello Kitty magnets for each NTD77 a consumer spends at 7-Eleven store. There are 41 different patterns of Hello Kitty magnets in total. Because the cover of each package of magnet is the same, it is reasonable to assume that the magnets are given randomly.

We now review the classical coupon collector's problem. Assume there are  $N$  distinct coupons in a collection, and a series of random draws is made with replacement from these. Let  $T$  denote the number of draws necessary for all  $N$  coupons to have been drawn at least once. Properties of  $T$  had been studied by many authors, see e.g. Goodwin (1949) and Feller (1950). Among others, expectation and variance of  $T$  can be obtained as follows. For  $i \geq 1$ , let  $C_i \in \{1, 2, \dots, N\}$  be the coupon obtained at the  $i$ -th draw. The  $i$ -th draw is called a success, if  $C_i$  has not been obtained before the  $i$ -th draw. For  $1 \leq i \leq N$ , let  $T_i$  denote the number of draws after the  $(i-1)$ -th success, till the  $i$ -th success. Then  $T = \sum_{i=1}^N T_i$ . Obviously,  $T_1, T_2, \dots, T_N$  are independent, and  $T_i$  has a geometric distribution with parameter  $p_i = (N-i+1)/N$ , then  $E(T_i) = N/(N-i+1)$ , and  $\text{Var}(T_i) = (1 - (N/(N-i+1)))/(N/(N-i+1))^2$ ,  $1 \leq i \leq N$ . Thus

$$E(T) = \sum_{i=1}^N E(T_i) = NH_N, \quad (1)$$

where for  $N \geq 1$ ,  $H_N = \sum_{i=1}^N 1/i$  is the  $N$ -th Harmonic number, and

$$\text{Var}(T) = \sum_{i=1}^N \text{Var}(T_i) = N^2 \sum_{i=1}^N \frac{1}{i^2} - NH_N. \quad (2)$$

The above coupon collector's problem can be generalized. Assume the  $i$ -th coupon has probability  $p_i$  of being drawn, where  $0 < p_i < 1$ ,  $i = 1, 2, \dots, N$ , such that  $\sum_{i=1}^N p_i = 1$ , the  $p_i$ 's are allowed to be unequal. This was studied by von Schelling (1954). Some limiting results were derived by Baum and Billingsley (1965) and Hoslt (1971), and others. Related problems had also been discussed, such as the collector's brotherhood problem. As an example, Foata et al.(2001) and Foata and Zeilberger (2003) considered the situation that the collector shares his harvest with his brothers. They answered the question that when the collection of the collector is completed, the number of coupons each brother still lacks.

For our present problem, the expected number of magnets needed for collecting a complete set of 41 magnets is  $41 \sum_{i=1}^{41} 1/i \doteq 176.42$ . For a particular consumer, if his spending is less than NTD77, then he gets 0 magnet, if his spending is at least NTD77 and less than NTD154, then he gets 1 magnet, if his spending is at least NTD154 and less than NTD231, then he gets 2 magnets on that purchase, and so on. Now what is the number of purchases needed in order to get magnets greater than or equal to 176.42? To solve this problem, first we introduce a new process and study some of its properties.

Let  $\{X_i, i \geq 1\}$  be a sequence of independent and identically distributed (i.i.d.) random variables,  $S_n = \sum_{i=1}^n X_i$ ,  $n \geq 1$ ,  $S_0 = 0$ . For every  $t \geq 0$ , define  $Q_t = \inf\{n \mid n \geq 0, S_n \geq t\}$ . For the magnets problem,  $X_i$  can be viewed as the number of magnets received on the  $i$ -th purchase,  $i \geq 1$ , and  $Q_t$  can be viewed as the minimal number of purchases, such that the total number of magnets is greater than or equal to  $t$ ,  $t \geq 0$ .

Recall that the renewal process  $\{N_t, t \geq 0\}$  generated by the same sequence  $\{X_i, i \geq 1\}$ , where for  $t \geq 0$ ,  $N_t = \sup\{n \mid n \geq 0, S_n \leq t\}$ ,  $N_t$  denotes the number of renewals in  $[0, t]$ .  $Q_t$  can be referred to as the minimal number of renewals in  $[t, \infty)$ , and we call  $\{Q_t, t \geq 0\}$  the suprenewal process. In Section 2, we compare the suprenewal process  $\{Q_t, t \geq 0\}$  with the renewal process. In Section 3, some fundamental properties of  $\{Q_t, t \geq 0\}$  are studied. Finally, limiting and some other related results are presented in Section 4.

## 2. Comparisons of $\{Q_t, t \geq 0\}$ and $\{N_t, t \geq 0\}$

Let  $X_1, X_2, \dots$  be i.i.d. random variables with the same distribution as  $X$ , where  $X$ , a nonnegative random variable, has the distribution function  $F$  with  $F(0-) = 0$  and  $F(0) < 1$ . Let  $S_n = \sum_{i=1}^n X_i$ ,  $n \geq 1$ ,  $S_0 = 0$ . Let  $\{Q_t, t \geq 0\}$  and  $\{N_t, t \geq 0\}$  be the suprenewal process and renewal process generated by  $\{X_i, i \geq 1\}$  respectively. Obviously,  $Q_0 = 0$  and  $Q_t \geq 1$ , if  $t > 0$ . Also  $Q_t \leq n$  if and only if  $S_n \geq t$ . Hence for every  $t > 0$  and integer  $n \geq 1$ ,

$$\begin{aligned} P(Q_t = n) &= P(Q_t \leq n) - P(Q_t \leq n-1) \\ &= P(S_n \geq t) - P(S_{n-1} \geq t) \\ &= P(S_{n-1} < t) - P(S_n < t). \end{aligned} \quad (3)$$

If  $F$  is continuous, then  $S_n$  is a continuous random variable for every integer  $n \geq 0$ . Consequently,

$$P(Q_t = n) = F_{n-1}(t) - F_n(t), t > 0, n \geq 1, \quad (4)$$

where  $F_n$  is the  $n$ -fold convolution of  $F$  with itself,  $n \geq 1$ , and  $F_0(t) = 1$ ,  $t \geq 0$ .

We now compare the two processes  $\{Q_t, t \geq 0\}$  and  $\{N_t, t \geq 0\}$ . First instead of having right continuous sample paths for  $\{N_t, t \geq 0\}$ ,  $\{Q_t, t \geq 0\}$  has left continuous sample paths. Next instead of (3) and (4), whether  $F$  is continuous or not,

$$P(N_t = n) = P(S_n \leq t) - P(S_{n+1} \leq t) = F_n(t) - F_{n+1}(t), t \geq 0, n \geq 0. \quad (5)$$

On the other hand,  $\{Q_t, t \geq 0\}$  and  $\{N_t, t \geq 0\}$  have the same jump times. Denote the sequence of jump times by  $0 = \tau_0 < \tau_1 < \tau_2 < \dots$ . Then

$$N_t = Q_t - 1, \text{ if } t \notin \{\tau_0, \tau_1, \tau_2, \dots\}, \quad (6)$$

and

$$N_{\tau_i} = Q_{\tau_i} + Y_{\tau_i} - 1, i \geq 0, \quad (7)$$

where

$$\begin{aligned} Y_{\tau_0} &= 1, \\ Y_{\tau_i} &= Q_{\tau_i+} - Q_{\tau_i} = N_{\tau_i} - N_{\tau_i-}, \quad i \geq 1, \end{aligned} \quad (8)$$

denotes the common jump size at  $\tau_i$  of the processes  $\{Q_t, t \geq 0\}$  and  $\{N_t, t \geq 0\}$ . It can be seen that for  $i \geq 1$ ,  $Y_{\tau_i}$  has a  $\mathcal{G}e(\lambda)$  distribution, where  $\lambda = 1 - F(0)$ , if  $F(0) > 0$ ; and  $Y_{\tau_i} \equiv 1$ , hence  $Q_{\tau_i} = N_{\tau_i}$ ,  $i \geq 1$ , if  $F(0) = 0$ .  $Y_{\tau_i}$  and  $Q_{\tau_i}$  are independent, and  $Y_{\tau_i}$  and  $N_{\tau_i-1}$  are also independent. If  $F$  is continuous, then  $F(0) = 0$ , and

$$N_t = Q_t - 1 \text{ almost everywhere on } [0, \infty). \quad (9)$$

As an example, let  $F(x) = 1 - e^{-\lambda x}$ ,  $\lambda > 0$ ,  $x > 0$ , then it is well known that  $N_t$  has a  $\mathcal{P}(\lambda t)$  distribution,  $t > 0$ . By (9),  $Q_t - 1$  is also  $\mathcal{P}(\lambda t)$  distributed for almost all  $t$  on  $[0, \infty)$ . Note that except (6) and (7), we also have the following relationship

$$Q_{\tau_i} \leq N_{\tau_i} \leq Q_{\tau_{i+1}}, \quad i \geq 0. \quad (10)$$

Although  $Q_t$  may be less than  $N_t$ , from the definitions of  $Q_t$  and  $N_t$ , we have

$$S_{N_t} \leq t \leq S_{Q_t}, \quad t \geq 0. \quad (11)$$

In particular

$$S_{N_{\tau_i}} = S_{Q_{\tau_i}} = \tau_i, \quad i \geq 0. \quad (12)$$

Recall that  $S_{N_{t+1}} - t$ ,  $t - S_{N_t}$ , and  $X_{N_{t+1}} = S_{N_{t+1}} - S_{N_t}$  are called residual life at time  $t$ , current life at time  $t$ , and total life at time  $t$ , respectively, for the renewal process  $\{N_t, t \geq 0\}$ . It is known that  $P(X_{N_{t+1}} > x) \geq P(X_1 > x)$ ,  $x \geq 0$ , and  $E(X_{N_{t+1}}) \geq E(X_1)$ ,  $t \geq 0$ . This is the so-called inspection paradox. Similarly, it can be shown

$$P(X_{Q_t} > x) \geq P(X_1 > x), \quad x \geq 0. \quad (13)$$

That is  $X_{Q_t}$  is stochastically larger than  $X_1$ . Consequently,

$$E(X_{Q_t}) \geq E(X_1), \quad t \geq 0. \quad (14)$$

Furthermore, using the fact that a renewal process probabilistically starts over when a renewal occurs, for every increasing function  $g$ , the following inequality is immediate:

$$E(g(N_{t+s} - N_t)) \leq E(g(N_s + 1)), \quad t, s \geq 0. \quad (15)$$

Similarly, we have

$$E(g(Q_{t+s} - Q_t)) \leq E(g(Q_s)), \quad t, s \geq 0. \quad (16)$$

In particular

$$E(Q_{t+s} - Q_t) \leq E(Q_s), \quad t, s \geq 0. \quad (17)$$

We give a typical sample paths of  $\{Q_t, t \geq 0\}$  and  $\{N_t, t \geq 0\}$ , respectively, to illustrate the relationships (6), (7) and (10). Assume  $X_1 = 2, X_2 = 0, X_3 = 1, X_4 = 4, X_5 = 0, X_6 = 0, X_7 = 3, \dots$ , then  $S_1 = 2, S_2 = 2, S_3 = 3, S_4 = 7, S_5 = 7, S_6 = 7, S_7 = 10, \dots$ . Figure 1 gives the sample paths of  $\{Q_t, t \geq 0\}$  and  $\{N_t, t \geq 0\}$ .

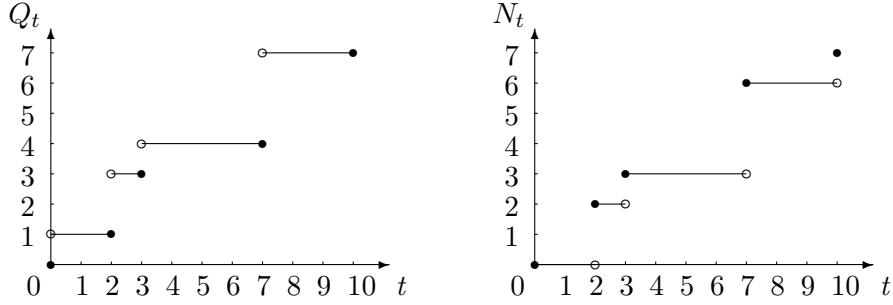


Figure 1. Sample paths of  $\{Q_t, t \geq 0\}$  and  $\{N_t, t \geq 0\}$

### 3. Some fundamental properties of $\{Q_t, t \geq 0\}$

There are many investigations for properties of renewal process in the literatures. In this section, we explore some basic properties of the process  $\{Q_t, t \geq 0\}$ , especially for the case that  $X$  takes on nonnegative integer values. Throughout this section, let  $P(X < \infty) = 1$ , and  $P(X = k) = p_k$ , where  $p_k \geq 0$ ,  $k = 0, 1, 2, \dots$ ,  $p_0 < 1$ , and  $\sum_{k=0}^{\infty} p_k = 1$ . Also let  $N = \sup\{i \mid i \geq 0, p_i > 0\}$ .

First we introduce some notation which will be used often in this work. Let  $\lceil t \rceil$  and  $\lfloor t \rfloor$  denote the ceiling function and the floor function, respectively, namely  $\lceil t \rceil$  = the least integer greater than or equal to  $t$ , and  $\lfloor t \rfloor$  = the greatest integer less than or equal to  $t$ . For example,  $\lceil 3.7 \rceil = 4$ ,  $\lfloor 3.7 \rfloor = 3$ , and  $\lceil 6 \rceil = \lfloor 6 \rfloor = 6$ . For integers  $a, b, c$ , with  $a \leq b \leq c$ , and nonnegative integers  $x_0, x_1, \dots, x_N$ , if  $N < \infty$ , let

$$\mathcal{A}_{a,b} = \{(x_0, x_1, \dots, x_N) \mid \sum_{i=0}^N x_i = a, \text{ and } \sum_{i=0}^N ix_i = b\},$$

$$\mathcal{B}_{a,b,c} = \{(x_1, x_2, \dots, x_N) \mid \sum_{i=1}^N x_i = a, \text{ and } \sum_{i=1}^N ix_i = b, b+1, \dots, c\},$$

and

$$\mathcal{C}_a = \{(x_1, x_2, \dots, x_N) \mid \sum_{i=1}^N x_i = a\};$$



if  $N = \infty$ , let

$$\mathcal{A}_{a,b}^1 = \{(x_0, x_1, \dots) \mid \sum_{i \geq 0} x_i = a, \text{ and } \sum_{i \geq 0} ix_i = b\},$$

and

$$\mathcal{B}_{a,b,c}^1 = \{(x_1, x_2, \dots) \mid \sum_{i \geq 1} x_i = a, \text{ and } \sum_{i \geq 1} ix_i = b, b+1, \dots, c\}.$$

Note that if  $(x_0, x_1, \dots, x_N) \in \mathcal{A}_{a+x_0,b}$ , then  $(x_1, x_2, \dots, x_N) \in \mathcal{B}_{a,b,b}$ ; if  $(x_0, x_1, \dots) \in \mathcal{A}_{a+x_0,b}^1$ , then  $(x_1, x_2, \dots) \in \mathcal{B}_{a,b,b}^1$ , and  $\mathcal{B}_{a,a,Na} = \mathcal{C}_a$ .

We give three simple examples in the following.

**Example 1.** Let  $p_1 = p_2 = p_3 = 1/3$ . Then the support of  $Q_{3.5}$  is  $\{2, 3, 4\}$ , and  $P(Q_{3.5} = 2) = 2/3$ ,  $P(Q_{3.5} = 3) = 8/27$ ,  $P(Q_{3.5} = 4) = 1/27$ .

**Example 2.** Let  $p_0, p_1, p_2 > 0$ , and  $p_0 + p_1 + p_2 = 1$ . Then  $P(Q_2 = i) = (i-1)p_0^{i-2}(1-p_0)p_1 + p_0^{i-1}p_2$ ,  $i \geq 1$ . In particular, if  $p_0 = 0.2, p_1 = 0.3, p_2 = 0.5$ , then  $P(Q_2 = 1) = 0.5$ ,  $P(Q_2 = 2) = 0.34$ ,  $P(Q_2 = 3) = 0.116$ ,  $P(Q_2 = 4) = 0.0328, \dots$ .

The next example shows that the family of the distribution  $Q_t$  contains geometric and negative binomial, the two common statistical distributions.

**Example 3.** Assume  $0 < p_0 < 1$  and  $p_1 = 1 - p_0$ . In this case  $\tau_i = i$ ,  $i \geq 1$ . That is  $X_1, X_2, \dots$  are i.i.d.  $\mathcal{Ber}(p_1)$  random variables, and  $S_n$  is  $\mathcal{B}(n, p_1)$  distributed,  $n \geq 1$ . Then obviously for every  $t > 0$ ,  $S_{Q_t} = \lceil t \rceil$ ,  $Q_t \sim \mathcal{NB}(\lceil t \rceil, p_1)$ , and for every integer  $k \geq 2$ ,  $Q_1, Q_2 - Q_1, \dots, Q_k - Q_{k-1}$ , are i.i.d. random variables with the common  $\mathcal{Ge}(p_1)$  distribution. Consequently, for every  $t > 0$ ,  $E(Q_t) = \lceil t \rceil / p_1$ ,  $E(S_{Q_t} - t) = \lceil t \rceil - t$ , and  $E(t - S_{Q_t-1}) = t - \lceil t \rceil + 1$ . Hence  $E(X_{Q_t}) = E(S_{Q_t} - S_{Q_t-1}) = 1 > p_1 = E(X_1)$ . Moreover, it can be seen easily, for the above  $\{X_i, i \geq 1\}$ , for any  $0 < p_0 < 1$ , there is an infinite number of positive  $t$ 's, such that  $E(S_{Q_t} - t) > E(X_1)$ .

**Remark 1.** As a comparison, for the  $\{X_i, i \geq 1\}$  defined in Example 3, we have  $S_{N_t} = \lfloor t \rfloor$ ,

$$P(N_t = n) = \binom{n}{\lfloor t \rfloor} p_1^{\lfloor t \rfloor + 1} p_0^{n - \lfloor t \rfloor}, \quad n \geq \lfloor t \rfloor,$$

and  $P(N_t = n) = 0$ , for  $n < \lfloor t \rfloor$ . That is  $N_t + 1 \sim \mathcal{NB}(\lfloor t \rfloor + 1, p_1)$ ,  $t > 0$ . This also can be seen by (6), (7) and Example 3. Now  $E(N_t) = (\lfloor t \rfloor + p_0)/p_1$ ,  $E(t - S_{N_t}) = t - \lfloor t \rfloor$ ,  $E(S_{N_t+1} - t) = \lfloor t \rfloor + 1 - t$ , and  $E(X_{N_t+1}) = 1 > E(X_1)$ ,  $t > 0$ .

Next we find the distribution of  $Q_t$ ,  $t > 0$ .

**Theorem 1.** For every integer  $n \geq 1$  and  $t > 0$ ,

$$P(Q_t = n) = \begin{cases} g_{n-1,t} - g_{n,t} & , p_0 = 0, \\ \sum_{m=0}^{\lceil t \rceil - 1} g_{m,t} p_0^{n-m-1} \left( \binom{n-1}{m} - \binom{n}{m} p_0 \right) & , 0 < p_0 < 1, \end{cases} \quad (18)$$

where if  $N < \infty$ ,

$$g_{m,t} = \begin{cases} (1 - p_0)^m & , 0 \leq m \leq \lfloor \frac{\lceil t \rceil - 1}{N} \rfloor, \\ \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{m,m, \lceil t \rceil - 1}} m! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) & , m \geq \lfloor \frac{\lceil t \rceil - 1}{N} \rfloor + 1, \end{cases} \quad (19)$$

if  $N = \infty$ ,

$$g_{m,t} = \begin{cases} 1 & , m = 0, \\ \sum_{(x_1, x_2, \dots) \in \mathcal{B}_{m,m, \lceil t \rceil - 1}^1} m! \left( \prod_{i \geq 1} \frac{p_i^{x_i}}{x_i!} \right) & , m \geq 1. \end{cases} \quad (20)$$

**Proof.** Obviously we only need to prove (18) holds for positive integer  $t$ . We prove this by induction. (i) First we prove the case  $p_0 = 0$  and  $N < \infty$ . That (18) holds for  $t = 1$  can be seen as following. From assumptions, we have  $P(S_n < 1) = 0$ ,  $n \geq 1$ ,  $P(S_0 < 1) = 1$ . Thus

$$\begin{aligned} P(Q_1 = n) &= P(S_{n-1} < 1) - P(S_n < 1) \\ &= \begin{cases} 1 & , n = 1, \\ 0 & , n \geq 2, \end{cases} \\ &= g_{n-1,1} - g_{n,1}, n \geq 1, \end{aligned}$$

where the last equality holds is because for every  $n \geq 1$ ,  $B_{n,n,0}$  is a null set, hence  $g_{n,1} = 0$ . This together with  $g_{0,1} = 1$  implies  $g_{0,1} - g_{1,1} = 1$ , and  $g_{n-1,1} - g_{n,1} = 0$ ,  $n \geq 2$ . Now suppose (18) is true for  $t = r \geq 1$ , i.e. we have

$$P(Q_r = n) = g_{n-1,r} - g_{n,r}.$$

Then

$$\begin{aligned} P(Q_{r+1} = n) &= P(S_{n-1} < r+1) - P(S_n < r+1) \\ &= [P(S_{n-1} < r) - P(S_n < r)] + [P(S_{n-1} = r) - P(S_n = r)] \\ &= [g_{n-1,r} - g_{n,r}] + [P(S_{n-1} = r) - P(S_n = r)] \\ &= \left\{ \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{n-1, n-1, r-1}} (n-1)! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) - \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{n, n, r-1}} n! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) \right\} \end{aligned}$$

$$\begin{aligned}
& + \left\{ \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{n-1, r, r}} (n-1)! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) - \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{n, r, r}} n! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) \right\} \\
& = \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{n-1, n-1, r}} (n-1)! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) - \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{n, n, r}} n! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) \\
& = g_{n-1, r+1} - g_{n, r+1}.
\end{aligned}$$

This proves (18) holds for  $t = r + 1$ . By the induction argument this completes the proof for the case  $p_0 = 0$  and  $N < \infty$ . The proof of (18) for the case  $p_0 = 0$  and  $N = \infty$  is similar to the proof for the case  $p_0 = 0$  and  $N < \infty$ , hence is omitted.

(ii) Next we prove the case  $0 < p_0 < 1$  and  $N < \infty$ . The proof of (18) for  $t = 1$  is as following.

$$\begin{aligned}
P(Q_1 = n) &= P(S_{n-1} < 1) - P(S_n < 1) = P(S_{n-1} = 0) - P(S_n = 0) \\
&= p_0^{n-1} - p_0^n = g_{0,1} p_0^{n-1} \left( \binom{n-1}{0} - \binom{n}{0} p_0 \right).
\end{aligned}$$

Now suppose (18) is true for  $t = r \geq 1$ , i.e. we have

$$P(Q_r = n) = \sum_{m=0}^{r-1} g_{m,r} p_0^{n-m-1} \left( \binom{n-1}{m} - \binom{n}{m} p_0 \right). \quad (21)$$

Then

$$\begin{aligned}
P(Q_{r+1} = n) &= P(S_{n-1} < r+1) - P(S_n < r+1) \\
&= [P(S_{n-1} < r) - P(S_n < r)] + [P(S_{n-1} = r) - P(S_n = r)] \\
&= P(Q_r = n) + [P(S_{n-1} = r) - P(S_n = r)],
\end{aligned} \quad (22)$$

and

$$\begin{aligned}
P(S_{n-1} = r) &= \sum_{(x_0, x_1, \dots, x_N) \in \mathcal{A}_{n-1, r}} (n-1)! \left( \prod_{i=0}^N \frac{p_i^{x_i}}{x_i!} \right) \\
&= \left( \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{\lfloor \frac{r-1}{N} \rfloor + 1, r, r}} \frac{(n-1)!}{(n - \lfloor \frac{r-1}{N} \rfloor - 2)!} p_0^{n - \lfloor \frac{r-1}{N} \rfloor - 2} \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) \right) + \dots \\
&\quad + \left( \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{r, r, r}} \frac{(n-1)!}{(n-r-1)!} p_0^{n-r-1} \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) \right) \\
&= \binom{n-1}{\lfloor \frac{r-1}{N} \rfloor + 1} \left\{ \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{\lfloor \frac{r-1}{N} \rfloor + 1, r, r}} \left( \lfloor \frac{r-1}{N} \rfloor + 1 \right)! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) \right\} p_0^{n - \lfloor \frac{r-1}{N} \rfloor - 2} + \dots \\
&\quad + \binom{n-1}{r} \left\{ \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{r, r, r}} r! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) \right\} p_0^{n-r-1}.
\end{aligned} \quad (23)$$

That the second equality of (23) holds is because if  $(x_0, x_1, \dots, x_N) \in \mathcal{A}_{n-1, r}$ , i.e.  $\sum_{i=1}^N x_i = n-1-x_0$  and  $\sum_{i=1}^N ix_i = r$ , then  $\sum_{i=1}^N x_i \leq r$ , hence  $x_0 \geq n-r-1$ . On the other hand, if  $x_0 \geq n - \lfloor (r-1)/N \rfloor - 1$ , then  $\sum_{i=1}^N x_i \leq \lfloor (r-1)/N \rfloor$ , and  $\sum_{i=1}^N ix_i \leq r-1$  follows. This proves  $n-r-1 \leq x_0 \leq n - \lfloor (r-1)/N \rfloor - 2$ . Hence  $\sum_{i=1}^N x_i = \lfloor (r-1)/N \rfloor + 1, \dots, r$ . This together with  $\sum_{i=1}^N ix_i = r$  implies  $(x_1, x_2, \dots, x_N) \in \mathcal{B}_{\lfloor (r-1)/N \rfloor + 1, r, r}, \dots, \mathcal{B}_{r, r, r}$ . Similarly,

$$\begin{aligned} P(S_n = r) &= \binom{n}{\lfloor \frac{r-1}{N} \rfloor + 1} \left\{ \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{\lfloor \frac{r-1}{N} \rfloor + 1, r, r}} (\lfloor \frac{r-1}{N} \rfloor + 1)! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) p_0^{n - \lfloor \frac{r-1}{N} \rfloor - 1} + \dots \right. \\ &\quad \left. + \binom{n}{r} \left\{ \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{r, r, r}} r! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) p_0^{n-r} \right\} \right\}. \end{aligned} \quad (24)$$

Substituting (21), (23) and (24) into (22), it yields

$$\begin{aligned} P(Q_{r+1} = n) &= \left\{ \binom{n-1}{0} p_0^{n-1} + \binom{n-1}{1} (1-p_0) p_0^{n-2} + \dots + \binom{n-1}{\lfloor \frac{r-1}{N} \rfloor} (1-p_0)^{\lfloor \frac{r-1}{N} \rfloor} \right. \\ &\quad \cdot p_0^{n-1 - \lfloor \frac{r-1}{N} \rfloor} + \binom{n-1}{\lfloor \frac{r-1}{N} \rfloor + 1} \left( \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{\lfloor \frac{r-1}{N} \rfloor + 1, \lfloor \frac{r-1}{N} \rfloor + 1, r}} (\lfloor \frac{r-1}{N} \rfloor + 1)! \right. \\ &\quad \cdot \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) p_0^{n - \lfloor \frac{r-1}{N} \rfloor - 2} + \dots + \binom{n-1}{r-1} \left( \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{r-1, r-1, r}} (r-1)! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) \right. \\ &\quad \cdot p_0^{n-r} + \binom{n-1}{r} \left( \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{r, r, r}} r! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) p_0^{n-r-1} \right) \left. \right\} - \left\{ \binom{n}{0} p_0^n \right. \\ &\quad + \binom{n}{1} (1-p_0) p_0^{n-1} + \dots + \binom{n}{\lfloor \frac{r-1}{N} \rfloor} (1-p_0)^{\lfloor \frac{r-1}{N} \rfloor} p_0^{n - \lfloor \frac{r-1}{N} \rfloor} \\ &\quad + \binom{n}{\lfloor \frac{r-1}{N} \rfloor + 1} \left( \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{\lfloor \frac{r-1}{N} \rfloor + 1, \lfloor \frac{r-1}{N} \rfloor + 1, r}} (\lfloor \frac{r-1}{N} \rfloor + 1)! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) p_0^{n - \lfloor \frac{r-1}{N} \rfloor - 1} \right. \\ &\quad + \dots + \binom{n}{r-1} \left( \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{r-1, r-1, r}} (r-1)! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) p_0^{n-r+1} \right. \\ &\quad + \binom{n}{r} \left( \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{r, r, r}} r! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) p_0^{n-r} \right) \left. \right\} \\ &= \sum_{m=0}^{\lfloor \frac{r-1}{N} \rfloor} (1-p_0)^m p_0^{n-m-1} \left( \binom{n-1}{m} - \binom{n}{m} p_0 \right) \\ &\quad + \sum_{m=\lfloor \frac{r-1}{N} \rfloor + 1}^r \left( \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{A}_{m, m, r}} r! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) p_0^{n-m-1} \left( \binom{n-1}{m} - \binom{n}{m} p_0 \right) \right) \end{aligned}$$

$$\begin{aligned}
&= \sum_{m=0}^{\lfloor \frac{r}{N} \rfloor} (1-p_0)^m p_0^{n-m-1} \left( \binom{n-1}{m} - \binom{n}{m} p_0 \right) \\
&\quad + \sum_{m=\lfloor \frac{r}{N} \rfloor + 1}^r \left( \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{m, m, r}} m! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) p_0^{n-m-1} \left( \binom{n-1}{m} - \binom{n}{m} p_0 \right) \right) \\
&= \sum_{m=0}^r g_{m, r+1} p_0^{n-m-1} \left( \binom{n-1}{m} - \binom{n}{m} p_0 \right), \tag{25}
\end{aligned}$$

where the third equality holds is because if  $\sum_{i=1}^N X_i = m$ , then  $m \leq \sum_{i=1}^N iX_i \leq Nm$ , hence if  $r \geq Nm$ , i.e.  $m \leq \lfloor r/N \rfloor$ , then  $\mathcal{B}_{m, m, r} = \mathcal{B}_{m, m, Nm}$  and

$$\begin{aligned}
\sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{m, m, r}} m! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) &= \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{B}_{m, m, Nm}} m! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) \\
&= \sum_{(x_1, x_2, \dots, x_N) \in \mathcal{C}_m} m! \left( \prod_{i=1}^N \frac{p_i^{x_i}}{x_i!} \right) \\
&= (p_1 + p_2 + \dots + p_N)^m \\
&= (1 - p_0)^m.
\end{aligned}$$

This proves (18) holds for  $t = r + 1$ , and the proof for the case  $0 < p_0 < 1$  and  $N < \infty$  is completed.

(iii) Finally we consider the case  $0 < p_0 < 1$  and  $N = \infty$ . The proof of (18) for  $t = 1$  is the same as in (ii). Now suppose the induction statement is true for  $t = r \geq 1$ , i.e. we have

$$P(Q_r = n) = \sum_{m=0}^{r-1} g_{m, r} p_0^{n-m-1} \left( \binom{n-1}{m} - \binom{n}{m} p_0 \right). \tag{26}$$

Then

$$\begin{aligned}
P(Q_{r+1} = n) &= P(S_{n-1} < r+1) - P(S_n < r+1) \\
&= P(Q_r = n) + [P(S_{n-1} = r) - P(S_n = r)], \tag{27}
\end{aligned}$$

and

$$\begin{aligned}
P(S_{n-1} = r) &= \sum_{(x_0, x_1, \dots) \in \mathcal{A}_{n-1, r}^1} (n-1)! \left( \prod_{i \geq 0} \frac{p_i^{x_i}}{x_i!} \right) \\
&= \left( \sum_{(x_1, x_2, \dots) \in \mathcal{B}_{1, r, r}^1} \frac{(n-1)!}{(n-2)!} p_0^{n-2} \left( \prod_{i \geq 1} \frac{p_i^{x_i}}{x_i!} \right) \right) + \dots \\
&\quad + \left( \sum_{(x_1, x_2, \dots) \in \mathcal{B}_{r, r, r}^1} \frac{(n-1)!}{(n-r-1)!} p_0^{n-r-1} \left( \prod_{i \geq 1} \frac{p_i^{x_i}}{x_i!} \right) \right)
\end{aligned}$$

$$\begin{aligned}
&= \binom{n-1}{1} \left\{ \sum_{(x_1, x_2, \dots) \in \mathcal{B}_{1,r,r}^1} \left( \prod_{i \geq 1} \frac{p_i^{x_i}}{x_i!} \right) \right\} p_0^{n-2} + \dots \\
&\quad + \binom{n-1}{r} \left\{ \sum_{(x_1, x_2, \dots) \in \mathcal{B}_{r,r,r}^1} r! \left( \prod_{i \geq 1} \frac{p_i^{x_i}}{x_i!} \right) \right\} p_0^{n-r-1}, \tag{28}
\end{aligned}$$

where the second equality holds is by using the same argument as in the discussion of the paragraph after (23). Similarly,

$$\begin{aligned}
P(S_n = r) &= \binom{n}{1} \left\{ \sum_{(x_1, x_2, \dots) \in \mathcal{B}_{1,r,r}^1} \left( \prod_{i \geq 1} \frac{p_i^{x_i}}{x_i!} \right) \right\} p_0^{n-1} + \dots \\
&\quad + \binom{n}{r} \left\{ \sum_{(x_1, x_2, \dots) \in \mathcal{B}_{r,r,r}^1} r! \left( \prod_{i \geq 1} \frac{p_i^{x_i}}{x_i!} \right) \right\} p_0^{n-r}. \tag{29}
\end{aligned}$$

Substituting (26), (28) and (29) into (27), it yields

$$\begin{aligned}
P(Q_{r+1} = n) &= \left\{ \binom{n-1}{0} p_0^{n-1} + \binom{n-1}{1} \left( \sum_{(x_1, x_2, \dots) \in \mathcal{B}_{1,1,r}^1} \left( \prod_{i \geq 1} \frac{p_i^{x_i}}{x_i!} \right) \right) p_0^{n-2} + \dots \right. \\
&\quad + \binom{n-1}{r} \left( \sum_{(x_1, x_2, \dots) \in \mathcal{B}_{r,r,r}^1} r! \left( \prod_{i \geq 1} \frac{p_i^{x_i}}{x_i!} \right) \right) p_0^{n-r-1} \Big\} \\
&\quad - \left\{ \binom{n}{0} p_0^n + \binom{n}{1} \left( \sum_{(x_1, x_2, \dots) \in \mathcal{B}_{1,1,r}^1} \left( \prod_{i \geq 1} \frac{p_i^{x_i}}{x_i!} \right) \right) p_0^{n-1} + \dots \right. \\
&\quad + \binom{n}{r} \left( \sum_{(x_1, x_2, \dots) \in \mathcal{B}_{r,r,r}^1} r! \left( \prod_{i \geq 1} \frac{p_i^{x_i}}{x_i!} \right) \right) p_0^{n-r} \Big\} \\
&= p_0^{n-1} \left( \binom{n-1}{0} - \binom{n}{0} p_0 \right) \\
&\quad + \sum_{m=1}^r p_0^{n-m-1} \left( \sum_{(x_1, x_2, \dots) \in \mathcal{B}_{m,m,r}^1} m! \left( \prod_{i \geq 1} \frac{p_i^{x_i}}{x_i!} \right) \right) \left( \binom{n-1}{m} - \binom{n}{m} p_0 \right) \\
&= \sum_{m=0}^r g_{m,r+1} p_0^{n-m-1} \left( \binom{n-1}{m} - \binom{n}{m} p_0 \right). \tag{30}
\end{aligned}$$

This proves (18) holds for  $t = r + 1$ . The proof is completed.

Let

$$\phi_t(s) = E(e^{-sQ_t}) = \sum_{n=1}^{\infty} P(Q_t = n) e^{-sn}, \quad s \geq 0,$$

be the Laplace transform of  $Q_t$ ,  $t \geq 0$ .  $\phi_t(s)$  and the moments of  $Q_t$  can be obtained immediately by using Theorem 1. We give the results in the following corollary.

**Corollary 1.** Let integer  $n \geq 1$ ,  $p_0 \geq 0$ , and  $t > 0$ .

(i).

$$\phi_t(s) = 1 - \frac{1 - e^{-s}}{1 - p_0 e^{-s}} \sum_{m=0}^{\lceil t \rceil - 1} g_{m,t} \left( \frac{e^{-s}}{1 - p_0 e^{-s}} \right)^m, \quad s \geq 0. \quad (31)$$

(ii).

$$E(Q_t) = \sum_{m=0}^{\lceil t \rceil - 1} \frac{g_{m,t}}{(1 - p_0)^{m+1}}. \quad (32)$$

(iii).

$$\text{Var}(Q_t) = \sum_{m=0}^{\lceil t \rceil - 1} (1 + 2m + p_0) \frac{g_{m,t}}{(1 - p_0)^{m+2}} - \left\{ \sum_{m=0}^{\lceil t \rceil - 1} \frac{g_{m,t}}{(1 - p_0)^{m+1}} \right\}^2, \quad (33)$$

where  $g_{m,t}$ ,  $m \geq 0$ , are as defined in (19) and (20).

**Proof.** (i). Due to the expression (18), the proof of (31) is divided into two parts:  $0 < p_0 < 1$ , and  $p_0 = 0$ . First we prove the case  $0 < p_0 < 1$ .

$$\begin{aligned} \phi_t(s) &= \sum_{n=1}^{\infty} \left\{ \sum_{m=0}^{\lceil t \rceil - 1} g_{m,t} p_0^{n-m-1} \left( \binom{n-1}{m} - \binom{n}{m} p_0 \right) \right\} e^{-sn} \\ &= \sum_{n=1}^{\infty} \sum_{m=0}^{\lceil t \rceil - 1} \binom{n-1}{m} g_{m,t} p_0^{n-m-1} e^{-sn} - \sum_{n=1}^{\infty} \sum_{m=0}^{\lceil t \rceil - 1} \binom{n}{m} g_{m,t} p_0^{n-m} e^{-sn} \\ &= \sum_{m=0}^{\lceil t \rceil - 1} \frac{g_{m,t}}{p_0^{m+1}} \left\{ \sum_{n=1}^{\infty} \binom{n-1}{m} (p_0 e^{-s})^n \right\} - \sum_{m=0}^{\lceil t \rceil - 1} \frac{g_{m,t}}{p_0^m} \left\{ \sum_{n=1}^{\infty} \binom{n}{m} (p_0 e^{-s})^n \right\} \\ &= A - B. \end{aligned} \quad (34)$$

Now

$$\begin{aligned} A &= \frac{g_{0,t}}{p_0} \left\{ \sum_{n=1}^{\infty} \binom{n-1}{0} (p_0 e^{-s})^n \right\} + \sum_{m=1}^{\lceil t \rceil - 1} \frac{g_{m,t}}{p_0^{m+1}} \left\{ \sum_{n=1}^{\infty} \binom{n-1}{m} (p_0 e^{-s})^n \right\} \\ &= \frac{g_{0,t} e^{-s}}{1 - p_0 e^{-s}} + \sum_{m=1}^{\lceil t \rceil - 1} \frac{g_{m,t}}{p_0^{m+1}} \left\{ \sum_{n=1}^{\infty} \frac{n-m}{m} \binom{n-1}{n-m} (p_0 e^{-s})^n \right\} \\ &= \frac{g_{0,t} e^{-s}}{1 - p_0 e^{-s}} + \sum_{m=1}^{\lceil t \rceil - 1} \frac{g_{m,t}}{p_0^{m+1}} \left\{ \sum_{n=1}^{\infty} \left( \frac{n}{m} - 1 \right) \binom{n-1}{n-m} (1 - p_0 e^{-s})^m (p_0 e^{-s})^{n-m} \left( \frac{p_0 e^{-s}}{1 - p_0 e^{-s}} \right)^m \right\} \\ &= \frac{g_{0,t} e^{-s}}{1 - p_0 e^{-s}} + \sum_{m=1}^{\lceil t \rceil - 1} \frac{g_{m,t}}{p_0^{m+1}} \left( \frac{1}{1 - p_0 e^{-s}} - 1 \right) \left( \frac{p_0 e^{-s}}{1 - p_0 e^{-s}} \right)^m \end{aligned}$$

$$\begin{aligned}
&= \frac{g_{0,t}e^{-s}}{1-p_0e^{-s}} + \sum_{m=1}^{\lceil t \rceil - 1} g_{m,t} \left( \frac{e^{-s}}{1-p_0e^{-s}} \right)^{m+1} \\
&= \sum_{m=0}^{\lceil t \rceil - 1} g_{m,t} \left( \frac{e^{-s}}{1-p_0e^{-s}} \right)^{m+1},
\end{aligned} \tag{35}$$

and

$$\begin{aligned}
B &= g_{0,t} \left\{ \sum_{n=1}^{\infty} \binom{n}{0} (p_0e^{-s})^n \right\} + \sum_{m=1}^{\lceil t \rceil - 1} \frac{g_{m,t}}{p_0^m} \left\{ \sum_{n=1}^{\infty} \binom{n}{m} (p_0e^{-s})^n \right\} \\
&= g_{0,t} \frac{p_0e^{-s}}{1-p_0e^{-s}} + \sum_{m=1}^{\lceil t \rceil - 1} \frac{g_{m,t}}{p_0^m} \left\{ \sum_{n=1}^{\infty} \binom{n}{m} \binom{n-1}{m-1} (p_0e^{-s})^n \right\} \\
&= g_{0,t} \frac{p_0e^{-s}}{1-p_0e^{-s}} + \sum_{m=1}^{\lceil t \rceil - 1} \frac{g_{m,t}}{p_0^m} \left\{ \sum_{n=1}^{\infty} \binom{n}{m} \binom{n-1}{m-1} (1-p_0e^{-s})^m (p_0e^{-s})^{n-m} \right\} \left( \frac{p_0e^{-s}}{1-p_0e^{-s}} \right)^m \\
&= \left( -g_{0,t} + \frac{g_{0,t}}{1-p_0e^{-s}} \right) + \sum_{m=1}^{\lceil t \rceil - 1} \frac{g_{m,t}}{p_0^m} \left( \frac{1}{1-p_0e^{-s}} \right) \left( \frac{p_0e^{-s}}{1-p_0e^{-s}} \right)^m \\
&= -g_{0,t} + \sum_{m=0}^{\lceil t \rceil - 1} \frac{g_{m,t}}{1-p_0e^{-s}} \left( \frac{e^{-s}}{1-p_0e^{-s}} \right)^m.
\end{aligned} \tag{36}$$

The assertion follows by substituting  $A$  and  $B$  in (35) and (36) into (34), and noting that  $g_{0,t} = 1$ .

The proof for the case  $p_0 = 0$  is given below.

$$\begin{aligned}
\phi_t(s) &= \sum_{n=1}^{\infty} (g_{n-1,t} - g_{n,t}) e^{-sn} \\
&= \sum_{m=0}^{\infty} g_{m,t} e^{-s(m+1)} - \left( \sum_{m=0}^{\infty} g_{m,t} e^{-sm} - g_{0,t} \right) \\
&= 1 - (1 - e^{-s}) \sum_{m=0}^{\lceil t \rceil - 1} g_{m,t} e^{-sm}, \quad s \geq 0.
\end{aligned} \tag{37}$$

(ii). Taking the derivative of  $\phi_t$  with respect to  $s$ , we obtain for  $s > 0$ ,

$$\begin{aligned}
\phi'_t(s) &= -\frac{e^{-s}(1-p_0e^{-s}) - (1-e^{-s})p_0e^{-s}}{(1-p_0e^{-s})^2} \sum_{m=0}^{\lceil t \rceil - 1} g_{m,t} \left( \frac{e^{-s}}{1-p_0e^{-s}} \right)^m \\
&\quad - \frac{1-e^{-s}}{1-p_0e^{-s}} \sum_{m=0}^{\lceil t \rceil - 1} m \left( \frac{e^{-s}}{1-p_0e^{-s}} \right)^{m-1} g_{m,t} \left\{ \frac{-e^{-s}(1-p_0e^{-s}) - e^{-s}(p_0e^{-s})}{(1-p_0e^{-s})^2} \right\} \\
&= -\frac{(1-p_0)e^{-s}}{(1-p_0e^{-s})^2} \sum_{m=0}^{\lceil t \rceil - 1} g_{m,t} \left( \frac{e^{-s}}{1-p_0e^{-s}} \right)^m + \frac{e^{-s}(1-e^{-s})}{(1-p_0e^{-s})^3} \sum_{m=0}^{\lceil t \rceil - 1} m \left( \frac{e^{-s}}{1-p_0e^{-s}} \right)^{m-1} g_{m,t}.
\end{aligned}$$



Hence

$$E(Q_t) = -\lim_{s \downarrow 0} \phi'_t(s) = \sum_{m=0}^{\lceil t \rceil - 1} \frac{g_{m,t}}{(1-p_0)^{m+1}}.$$

(iii). As in (ii),

$$E(Q_t^2) = \lim_{s \downarrow 0} \phi''_t(s) = \sum_{m=0}^{\lceil t \rceil - 1} (1 + 2m + p_0) \frac{g_{m,t}}{(1-p_0)^{m+2}},$$

hence

$$\begin{aligned} \text{Var}(Q_t) &= E(Q_t^2) - (E(Q_t))^2 \\ &= \sum_{m=0}^{\lceil t \rceil - 1} (1 + 2m + p_0) \frac{g_{m,t}}{(1-p_0)^{m+2}} - \left\{ \sum_{m=0}^{\lceil t \rceil - 1} \frac{g_{m,t}}{(1-p_0)^{m+1}} \right\}^2 \end{aligned}$$

as required.

**Example 3.(Continued)** We use Theorem 1 and (i) of Corollary 1, respectively, to demonstrate  $Q_t \sim \mathcal{NB}(\lceil t \rceil, p_1)$ ,  $t > 0$ .

By letting  $N = 1$  in Theorem 1, it yields

$$g_{m,t} = \begin{cases} p_1^m & , 0 \leq m \leq \lceil t \rceil - 1, \\ 0 & , m \geq \lceil t \rceil. \end{cases} \quad (38)$$

Hence

$$\begin{aligned} P(Q_t = n) &= \sum_{m=0}^{\lceil t \rceil - 1} g_{m,t} p_0^{n-m-1} \left( \binom{n-1}{m} - \binom{n}{m} p_0 \right) \\ &= \sum_{m=0}^{\lceil t \rceil - 1} p_1^m p_0^{n-m-1} \left\{ \binom{n-1}{m} - \left( \binom{n-1}{m} + \binom{n-1}{m-1} \right) p_0 \right\} \\ &= \sum_{m=0}^{\lceil t \rceil - 1} \binom{n-1}{m} p_1^{m+1} p_0^{n-m-1} - \sum_{m=0}^{\lceil t \rceil - 2} \binom{n-1}{m} p_1^{m+1} p_0^{n-m-1} \\ &= \binom{n-1}{\lceil t \rceil - 1} p_1^{\lceil t \rceil} p_0^{n-\lceil t \rceil}, \quad n \geq \lceil t \rceil, \end{aligned}$$

where  $\binom{n-1}{-1}$  is defined to be 0. This shows  $Q_t \sim \mathcal{NB}(\lceil t \rceil, p_1)$ ,  $t > 0$ .

Next from (i) of Corollary 1 and (38), we have

$$\begin{aligned} \phi_t(s) &= 1 - \frac{1 - e^{-s}}{1 - p_0 e^{-s}} \sum_{m=0}^{\lceil t \rceil - 1} g_{m,t} \left( \frac{e^{-s}}{1 - p_0 e^{-s}} \right)^m = 1 - \frac{1 - e^{-s}}{1 - p_0 e^{-s}} \sum_{m=0}^{\lceil t \rceil - 1} \left( \frac{p_1 e^{-s}}{1 - p_0 e^{-s}} \right)^m \\ &= 1 - \frac{1 - e^{-s}}{1 - p_0 e^{-s}} \left\{ \frac{1 - \left( \frac{p_1 e^{-s}}{1 - p_0 e^{-s}} \right)^{\lceil t \rceil}}{\frac{1 - e^{-s}}{1 - p_0 e^{-s}}} \right\} = \left( \frac{p_1 e^{-s}}{1 - p_0 e^{-s}} \right)^{\lceil t \rceil}, \quad s \geq 0, \end{aligned}$$

which is exactly the Laplace transform of a  $\mathcal{NB}(\lceil t \rceil, p_1)$  distributed random variable.

Note that for integers  $a, b$ , with  $a \leq b$ ,  $\mathcal{B}_{a,a,b}^1 = \bigcup_{i=0}^{b-a} \mathcal{B}_{a,a+i,a+i}^1$ , where  $\mathcal{B}_{a,a,a}^1, \mathcal{B}_{a,a+1,a+1}^1, \dots, \mathcal{B}_{a,b,b}^1$ , are disjoint. Also let  $H_k^n = \binom{n-1+k}{k}$ ,  $n \geq 1$ ,  $k \geq 0$ , denote the number of  $k$ -combinations with repetition of  $n$  distinct things. Before giving Example 4, we need the following lemma.

**Lemma 1.** For integers  $n \geq 1$ , and  $k \geq 0$ , we have

$$\sum_{(x_1, x_2, \dots) \in \mathcal{B}_{n, n+k, n+k}^1} \frac{n!}{\prod_{i \geq 1} x_i!} = \binom{n-1+k}{k}. \quad (39)$$

**Proof.** For  $n \geq 1$ , let  $y_1, y_2, \dots, y_n$  be any  $n$  positive integers. For every  $i \geq 1$ , let  $z_i = \sum_{j=1}^n I_{\{y_j=i\}}$ , where

$$I_{\{y_j=i\}} = \begin{cases} 1 & , y_j = i, \\ 0 & , y_j \neq i, \end{cases}$$

is the indicator function. Then  $\sum_{i \geq 1} z_i = n$ , and  $\sum_{i \geq 1} i z_i = \sum_{j=1}^n y_j$ . Conversely, for every  $(z_1, z_2, \dots) \in \mathcal{B}_{n, n+k, n+k}^1$ , i.e.  $\sum_{i \geq 1} z_i = n$ , and  $\sum_{i \geq 1} i z_i = n+k$ ,  $k \geq 0$ , there exists exactly one multiset  $\{y_1, y_2, \dots, y_n\}$  satisfying  $\sum_{j=1}^n y_j = n+k$ . Hence for every  $(z_1, z_2, \dots) \in \mathcal{B}_{n, n+k, n+k}^1$ ,  $(n! / \prod_{i \geq 1} z_i!)$  is the number of distinct permutations of the corresponding multiset  $\{y_1, y_2, \dots, y_n\}$ . For  $n \geq 1$ ,  $k \geq 0$ ,  $\sum_{(z_1, z_2, \dots) \in \mathcal{B}_{n, n+k, n+k}^1} n! / (\prod_{i \geq 1} z_i!)$  is the total number of combinations of  $(y_1, y_2, \dots, y_n)$ , such that  $\sum_{j=1}^n y_j = n+k$ , i.e.

$$\sum_{(z_1, z_2, \dots) \in \mathcal{B}_{n, n+k, n+k}^1} \frac{n!}{\prod_{i \geq 1} z_i!} = H_{(n+k)-n}^n = H_k^n = \binom{n-1+k}{k},$$

as desired.

**Example 4.** Assume  $p_0 = 0$  and  $p_k = p(1-p)^{k-1}$ ,  $k \geq 1$ , where  $0 < p < 1$ . Then for  $\lceil t \rceil \geq n$ ,

$$P(Q_t = n) = \sum_{k=0}^{\lceil t \rceil - n} \binom{n-2+k}{k} p^{n-1} (1-p)^k - \sum_{k=0}^{\lceil t \rceil - n - 1} \binom{n-1+k}{k} p^n (1-p)^k,$$

and  $P(Q_t = n) = 0$ , for  $\lceil t \rceil < n$ .

**Proof.** For  $\lceil t \rceil < n$ , the result is obvious. We now prove the case for  $\lceil t \rceil \geq n$ . By letting  $p_0 = 0$  and  $p_k = p(1-p)^{k-1}$ ,  $k \geq 1$ , in Theorem 1, and from Lemma 1, we obtain

$$P(Q_t = n) = g_{n-1, t} - g_{n, t}$$

$$\begin{aligned}
&= \sum_{(x_1, x_2, \dots) \in \mathcal{B}_{n-1, n-1, \lceil t \rceil - 1}^1} (n-1)! \left( \prod_{i \geq 1} \frac{(p(1-p)^{i-1})^{x_i}}{x_i!} \right) \\
&\quad - \sum_{(x_1, x_2, \dots) \in \mathcal{B}_{n, n, \lceil t \rceil - 1}^1} n! \left( \prod_{i \geq 1} \frac{(p(1-p)^{i-1})^{x_i}}{x_i!} \right) \\
&= \sum_{k=0}^{\lceil t \rceil - n} \left\{ \sum_{(x_1, x_2, \dots) \in \mathcal{B}_{n-1, n-1+k, n-1+k}^1} \frac{(n-1)!}{\prod_{i \geq 1} x_i!} p^{\sum_{i \geq 1} x_i} (1-p)^{\sum_{i \geq 2} (i-1)x_i} \right\} \\
&\quad - \sum_{k=0}^{\lceil t \rceil - n - 1} \left\{ \sum_{(x_1, x_2, \dots) \in \mathcal{B}_{n, n+k, n+k}^1} \frac{n!}{\prod_{i \geq 1} x_i!} p^{\sum_{i \geq 1} x_i} (1-p)^{\sum_{i \geq 2} (i-1)x_i} \right\} \\
&= \sum_{k=0}^{\lceil t \rceil - n} \binom{n-2+k}{k} p^{n-1} (1-p)^k - \sum_{k=0}^{\lceil t \rceil - n - 1} \binom{n-1+k}{k} p^n (1-p)^k. \quad (40)
\end{aligned}$$

That the last equality of (40) holds is because if  $(x_1, x_2, \dots) \in \mathcal{B}_{n-1, n-1+k, n-1+k}^1$ ,  $0 \leq k \leq \lceil t \rceil - n$ , i.e.  $\sum_{i \geq 1} x_i = n-1$  and  $\sum_{i \geq 1} ix_i = n-1+k$ , then  $\sum_{i \geq 2} (i-1)x_i = k$ . Similarly, if  $(x_1, x_2, \dots) \in \mathcal{B}_{n, n+k, n+k}^1$ ,  $0 \leq k \leq \lceil t \rceil - n - 1$ , i.e.  $\sum_{i \geq 1} x_i = n$  and  $\sum_{i \geq 1} ix_i = n+k$ , then  $\sum_{i \geq 2} (i-1)x_i = k$ .

**Remark 2.** As a comparison, for the  $p_k$ ,  $k \geq 0$ , defined in Example 4, we have

$$P(N_t = n) = \sum_{k=0}^{\lceil t \rceil - n} \binom{n-1+k}{k} p^n (1-p)^k - \sum_{k=0}^{\lceil t \rceil - n - 1} \binom{n+k}{k} p^{n+1} (1-p)^k, \quad \lceil t \rceil \geq n,$$

and  $P(N_t = n) = 0$ , for  $\lceil t \rceil < n$ .

#### 4. Limiting and some other related results

For the renewal process  $\{N_t, t \geq 0\}$ , it is known that for every  $t > 0$  and  $r > 0$ ,  $E(N_t^r) < \infty$ , and as  $t \rightarrow \infty$ ,  $N_t \xrightarrow{\text{a.s.}} \infty$ ,  $N_t/t \xrightarrow{\text{a.s.}} 1/\mu$ , and  $E(N_t)/t \rightarrow 1/\mu$ , where  $\mu = E(X_1) < \infty$ . By using (6), (7) and (10), the following consequence is immediate.

**Theorem 2.** Let  $\mu = E(X_1)$ .

- (i). For  $t > 0$  and  $r > 0$ ,  $E(Q_t^r) < \infty$ .
- (ii). If  $t \rightarrow \infty$ , then  $Q_t \xrightarrow{\text{a.s.}} \infty$ .
- (iii). Let  $\mu < \infty$ . If  $t \rightarrow \infty$ , then  $Q_t/t \xrightarrow{\text{a.s.}} 1/\mu$ .
- (iv). Let  $\mu < \infty$ . If  $t \rightarrow \infty$ , then  $E(Q_t)/t \rightarrow 1/\mu$ .

Along the lines of the proof for the renewal process  $\{N_t, t \geq 0\}$ , the central limit theorem also holds for  $\{Q_t, t \geq 0\}$ .

**Theorem 3.** Let  $\mu = E(X_1) < \infty$ , and  $\sigma^2 = \text{Var}(X_1) < \infty$ , then

$$\frac{Q_t - t/\mu}{\sigma\sqrt{t/\mu^3}} \xrightarrow[t \rightarrow \infty]{d} \mathcal{N}(0, 1). \quad (41)$$

For the process  $\{N_t, t \geq 0\}$ , it is well known

$$E(S_{N_t+1}) = E(X_1 + \cdots + X_{N_t+1}) = E(X_1)E(N_t + 1), \quad t \geq 0. \quad (42)$$

Although  $N_t$  is not a stopping time,  $Q_t$  nevertheless is a stopping time. Hence by the Wald equality and (i) of Theorem 2, we have

$$E(S_{Q_t}) = E(X_1)E(Q_t), \quad t \geq 0. \quad (43)$$

We use Example 3 to illustrate (43).

**Example 3.(Continued)** For  $t = 0$ ,  $Q_0 = 0$ , and  $S_{Q_0} = 0$ , hence (43) holds. For  $t > 0$ , as  $E(S_{Q_t}) = \lceil t \rceil$ ,  $E(Q_t) = \lceil t \rceil / p_1$ , and  $E(X_1) = p_1$ , (43) holds again.

We now give an example to present a partial answer of the Hello Kitty magnets example mentioned in Introduction. Note that if one magnet is given at a time, the expected number of purchases to collect a complete set of 41 magnets is  $t = 176.42$ .

**Example 5.** Let  $Y_1, Y_2, \dots$  be i.i.d. random variables with the same distribution as  $Y$ , where for every  $i \geq 1$ ,  $Y_i$  denotes the amount that a consumer spends at the  $i$ -th purchase at 7-Eleven. Assume that  $Y \sim \text{Uniform}\{1, 2, \dots, 250\}$ . Then

$$\begin{aligned} p_0 &= P(X_1 = 0) = P(1 \leq Y \leq 76) = \frac{76}{250}, \\ p_1 &= P(X_1 = 1) = P(77 \leq Y \leq 153) = \frac{77}{250}, \\ p_2 &= P(X_1 = 2) = P(154 \leq Y \leq 230) = \frac{77}{250}, \\ p_3 &= P(X_1 = 3) = P(231 \leq Y \leq 250) = \frac{20}{250}, \end{aligned}$$

$E(X_1) = 1.164$ ,  $\text{Var}(X_1) = 0.905104$ , and  $N = \sup\{i \mid i \geq 0, p_i > 0\} = 3$ . Thus by routine computations, we obtain

$$\begin{aligned} P(Q_t = n) &= \sum_{m=0}^{\lfloor \frac{\lceil t \rceil - 1}{N} \rfloor} \left(1 - \frac{76}{250}\right)^m \left(\frac{76}{250}\right)^{n-m-1} \left(\binom{n-1}{m} - \binom{n}{m} \frac{76}{250}\right) \\ &+ \sum_{m=\lfloor \frac{\lceil t \rceil - 1}{N} \rfloor + 1}^{\lceil t \rceil - 1} \sum_{k=0}^{\lceil t \rceil - 1 - m} \sum_{v=m-k}^{m - \lceil \frac{k}{2} \rceil} \frac{m!}{v!(2m - 2v - k)!(k - m + v)!} \left(\frac{77}{250}\right)^{2m-v-k} \\ &\cdot \left(\frac{20}{250}\right)^{k-m+v} \left(\frac{76}{250}\right)^{n-m-1} \left(\binom{n-1}{m} - \binom{n}{m} \frac{76}{250}\right), \quad n \geq 59, \end{aligned}$$

$$E(Q_t) = \frac{\lfloor \frac{[t]-1}{N} \rfloor + 1}{1 - P_0} + \sum_{m=\lfloor \frac{[t]-1}{N} \rfloor + 1}^{[t]-1} \sum_{k=0}^{[t]-1-m} \sum_{v=m-k}^{m-\lceil \frac{k}{2} \rceil} \frac{m!}{v!(2m-2v-k)!(k-m+v)!} \cdot \frac{250 \times 77^{2m-v-k} \times 20^{k-m+v}}{174^{m+1}},$$

and

$$\text{Var}(Q_t) = \sum_{m=0}^{\lfloor \frac{[t]-1}{N} \rfloor} \frac{81500 + 125000m}{174^2} + \sum_{m=\lfloor \frac{[t]-1}{N} \rfloor + 1}^{[t]-1} \sum_{k=0}^{[t]-1-m} \sum_{v=m-k}^{m-\lceil \frac{k}{2} \rceil} \frac{m!}{v!(2m-2v-k)!(k-m+v)!} \cdot \frac{250^2 \times 77^{2m-v-k} \times 20^{k-m+1}}{174^{m+2}} - (E(Q_t))^2.$$

Now for  $t = 176.42$ ,  $E(Q_{176.42}) \doteq 152.466$ , and  $\text{Var}(Q_{176.42}) \doteq 101.888$ . Hence the expected number of purchases to get magnets greater than or equal to 176.42 is about 152.466. Furthermore from (43), we have

$$E(S_{Q_{176.42}}) = E(X_1)E(Q_{176.42}) \doteq 1.164 \times 152.466 \doteq 177.470,$$

which is slightly greater than 176.42. Recall that, by the definition of  $Q_t$ ,  $S_{Q_t} \geq t$ ,  $t \geq 0$ .

Finally, we give the curve of the probability density function of  $Q_{176.42}$  in Figure 2, and plot the probability density functions of  $Z_{176.42}$  and  $\mathcal{N}(0, 1)$  in Figure 3, where

$$Z_{176.42} = \frac{Q_{176.42} - 176.42/1.164}{\sqrt{0.905104}\sqrt{176.42/1.164^3}}$$

is the normalized  $Q_{176.42}$ . As expected, due to Theorem 3, the normal approximation to the probability density function of  $Z_{176.42}$  is very accurate.

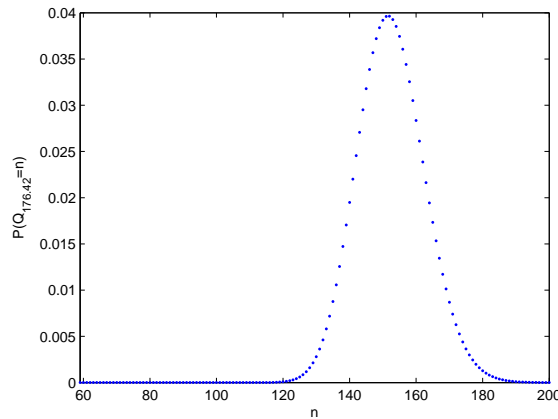


Figure 2. The probability density function of  $Q_{176.42}$

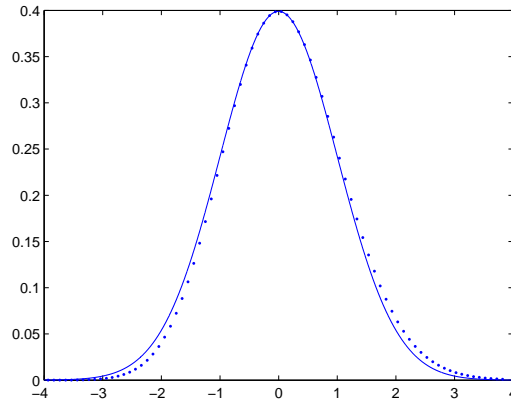


Figure 3. The probability density functions of  $Z_{176.42}$ (dotted line) and  $\mathcal{N}(0, 1)$ (solid line)

## References

1. Baum, L.E., Billingsley, P. (1965). Asymptotic distributions for the coupon collector's problem. *Ann. Math. Statist.* **36**, 1835-1839.
2. Feller, W. (1950). *Probability Theory and Its Applications*. Vol. 1. John Wiley & Sons, New York.
3. Foata, D., Han GN., Lass, B. (2001). Les nombres hyperharmoniques et la fratrie du collectionneur de vignettes. *Séminaire Lotharingien de Combinatoire* **47**, B47a, 20 pp.
4. Foata, D., Zeilberger, D. (2003). The collector's brotherhood problem using the Newman-Shepp symbolic method. *Algebra Universalis* **49**, 387-359.
5. Goodwin, H.J. (1949) On cartophily and motor cars. *Math. Gazette.* **33**, 169-171.
6. Hoslt, L. (1971). Limit theorems for some occupancy and sequential occupancy problems. *Ann. Math. Statist.* **42**, 1671-1680.
7. Von Schelling, H. (1954). Coupon collecting for unequal probabilities. *Amer. Math. Monthly* **61**, 306-311.

## 小傳

賴佳鈴於民國70年12月16日生於高雄。民國93年6月畢業於國立高雄大學應用數學系，民國93年9月進入國立高雄大學統計學研究所就讀。