

國立交通大學

應用數學系

碩士論文

更新理論在隨機 Tries 上的應用

An Introduction to Renewal Theory with  
Applications to Random Tries

研究生: 高名世

指導教授: Michael Fuchs 教授

中華民國一百零三年九月

# 前言

在電腦科學中，trie 是相當重要的資料結構，由於它的特殊生成方式，在生活中有很多很好的應用，像是搜索物件，輸入一個開頭來找出所有的可能，而 trie 的演算法很容易修改也是被廣泛應用的原因之一，其中 2-protected nodes 更是在近期很多人研究的重點。

在這篇論文裡，我們的主要目標是用機率中的更新理論來分析 trie 的一些性質，我們先證明更新理論的一些定理，並證明關鍵更新定理，接著我們介紹一些 Svante Janson 在他的 paper 中提出的一些結果並補足證明，最後我們利用更新理論算出 2-protected nodes 和 3-protected nodes 在 trie 中的期望值。

接著我們介紹一下整篇論文的架構。

第一章我們先從電腦科學中的樹開始介紹，介紹樹的定義，樹中一些有用到的名詞和各種樹的畫法與優缺點，最後介紹的 trie 與 patricia trie 的生成方式和比較，並提出了接下來要研究的方向。

第二章我們討論要用到的研究工具—更新理論，從更新過程開始，中間證明了一些接下來要使用的定理，最後證明了基本更新定理和關鍵更新定理，並區分了算術型與  $d$ -算術型的結果。並把這些結果推廣到延遲更新過程作為章節的收尾。

第三章我們把第二章的更新理論應用在 trie 裡，我們先介紹了 Svante Janson 提過的一些結果並把證明補完整，再來我們提出了一些其它有趣的結果，並給出如何用更新理論來研究這類問題的一般解法。

第四章我們用了前一章的方法來計算 2-protected nodes 和 3-protected nodes 在 trie 中算術型與非算術型的期望值，由於計算量較大，所以我們也借助了 maple 來幫助我們計算出結果。

第五章我們將整篇文章作了總結。

# Preface

The trie is a data structure of fundamental importance in Computer Science. This is reflected by the fact that tries have found numerous applications, such as in searching, sorting, contention tree algorithms, retrieving IP addresses and satellite data, internet routing, etc.

In this master's thesis, our primary target is to analyze some properties of tries using renewal theory. First, we will give a brief, self-contained introduction into renewal theory. Then, we will introduce some results of a recent paper of S. Janson in which the application of renewal theory to the analysis of tries was pioneered. Finally, we will add some new results.

The following is an outline of the structure of the thesis.

In Chapter 1, we will give some basics from Computer Science. More precisely, we will give the definitions of some data structures (including tries) and discuss their advantages and disadvantages. Moreover, the aim of this thesis will be stated in Chapter 1.

In Chapter 2, we give a detailed review of renewal theory. The main focus in this chapter will be on limit laws and we will discuss both the elementary and the key renewal theorem in details. In particular, the latter will turn out to be of key importance for the analysis of trie parameters. At the end of the chapter, we will also briefly discuss results for the delayed renewal process.

In Chapter 3, the renewal theory from Chapter 2 will be applied to tries. First we will introduce some results from a recent paper of S. Janson and prove them. Next, we will propose other interesting results which can also be handled by S. Janson's tools.

In Chapter 4, we use the method from Chapter 3 to calculate the asymptotic mean of 2-protected and 3-protected nodes in tries. These kind of nodes

have been recently introduced and have already been studied in many papers. Parts of the extensive computations in this chapter will be done with the help of Maple.

Finally, in Chapter 5, we give a conclusion to the thesis.



## 致謝

這篇論文能夠完成，首先要感謝我的指導教授 Michael Fuchs 教授。感謝老師在指導論文上的時間與用心，也讓我在長時間的學習與研究期間學習到老師做研究的認真態度與嚴謹的精神，而在完成後的修改更是感謝老師在周末或是出國期間依然抽空與我討論與修訂我沒注意到的錯誤與細節。

另外要非常感謝我的女朋友，由於健康因素，我無法使用電腦太長的時間，以至於不能用電腦來完成我的論文，所以我的論文完全是我手寫再交由我女朋友使用 latex 來完成，而她原是完全沒接觸過 latex 的人也必非數學相關系所的學生，一路從頭學起，中間會遇到的困難絕非文字能夠敘述，而她一一克服且期間還必須包容我因為壓力而亂發的脾氣，最終完成了這篇論文，實在是讓我非常的感動，沒有她就沒有這篇論文，若這篇論文能有任何貢獻，都應該與她共享。

再來要感謝我的研究室室友、同屆的很多好朋友和很多的學長姐學弟妹，陪伴我度過研究生的生涯，也在我遇到困難時對我適時地伸出援手，與我討論並一起解學論文或課業上的困難，我們也一起把數學的專業應用在生活、興趣之中，讓我的碩士生涯能多采多姿。

最後我要感謝我的家人，長年來對我的照顧與支持才能讓我無後顧之憂的專注在課業與研究之中。所有的喜悅與成果我希望能和所有人分享！

# Contents

前言	1
1 Introduction	6
2 更新理論的介紹	11
3 更新理論在 trie 上的應用	28
4 2-Protected Nodes 和 3-Protected Nodes	41
5 結論	48
附錄	49

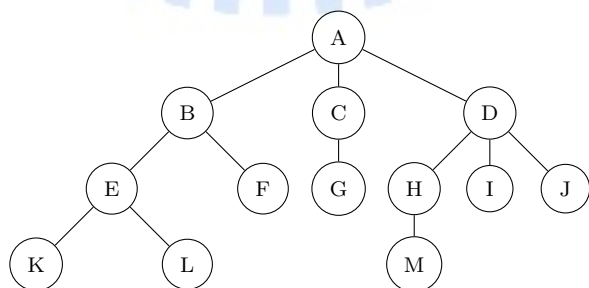
# Chapter 1

## Introduction

這篇論文的主要目的，是介紹更新理論 (renewal theory) 和了解更新理論在資料結構 trie 上的應用並分析其優缺點。首先我們先從資料結構 (data structure) 開始介紹，在電腦科學 (computer science) 中，資料結構是電腦中儲存、組織資料的方式，一個設計良好的資料結構，應該在盡可能使用較少的時間與空間資源的前提下，來儲存、刪除、插入或是找尋資料。不同種類的資料結構適合於不同種類的應用，基本的資料結構有陣列 (array)、串列 (list)、堆疊 (stack)、佇列 (queue)、樹 (tree) 和圖 (graph)，而我們要討論的是樹這種類型的資料結構。

先給出樹的定義，任意兩個點間只有一條路徑連接的圖就稱作”樹”，所以一個樹的圖是沒有迴路 (loop) 的，因此可以看出一些等價的定義，如在樹裡任意加一條邊就會形成迴路，在樹裡任意拿掉一條邊，則就不再是連通 (connect) 圖、樹中的任意兩個頂點都能被唯一的路徑連通。如果我們指一個點叫做根，則這個圖稱為有根樹。

而一般樹圖的畫法，是在最上面先畫一個點，而這個特殊的點稱為根 (root)，接著在下一層畫出數個點稱為節點 (node)，並把節點與根用分支 (branch) 連接，再把這層的節點當作原本樹裡的子樹 (subtree) 的根，用前述的方式繼續畫下去直到所有的點畫完為止，像圖一



圖一

而一個節點與子樹連接的分支數目稱為分支度 (degree)，像圖一的 A 分支度為 3，C 的分支度為 1，F 的分支度為 0。所有分支度為 0 的節點我們稱為樹葉 (leaf) 或是終節點 (external node)，而不是終節點的稱為中間節點 (internal node)，我們又將節點的子樹樹根都稱做孩子 (children)，像圖一 H、I、J 就是 D 的孩子。另外我們定義階度 (level) 表示節點所在第幾層，樹根的階度就是 0，最後我們定義一棵樹的高度 (height) 為此樹的最大階度，像圖一的高度為 3。

再來我們要介紹其中一種樹的型態 — 二元樹 (binary tree)，二元樹的定義為：每個節點最多只可以有兩個分支，也就是節點的分支度最多只有 2，二元樹也可以是空集合，另外每個節點左邊的子樹稱之為左子樹 (left subtree)，另一邊則稱做右子數 (right subtree)。而二元樹的孩子是有順序的，像圖二左右兩個圖示不同的。



圖二

這裡我們不難看出，一個二元樹如果階度為  $i$  則第  $i$  層的最大節點數為  $2^i$ ，而一個二元樹如果高度為  $k$ ，則樹的最大節點數為  $2^{k+1} - 1$ 。這裡補充一個樹葉的數目與分支度為 2 的節點之間的關係：如果一個非空的二元樹有  $n$  個樹葉， $m$  個分支度為 2 的節點，則  $n = m + 1$ 。最後介紹一個詞作為二元樹介紹的收尾，如果一個深度為  $k$  的二元樹且有  $2^{k+1} - 1$  個節點，則此樹稱為完滿二元樹 (full binary tree or complete tree)，而我們也可以看出一個有  $n$  個節點的完滿二元樹深度為  $\log_2(n + 1) - 1$ 。

再介紹二元搜尋樹之前，我們先把樹與資料結構 (data structure) 的關係做連接，資料結構主要是在討論如何把原始的資料 (data) 來作組織、安排儲存到電腦中的一門學問，而樹圖就是組織資料的一種方式，其中節點通常代表資料或是指標 (suffix)，而資料的種類有很多種，如數字或是字串 (string)，這篇論文討論的資料主要都是由 0,1 組成的字串且長度是無限 (infinite) 的。

我們再來介紹 -二元搜尋樹 (binary search tree) 或是有序二元樹 (ordered binary tree)，定義是有以下性質的二元樹就稱為二元搜尋樹，注意這裡討論的資料是數字且所有的資料都儲存在節點中：

1. 若任意節點的左子樹不空，則左子樹上所有節點的值均小於它的根的值。
2. 任意節點的右子樹不空，則右子樹上所有節點的值均大於它的根的值。
3. 任意節點的左、右子樹也分別為二元搜尋樹。



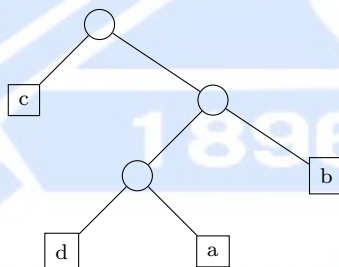
#### 4 沒有鍵值相等的節點 (no duplicate nodes)。

接者要介紹的是一個特殊的資料結構 — trie，又稱做前綴樹或是字典樹，Trie 這個詞源自於 retrieval，trie 的發明者 Edward Fredkin 把它讀做 "tree"，但其後大部分的人都讀做 "try"。就如同他的名字，可以像字典一樣依照字母找尋資料或是搜尋資料，當字串的量很大的時候這個資料結構就顯的非常有效，因為接下來討論的資料都是字串，而 Trie 的特性是利用每個共同前綴 (common prefix) 當作指標來搜尋或是插入資料，所以可以大量節省搜尋時間。接下來就來正式定義 trie 的排列方式: 如果沒有資料則不用討論，如果資料只有一筆，則資料會存在根裡，而資料如果超過二筆，首先樹根的部分是沒有任何資料，再來把所有資料是由 0 開頭的放在樹根左邊的孩子節點，1 開頭的放在右邊的孩子節點，接者把第二階的節點當作子樹的樹根，把資料第二個數字是 0 開頭的放在左邊節點，1 開頭的放在右邊，以這樣的分析法，直到每個資料都恰好分在不同的葉子為止，也就是說每個葉子只會有一筆資料，依此類推直到把所有資料分類完成，最後所有的資料都會被儲存在終節點裡，而中間節點只是指標裡面並不會有任何資料。所以 trie 跟其他的二元樹有些不同。

舉例來說

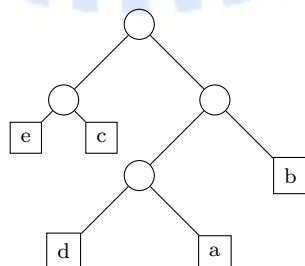
$$\{a = 101\dots, b = 1100101\dots, c = 010101\dots, d = 1000\dots\}$$

則 trie 的圖為



圖三

如果我們在插入一個資料  $e=001\dots$  則 trie 的圖為

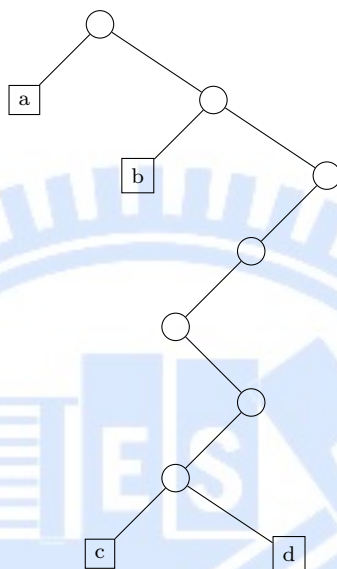


圖四

再舉第二個例子，如果資料為

$$\{a = 010101 \dots, b = 1000 \dots, c = 1100100 \dots, d = 1100101 \dots\}$$

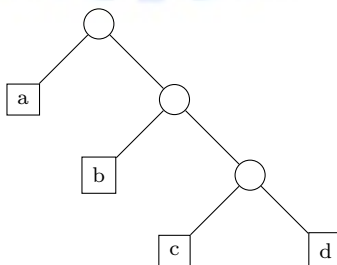
則 trie 的圖為



圖五

由這幾張圖我們可以輕易的看出 trie 這種資料結構的優缺點，這樣的資料結構可以以最快的速度找到資料，且插入資料的順序與最後的圖無關，不會受到插入的順序而改變圖的樣子，但缺點就是多了很多沒有用的空間（看最後一張圖可以看到很多多餘的節點）。

另外 trie 還有不少種類，首先是 b-trie: 每個節點最多可以儲存 b 個資料，也就是說如果有個點是中間節點，那就表示至少有 b+1 個點是由這個中間節點的指標作為開頭的。另外一個是 patricia trie: 跟原本的 trie 很類似，但是要刪掉所有只有一個孩子的節點，也就是說所有的中間節點都剛好有兩個孩子，舉例來說，如果把上面第二個例子的資料用 patricia trie 來儲存圖就會變成圖六。



圖六

從這個圖也可以看出 patricia trie 跟 trie 比起來在儲存空間上是節省很多的，但也可以看出 patricia trie 的缺點，就是不再和 trie 一樣，儲存資料時，必須為資料另外編寫指標的難度變高了，因為每個節點不一定是完整的指標，像圖六的兩個節點指標分為  $10\dots$  和  $110010\dots$ ，這在建構一個搜尋樹時會變得困難，這也是 patricia trie 無法取代 trie 的原因。

最後我們先給出一個簡單的隨機模型(random model)，假設我們的資料  $\Xi = \xi_1\xi_2\dots$  是 i.i.d 的隨機無窮字串，而  $\xi_i \in \{0,1\}$  且  $\xi_i$  也是 i.i.d 其中  $\xi_i \sim Be(p), p \in (0,1)$ ，也就是  $\mathbb{P}(\xi_i = 1) = p$  and  $\mathbb{P}(\xi_i = 0) = 1 - p = q$ 。

所以

$$\mathbb{P}(\xi_1 \dots \xi_n) = \prod_{i=1}^n \mathbb{P}(\xi_i) = \prod_{i=1}^n p^{\xi_i} q^{1-\xi_i}.$$

有了這個我們就可以來看這篇論文要討論的問題。

主要的問題就是如何看出一個 trie 的圖的走向，哪些地方會有多餘的節點？多出來的節點的部分如何？一個圖裡有多少特殊條件的節點等等問題。分析上述問題已經有一些方法了，此篇是要介紹一個數學工具：隨機過程中的更新定理。後面的章節會先介紹更新定理的原理與證明、用方法與優缺點，有哪些已經可以應用的證明。很快會發現更新定理在計算節點數量的期望值時非常的好用與快速，但算變異數時則不是個好方法，也會在證明中看到為何不好用在變異數上。

## Chapter 2

### 更新理論的介紹

更新理論最早是用來研究一些物品用壞後需要更換的問題，這些問題裡有不少機率的部份牽扯在其中，例如我們固定使用某牌生產的電池，只要沒電了就立即更換新的，而更換的每個時刻蒐集起來就是一個更新過程。

後來發現在機率理論裡有許多應用問題和更新理論研究的問題相關，而近年來更新理論已經不只在研究更換的問題，在很多領域都有很多有用的發展，像是我們假設某個商店只賣一種商品，並假設顧客每次來買的量為 i.i.d，若存貨少於一定量就補滿，沒少於定量則不補貨，則補貨次數也能看成一個更新過程。

另外在保險的領域研究生與死的過程或是平均壽命和剩餘壽命等問題也有廣泛的發展，而我們這篇文章主要是要研究更新理論在電腦科學裡的應用。

再介紹更新理論之前，我們先介紹更新過程 (renewal process) 的定義和一些相關的結果。

#### Definition 2.1.

令  $\{X_n; n \in \mathbb{N}\}$  是獨立且分佈相同的非負隨機變數數列，分佈函數為  $F(x)$ ,  $F(0) = \mathbb{P}(X_n = 0) < 1$ ,

$$(1) S_n = \sum_{i=1}^n X_i,$$

$$(2) v(t) = \sup \{n; S_n \leq t\}.$$

則我們稱  $\{v(t); t \geq 0\}$  為一個更新過程。

我們用換燈泡的問題來看更新過程裡  $X_i$ 、 $S_n$  和  $v(t)$  的意義，假設在時間  $t = 0$  的時候我們裝了一個新的燈泡，使用到  $t = X_1$  的時候燈泡壞了，我們在此時立刻更換一個新的，第二個燈泡在  $t = X_1 + X_2$  時壞掉則再換上一個新的，依此類推，第  $n$  個燈泡會在  $S_n$  時壞掉且更換一個新的，如果沒有外力的影響且燈泡的品質很穩定，則我們就可以假設燈泡的使用期限是獨立且有共同的分佈，而  $\{v(t); t \geq 0\}$  就表示時間  $t$  內更換了多少燈泡的過程。

另外我們也可以看一下  $S_{v(t)}$  和  $S_{v(t)+1}$  的意義，如果  $v(t)$  是一個更新過程，則  $S_{v(t)}$  就代表更新過程在時間  $t$  以前最後一次的更新時間。所以  $S_{v(t)+1}$  就是時間  $t$  以後第一次更新的時間，也有  $S_{v(t)} \leq t < S_{v(t)+1}$  的關係。

接著我們來看  $v(t)$  的分佈，首先根據強大數法則我們知道

$$\frac{S_n}{n} \rightarrow \mathbb{E}(X_1) =: \mu, \text{ as } n \rightarrow \infty.$$

且因為  $\mu > 0$ ，所以當  $n$  趨近無限大時， $S_n$  也會趨近於無限大，所以我們知道只有有限個  $n$  可以使得  $S_n \leq t$ ，而  $v(t) = \sup\{n | S_n \leq t\}$  所以我們有了以下的結論。

### Lemma 2.2.

$\mathbb{P}(v(t) < \infty) = 1$ ，所以  $v(t)$  是一個 proper random variable。

另外我們由定義可以知道對所有的  $n \geq 0$ ，則

$$v(t) \geq n \Leftrightarrow S_n \leq t.$$

所以如果  $t \geq 0$  且  $n \in \mathbb{N}$ ，

$$\begin{aligned} \mathbb{P}(v(t) = n) &= \mathbb{P}(v(t) \geq n) - \mathbb{P}(v(t) \geq n + 1) \\ &= \mathbb{P}(S_n \leq t) - \mathbb{P}(S_{n+1} \leq t) \\ &= F_n(t) - F_{n+1}(t). \end{aligned}$$

這裡  $F_n$  是  $S_n$  的分佈函數且對所有的  $t \geq 0$  時  $F_0(t) = 1$ 。

現在我們給出一個重要的定義。

### Definition 2.3.

$U(t) = \mathbb{E}v(t)$ ，我們稱  $U(t)$  為更新函數 (renewal function)。

更新函數  $U(t)$  是接下來要討論的重點之一，我們先從一些比較直觀的性質來了解更新函數。

### Lemma 2.4.

對所有的  $0 \leq t < \infty$ ，則

$$(1) U(t) = \sum_{n=1}^{\infty} F_n(t),$$

$$(2) U(t) < \infty.$$

*Proof.* 首先證明第一部分，由定義就可以得到

$$\begin{aligned} U(t) &= \mathbb{E}v(t) = \sum_{n=1}^{\infty} n\mathbb{P}(v(t) = n) \\ &= \sum_{n=1}^{\infty} \mathbb{P}(v(t) \geq n) \\ &= \sum_{n=1}^{\infty} \mathbb{P}(S_n \leq t) \\ &= \sum_{n=1}^{\infty} F_n(t). \end{aligned}$$

再來我們要證明更新函數是有限的，因為  $\mathbb{P}(X_n = 0) < 1$  且  $F_n$  是右連續的函數，所以存在一個  $\alpha > 0$  使得  $\mathbb{P}(X_n \geq \alpha) > 0$ 。

現在我們定義一個新的更新過程。

$$\{\tilde{X}_n, n \geq 1\} \text{ 其中 } \tilde{X}_n = \begin{cases} 0, & \text{if } X_n < \alpha; \\ \alpha, & \text{if } X_n \geq \alpha. \end{cases}$$

再來我們定義  $\tilde{v}(t) = \sup\{n : \sum_{i=1}^n \tilde{X}_i \leq t\}$ 。

在這很容易看出新的更新過程只在  $t = n\alpha$  的時候更新，這裡  $n$  是整數，且更新次數是 i.i.d 的幾何分佈，它的期望值為

$$\frac{1}{\mathbb{P}(X_n \geq \alpha)}.$$

所以

$$\mathbb{E}\tilde{v}(t) \leq \frac{\frac{t}{\alpha} + 1}{\mathbb{P}(X_n \geq \alpha)} < \infty.$$

又因為  $\tilde{X}_n \leq X_n$ ，所以我們就得到了  $U(t) = \mathbb{E}v(t) \leq \mathbb{E}\tilde{v}(t) < \infty$ 。 ■

接著我們來介紹何謂更新方程式 (renewal equation)。

#### Definition 2.5.

如果  $K(t)$  為一個函數且  $a(t)$  為一個已知的函數， $F$  是一個在  $(0, \infty)$  上的分佈函數，則

$$K(t) = a(t) + \int_0^t K(t-s)dF(s),$$

稱為更新方程式。

現在我們要把更新方程式與更新函數作連接，之後很多定理都會使用到這個性質。

#### Proposition 2.6.

$U(t)$  為更新方程式的解。

*Proof.* 在更新理論裡，大多先給定第一個更新的時間  $X_1$ ，因為過了  $X_1$  以後一切就像重新開始。

另外

$$\mathbb{E}(v(t) | X_1 = x) = \begin{cases} 0, & \text{if } x > t; \\ 1 + U(t-x), & \text{if } x \leq t. \end{cases}$$

這個式子的意義是如果  $X_1 > t$ ，則時間  $t$  內還沒更新，而  $X_1 \leq t$  則是表示在時間 0 到  $t$  之間平均更新了  $x$  那次加上  $U(t-x)$  次。

加上總機率法則

$$\mathbb{E}(\mathbb{E}(X | Y)) = \mathbb{E}(X),$$

我們得到

$$\begin{aligned}
 U(t) &= \mathbb{E}(v(t)) = \mathbb{E}(\mathbb{E}(v(t) \mid X_1 = x)) = \int_0^t \mathbb{E}(v(t) \mid X_1 = x) dF(x) \\
 &= \int_0^t (1 + U(t-x)) dF(x) \\
 &= F(x)|_0^t + \int_0^t U(t-x) dF(x) \\
 &= F(t) + \int_0^t U(t-x) dF(x).
 \end{aligned}$$

$U(t)$  滿足更新方程式的條件。 ■

接下來我們要開始介紹一些極限的定理，在研究的過程中，我們發現當時間  $t$  趨近無限大時，有很多很好的性質，所以接著要先看更新過程中的極限定理，最後來看更新定理作為這章的收尾。

接著我們要介紹一些證明基本更新定理 (elementary renewal theorem) 的工具，我們要從更新方程式著手。

### Theorem 2.7.

假設  $a(t)$  為一個有界函數， $K(t) = a(t) + \int_0^t K(t-s) dF(s)$  是一個更新方程式，則滿足上述更新方程式且在任意區間中皆為有界的唯一解為

$$K(t) = a(t) + \int_0^t a(t-x) dU(x).$$

*Proof.* 因為  $U(t)$  是一個非遞減的函數，又因為  $a(t)$  是有界限的，所以對所有的  $M > 0$ ，我們有

$$\begin{aligned}
 \sup_{0 \leq t \leq M} |K(t)| &\leq \sup_{0 \leq t \leq M} |a(t)| + \int_0^M \sup_{0 \leq y \leq M} |a(y)| dU(t) \\
 &= \sup_{0 \leq t \leq M} |a(t)| (1 + U(M)) < \infty.
 \end{aligned}$$

由上述不等式我們得到

$$K(t) = a(t) + \int_0^t a(t-x) dU(x) \text{ 在有限區間中有界 — (1)}$$



另外，我們知道  $F_k(t) = (F * F_{k-1})(t)$ ，所以

$$\begin{aligned}
 K(t) &= a(t) + (U * a)(t) \\
 &= a(t) + \left( \sum_{k=1}^{\infty} F_k(t) \right) * a(t) \\
 &= a(t) + F(t) * a(t) + \left( \sum_{k=2}^{\infty} F_k(t) \right) * a(t) \\
 &= a(t) + F(t) * \left( a(t) + \left( \sum_{k=1}^{\infty} F_k(t) \right) * a(t) \right) \\
 &= a(t) + (F * K)(t).
 \end{aligned}$$

所以

$$K(t) = a(t) + \int_0^t a(t-x) dU(x).$$

為滿足更新方程式的解——(2)

最後我們假設  $J(t)$  也是滿足條件的解，我們令  $L(t) := J(t) - K(t)$ ， $L(t) = \int_0^t L(t-x) dF(x) = (F * L)(t)$ ，

所以

$$L(t) = (F_n * L)(t)，對所有的  $n \geq 1$ 。$$

因為  $J(t)$  和  $K(t)$  在有限區間中有界，所以  $L(t)$  在有限區間中也有界，且根據 Lemma 2.4 中  $U(t) < \infty$ ，我們可以得到當  $n$  趨近於無限大時  $F_n$  會趨近於 0，所以

$$|(F_n * L)(t)| \leq \left( \sup_{0 \leq u \leq t} |L(u)| \right) F_n(t) \rightarrow 0, \forall t \geq 0.$$

所以  $L(t) = 0$  也就是  $J(t) = K(t)$ ——(3)

結合 (1)，(2) 和 (3)，我們證明了

$$K(t) = a(t) + \int_0^t a(t-x) dU(x).$$

為滿足更新方程式且在所有有限區間裡皆為有界的唯一解。 ■

有了以上的定理，我們就可以證明一個重要的結果和特例 — Wald's Equation。

另外，如果以下的定理如果沒有別表示，則我們都令  $\{X_n, n \geq 1\}$  為一更新過程， $\mathbb{E}(X_1)$  都用  $\mu$  表示且  $0 < \mu < \infty$ ，而  $X_n$  的共同分佈函數為  $F(t)$ ， $U(t)$  為更新函數，就不再列入定理的條件裡。

**Corollary 2.8.**

對所有的  $t \geq 0$ ，則

$$\mathbb{E}(S_{v(t)+1}) = \mathbb{E}(X_1)(U(t) + 1).$$

*Proof.* 我們要證明  $\mathbb{E}(S_{v(t)+1})$  滿足更新方程式的條件，首先令  $\mathbb{E}(S_{v(t)+1}) = K(t)$

$$\mathbb{E}(S_{v(t)+1} \mid X_1 = x) = \begin{cases} x, & \text{if } x > t; \\ x + K(t - x), & \text{if } x \leq t. \end{cases}$$

應用總機率法則我們可以得到

$$\begin{aligned} \mathbb{E}(S_{v(t)+1}) &= \int_0^\infty \mathbb{E}(S_{v(t)+1} \mid X_1 = x) dF(x) \\ &= \int_0^t (x + K(t - x)) dF(x) + \int_t^\infty x dF(x) \\ &= \int_0^\infty x dF(x) + \int_0^t K(t - x) dF(x) \\ &= \mathbb{E}(X_1) + \int_0^t K(t - x) dF(x). \end{aligned}$$

所以  $K(t)$  滿足更新方程式的條件，所以可以使用 Theorem 2.7 得到

$$\begin{aligned} K(t) &= a(t) + \int_0^t a(t - x) dU(x), \text{ 其中 } a(t) = \mathbb{E}(X_1) \\ &= \mathbb{E}(X_1) + \int_0^t \mathbb{E}(X_1) dU(x) \\ &= \mathbb{E}(X_1)(U(t) + 1). \end{aligned}$$

我們就證完了 Wald's Equation。 ■

再來我們來看一些  $v(t)$  的性質，很直觀的我們就能看出如果一家工廠出產的燈泡壽命不是無限大的話，如果我們希望維持一個地方一直是亮的，就必須無限多次的更換燈泡，所以我們有下述的引理。

**Lemma 2.9.**

當  $t$  趨近於無限大時， $v(t) \xrightarrow{a.s.} \infty$ 。

*Proof.* 如果  $\lim_{t \rightarrow \infty} v(t) < \infty$ ，則至少在一個  $X_K$  使得  $X_K = \infty$ ，所以

$$\begin{aligned} \mathbb{P}(\lim_{t \rightarrow \infty} v(t) < \infty) &= \mathbb{P}(\text{存在 } \{X_K\}_{K=1}^{\infty} \text{ 使得 } X_K = \infty) \\ &= \mathbb{P}(\cup_{K=1}^{\infty} (X_K = \infty)) \\ &\leq \sum_{k=1}^{\infty} \mathbb{P}(X_K = \infty) = 0. \end{aligned}$$

我們得到當  $t$  趨近無限大時， $v(t) \rightarrow \infty$ 。 ■

有了這個引理，那我們就要接著關心  $v(t)$  趨近到無限大的速率為何？我們透過下面的極限定理可以得知。

**Theorem 2.10.**

當  $t$  趨近無限大時， $\frac{v(t)}{t} \xrightarrow{a.s.} \frac{1}{\mu}$ 。

*Proof.* 由定義我們知道  $S_{v(t)} \leq t < S_{v(t)+1}$

所以我們也能得到

$$\frac{S_{v(t)}}{v(t)} \leq \frac{t}{v(t)} < \frac{S_{v(t)+1}}{v(t)}.$$

接下來我們要用夾擠來證明  $\frac{t}{v(t)} \rightarrow \mu$ 。

因為

$$\frac{S_{v(t)}}{v(t)} = \frac{\sum_{i=1}^{v(t)} X_i}{v(t)}.$$

根據強大數法則且當  $t \rightarrow \infty$  時  $v(t) \xrightarrow{a.s.} \infty$ ，所以

$$\frac{\sum_{i=1}^{v(t)} X_i}{v(t)} \xrightarrow{a.s.} \mu.$$

另外

$$\frac{S_{v(t)+1}}{v(t)} = \frac{S_{v(t)+1} (v(t) + 1)}{(v(t) + 1)v(t)} = \frac{S_{v(t)+1}}{v(t) + 1} \cdot \frac{v(t) + 1}{v(t)} \rightarrow \mu.$$

所以根據夾擠定理我們得到

$$\frac{t}{v(t)} \xrightarrow{a.s.} \mu \Rightarrow \frac{v(t)}{t} \xrightarrow{a.s.} \frac{1}{\mu}.$$

完成了證明。 ■

最後我們從看更新函數的速率來進入本章的重點之一——基本更新定理。

**Theorem 2.11.** (*Elementary renewal theorem*)

$$\lim_{t \rightarrow \infty} \frac{U(t)}{t} = \frac{1}{\mu}.$$

*Proof.* 因為  $S_{v(t)+1} > t$ ，根據 Corollary 2.8 我們得到，

$$\mu(U(t) + 1) > t,$$

也就有以下不等式，

$$\liminf_{t \rightarrow \infty} \frac{U(t)}{t} \geq \frac{1}{\mu}. \quad (1)$$

另一個方向，我們先定義一個新的更新過程  $\{\tilde{X}_n, n \in \mathbb{N}\}$ ，

而

$$\tilde{X}_n = \begin{cases} X_n, & \text{if } X_n \leq M; \\ M, & \text{if } X_n > M. \end{cases} \quad \text{其中 } M \text{ 是我們固定的一個常數。}$$

我們使用類似的符號來看新的更新過程，

$$\tilde{S}_n = \sum_{i=1}^n X_i, \quad \tilde{v}(t) = \sup\{n \mid \tilde{S}_n \leq t\}, \quad \tilde{U}(t) = \mathbb{E}\tilde{v}(t).$$

因為

$$\tilde{S}_{v(t)+1} \leq t + M,$$

所以根據 Corollary 2.8，

$$(\tilde{U}(t) + 1)\mathbb{E}(\tilde{X}_1) \leq t + M,$$

也就有以下的不等式

$$\limsup_{t \rightarrow \infty} \frac{\tilde{U}(t)}{t} \leq \frac{1}{\mathbb{E}(\tilde{X}_1)},$$

又因為  $\tilde{S}_n \leq S_n$ ，所以  $\tilde{v}(t) \geq v(t)$  且  $\tilde{U}(t) \geq U(t)$ 。

也就得到

$$\limsup_{t \rightarrow \infty} \frac{U(t)}{t} \leq \frac{1}{\mathbb{E}(\tilde{X}_1)}.$$

如果  $M$  趨近於無限大，則

$$\limsup_{t \rightarrow \infty} \frac{U(t)}{t} \leq \frac{1}{\mu}, \quad (2)$$

結合 (1) 和 (2)，我們得到  $\lim_{t \rightarrow \infty} \frac{U(t)}{t} = \frac{1}{\mu}$ 。 ■

在進入關鍵更新定理 (key renewal theorem) 前，我們要先介紹一些工具與定理才能證明關鍵更新定理，然後透過此定理來看更多更新函數的有用性質，並在下一章節把這些性質應用在資料結構 trie 的分析上。

首先我們先給出一些隨機變數的型態分類的定義。

### Definition 2.12.

存在一個  $d \in \mathbb{R}^+$  使得  $\frac{X_0}{d} \in \mathbb{N}$ ，我們取可能的  $d$  裡最大的，稱  $X_0$  為  $d$ -算術型 (d-arithmetic)，若  $d$  不存在，則稱為  $X_0$  為非算術型的。

### Definition 2.13.

設  $f$  是一個定義在  $(-\infty, \infty)$  區間中的一個非負的函數，對所有的  $\delta > 0$ ， $n \in \mathbb{Z}$ ，

又設

$$\underline{m}_n = \inf \{f(t) \mid (n-1)\delta \leq t \leq n\delta\},$$

$$\overline{m}_n = \sup \{f(t) \mid (n-1)\delta \leq t \leq n\delta\}.$$

$$\underline{\sigma}(\delta) = \delta \sum_{n=-\infty}^{\infty} \underline{m}_n,$$

$$\overline{\sigma}(\delta) = \delta \sum_{n=-\infty}^{\infty} \overline{m}_n.$$

如果  $\forall \delta > 0$ ,  $\bar{\sigma}(\delta)$  和  $\underline{\sigma}(\delta)$  收斂, 且如果  $\delta \rightarrow 0$  時  $\bar{\sigma}(\delta) - \underline{\sigma}(\delta)$  趨近於 0, 則稱  $f$  為直接 Riemann 可積函數 (directly Riemann integrable function), 簡寫為  $f \in D$ 。

另外, 我們再給出一個有用的定理, 我們也將使用此定理來證明關鍵更新定理。

**Theorem 2.14.** (*Blackwell theorem*).

若對所有的  $h$  大於 0, 則

- (i) 如果  $X_1$  為非算術型, 則  $\lim_{t \rightarrow \infty} (U(t+h) - U(t)) = \frac{h}{\mu}$ ,
- (ii) 如果  $X_1$  為  $d$ -算術型的, 則  $\lim_{t \rightarrow \infty} (U(t+kd) - U(t)) = \frac{kd}{\mu}, k \in \mathbb{N}$ 。

有關 Blackwell theorem 的證明, 就不附在此篇文章裡, 可以參考 D. Blackwell (1948). A renewal theorem, *Duke Math. J.*, **15**, 145-151。

現在我們要介紹本章的重點 — 關鍵更新定理。

**Theorem 2.15.** (*Key renewal theorem*).

設  $f(t) \in D$ , 則

- (i) 如果  $X_1$  是非算術型的, 則  $\lim_{t \rightarrow \infty} \int_0^\infty f(s-t) dU(s) = \frac{1}{\mu} \int_{-\infty}^\infty f(s) ds$ ,
- (ii) 如果  $X_1$  是  $d$ -算術型的, 則  $\lim_{t \rightarrow \infty} \int_0^\infty f(s-t) dU(s) = \frac{1}{\mu} \psi(t) + o(1)$ ,

其中  $\psi(t) = d \sum_{k=-\infty}^{\infty} f(kd-t)$ 。

*Proof.* 因為我們知道  $U(t) * (1 - F(t)) = F(t), \forall t \geq 0$ , 所以對所有在 0 到  $t$  中間的  $b$ , 我們有

$$\begin{aligned} 1 \geq F(t) &= \int_0^t (1 - F(s-t)) dU(s) \\ &\geq \int_{t-b}^t (1 - F(s-t)) dU(s) \\ &\geq (1 - F(b))(U(t) - U(t-b)). \end{aligned}$$

若任取一個  $\delta$  在  $(0, t)$  之間使得  $F(\delta) < 1$ , 則根據上述的不等式, 我們得到

$$\sup_t (U(t) - U(t-\delta)) < \infty, \text{———(1)}$$

我們另其為  $p(\delta)$ , 也因為任何一個區間我們都可以分成有限個小區間, 所以對所有的  $\delta > 0$ ,  $p(\delta) < \infty$ 。

我們先假設  $F$  為非算術型的, 對所有的  $\delta > 0$  和  $n \in \mathbb{Z}$ , 我們令

$$i_n(x) = \begin{cases} 1, & \text{if } x \in [(n-1)\delta, n\delta]; \\ 0, & \text{if } x \notin [(n-1)\delta, n\delta]. \end{cases}$$

則

$$\int_0^\infty i_n(s-t)dU(s) = U(t - (n-1)\delta) - U(t - n\delta).$$

而根據 Theorem 2.14, 上述式子的極限為

$$\lim_{t \rightarrow \infty} \int_0^\infty i_n(s-t)dU(s) = \frac{\delta}{\mu}. \quad (2)$$

現在我們取一組常數數列  $C_n \geq 0$  使得  $\sum_{n=-\infty}^\infty C_n < \infty$ , 然後設  $u(t) =$

$$\sum_{n=-\infty}^\infty C_n i_n(t).$$

由於(1)的關係我們知道  $\int_0^\infty i_n(s-t)dU(s) \leq p(\delta)$ ,

所以

$$\begin{aligned} \sum_{n=-\infty}^k C_n \int_0^\infty i_n(s-t)dU(s) &\leq \int_0^\infty u(s-t)dU(s)(t) \\ &\leq \sum_{n=-\infty}^k C_n \int_0^\infty i_n(s-t)dU(s) + p(\delta) \sum_{n=-\infty}^\infty C_n. \end{aligned}$$

因為(2)的關係, 如果  $t \rightarrow \infty$ , 則

$$\lim_{t \rightarrow \infty} \int_0^\infty u(s-t)dU(s) = \frac{\delta}{\mu} \sum_{n=-\infty}^\infty C_n.$$

我們現在要用  $f$  是直接 Riemann 可積的性質來完成證明。

先令  $\underline{u}_\delta(t) = \sum_{n=-\infty}^\infty \underline{m}_n i_n(t)$ ,  $\bar{u}_\delta(t) = \sum_{n=-\infty}^\infty \bar{m}_n i_n(t)$ , 其中  $\underline{m}_n$  與  $\bar{m}_n$  是

來自於 Definition 2.13, 則

$$\int_0^\infty \underline{u}_\delta(s-t)dU(s) \leq \int_0^\infty f(s-t)dU(s) \leq \int_0^\infty \bar{u}_\delta(s-t)dU(s). \quad (3)$$

根據 (2) 的等式，我們得到

$$\left. \begin{aligned} \lim_{t \rightarrow \infty} \int_0^\infty \underline{u}_\delta(s-t) dU(s) &= \frac{\delta}{\mu} \sum_{n=-\infty}^{\infty} \underline{m}_n = \frac{1}{\mu} \underline{\sigma}(\delta). \\ \lim_{t \rightarrow \infty} \int_0^\infty \bar{u}_\delta(s-t) dU(s) &= \frac{\delta}{\mu} \sum_{n=-\infty}^{\infty} \bar{m}_n = \frac{1}{\mu} \bar{\sigma}(\delta). \end{aligned} \right\} \text{---(4)}$$

由於  $f \in D$ ，所以  $\lim_{\delta \rightarrow 0} \underline{\sigma}(\delta) = \lim_{\delta \rightarrow 0} \bar{\sigma}(\delta) = \int_{-\infty}^{\infty} f(t) dt$ . ---(5)

所以結合 (3)，(4)，(5)，我們就證明了

$$\lim_{t \rightarrow \infty} \int_0^\infty f(s-t) dU(s) = \frac{1}{\mu} \int_{-\infty}^{\infty} f(s) ds.$$

而如果  $F$  為  $d$ -算術型的，則把前述證明的  $\delta = d$ ，且在  $\{x + nd, n \in \mathbb{N}\}$  上取極限即可。 ■

有了關鍵更新定理，就可以找出很多有用的工具來分析 *trie* 的一些性質，接下來就介紹一些下一章節要使用的定理。

### Theorem 2.16.

設  $\mathbb{E}X_1^2 < \infty$ ，則當  $t \rightarrow \infty$  時，

(i) 如果  $X_1$  是非算術型，

$$\mathbb{E}v(t) = \frac{t}{\mu} + \frac{\mathbb{E}X_1^2}{2\mu^2} + o(1).$$

(ii) 如果  $X_1$  是  $d$ -算術型的，

$$\mathbb{E}v(t) = \frac{t}{\mu} + \frac{\mathbb{E}X_1^2}{2\mu^2} + \frac{d}{\mu} \left( \frac{1}{2} - \left\{ \frac{t}{d} \right\} \right) + o(1).$$

*Proof.* 假設  $X_1$  是非算術型的，我們另  $K(t) = U(t) + 1 - \frac{t}{\mu}$ ，首先我們要證明  $K(t)$  滿足更新方程式。

$$\begin{aligned} K(t) &= U(t) + 1 - \frac{t}{\mu} \\ &= \frac{\mathbb{E}(S_{v(t)+1})}{\mu} - \frac{t}{\mu} \quad (\text{根據 Corollary 2.8}) \\ &= \frac{\mathbb{E}(r_t)}{\mu}, \quad (\text{另 } r_t = S_{v(t)+1} - t). \end{aligned}$$



我們給定第一次更新的時刻  $X_1 = x$ ，則我們有

$$\mathbb{E}(r_t | X_1 = x) = \begin{cases} x - t, & \text{if } x \geq t; \\ \mu K(t - x), & \text{if } x < t. \end{cases}$$

所以根據總機率法則

$$\begin{aligned} \mu K(t) &= \int_0^\infty \mathbb{E}(r_t | X_1 = x) dF(x) \\ &= \int_t^\infty (x - t) dF(x) + \mu \int_0^t K(t - x) dF(x), \end{aligned}$$

且其中

$$\begin{aligned} \int_t^\infty (x - t) dF(x) &= \int_0^\infty y dF(t + y) \\ &= \int_0^\infty (1 - F(t + y)) dy, \end{aligned}$$

另  $a(t) = \frac{1}{\mu} \int_0^\infty (1 - F(t + y)) dy$ ，

則我們得到

$$K(t) = a(t) + \int_0^t K(t - x) dF(x) \text{ 滿足更新方程式。}$$

又因為  $a(t)$  是單調 (monotonic) 函數且  $1 - F(t + y) = \int_{t+y}^\infty dF(z)$ ，所以

$$\begin{aligned} &\int_0^\infty \left( \int_0^\infty (1 - F(t + y)) dy \right) dt \\ &= \int_0^\infty \int_0^\infty \int_{t+y}^\infty dF(z) dy dt \\ &= \int_0^\infty \int_t^\infty \left( \int_0^{z-t} dy \right) dF(z) dt \\ &= \int_0^\infty \int_t^\infty (z - t) dF(z) dt \\ &= \int_0^\infty \int_0^z (z - t) dt dF(z) \\ &= \frac{1}{2} \int_0^\infty z^2 dF(z) = \frac{1}{2} \mathbb{E}X_1^2 < \infty. \end{aligned}$$

根據 Theorem 2.7  $a(t)$  為更新方程式  $K(t)$  的解，現在根據關鍵更新定理，

$$\begin{aligned}\lim_{t \rightarrow \infty} (U(t) + 1 - \frac{t}{\mu}) &= \frac{1}{\mu} \int_0^{\infty} a(x) dx \\ &= \frac{1}{2\mu^2} \mathbb{E}X_1^2.\end{aligned}$$

另外如果  $X_1$  為  $d$ -算術型的，用類似的證法可得證。 ■

接著我們透過中央極限定理 (central limit theorem) 來證明  $v(t)$  在  $t$  趨近無限大時有常態分佈。

**Theorem 2.17.**

設  $\sigma^2 = \text{Var}(X_1)$  且  $0 < \sigma^2 < \infty$ ，則

$$\lim_{t \rightarrow \infty} \mathbb{P} \left( \frac{v(t) - \frac{t}{\mu}}{\sqrt{\frac{t\sigma^2}{\mu^3}}} < x \right) = \Phi(x),$$

其中  $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy$ 。

*Proof.* 首先我們設  $r_t = \frac{t}{\mu} + y\sigma\sqrt{\frac{t}{\mu^3}}$ ，因為  $v(t) < r_t \Leftrightarrow S_{r_t} > t$ ，所以

$$\begin{aligned}\mathbb{P} \left( \frac{v(t) - \frac{t}{\mu}}{\sigma\sqrt{\frac{t}{\mu^3}}} < y \right) &= \mathbb{P}(v(t) < r_t) \\ &= \mathbb{P}(S_{r_t} > t) \\ &= \mathbb{P} \left( \frac{S_{r_t} - r_t\mu}{\sigma\sqrt{r_t}} > \frac{t - r_t\mu}{\sigma\sqrt{r_t}} \right) \\ &= \mathbb{P} \left( \frac{S_{r_t} - r_t\mu}{\sigma\sqrt{r_t}} > -y \left( 1 + \frac{y\sigma}{\sqrt{t\mu}} \right)^{-\frac{1}{2}} \right).\end{aligned}$$

而根據中央極限定理，當  $t \rightarrow \infty$  時，

$$\frac{S_{r_t} - r_t\mu}{\sigma\sqrt{r_t}} \rightarrow \mathbb{N}(0, 1).$$

另外如果  $t \rightarrow \infty$ ，

$$-y \left( 1 + \frac{y\sigma}{\sqrt{t\mu}} \right)^{-\frac{1}{2}} \rightarrow -y,$$

所以

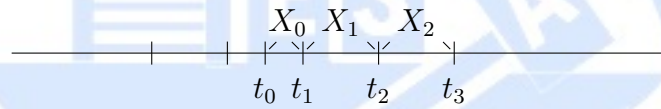
$$p\left(\frac{v(t) - \frac{t}{\mu}}{\sigma\sqrt{\frac{t}{\mu^3}}} < y\right) \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-y}^{\infty} e^{-\frac{x^2}{2}} dx.$$

又因為

$$\int_{-y}^{\infty} e^{-\frac{x^2}{2}} dx = \int_{-\infty}^y e^{-\frac{x^2}{2}} dx,$$

所以完成了證明。 ■

現在我們來推廣更新過程，實際的生活應用中，第一個間距可能與其它的間距不同分佈，以前面換燈泡的例子來說，我們從  $t_0$  這個時間開始觀察燈泡且  $t_0$  時沒有更新的行為，直到  $t_1$  才開始第一次更新，則後面更新的分佈就可能與一開始的分佈不同，如下圖，



圖七

所以我們在這定義一個延遲更新過程 (delayed renewal process)。

### Definition 2.18.

設  $X_0$  是獨立在  $X_1, X_2, \dots$  之外，且  $X_0$  可能與  $X_1, X_2$  不同分佈，則  $\bar{S}_n := \sum_{n=0}^{\infty} X_n = X_0 + S_n$ ，而  $\bar{v}(t) = \sup\{n \mid \bar{S}_n \leq t\}$  稱為延遲更新過程且  $\bar{U}(t) = \mathbb{E}\bar{v}(t)$ 。

延遲更新過程也有類似上述那些定理的性質，由於證明手法類似，我們將在下面列出結果來結束這個章節。

### Remark 2.19.

- (1) 當  $t$  趨近無限大時， $\frac{\bar{v}(t)}{t} \xrightarrow{a.s.} \frac{1}{\mu}$ 。
- (2) 設  $\mathbb{E} |X_0| < \infty$ ，則  $\lim_{t \rightarrow \infty} \frac{\bar{U}(t)}{t} = \frac{1}{\mu}$ 。
- (3) 設  $\mathbb{E}X_1^2 < \infty$ ， $\mathbb{E} |X_0| < \infty$ ，則當  $t$  趨近於無限大時。

(i) 若  $X_1$  是非算術型的，

$$\mathbb{E}\bar{v}(t) = \frac{t}{\mu} + \frac{\mathbb{E}X_1^2}{2\mu^2} - \frac{\mathbb{E}X_0}{\mu} + o(1).$$

(ii) 若  $X_1$  是  $d$ -算術型的，

$$\mathbb{E}\bar{v}(t) = \frac{t}{\mu} + \frac{\mathbb{E}X_1^2}{2\mu^2} + \frac{d}{\mu} \left( \frac{1}{2} - \mathbb{E} \left\{ \frac{t - X_0}{d} \right\} \right) - \frac{\mathbb{E}X_0}{\mu} + o(1).$$

(4) 設  $\sigma^2 = \text{Var}X_1 < \infty$ ，則當  $t$  趨近於無限大時，

$$\frac{\bar{v}(t) - \frac{t}{\mu}}{\sqrt{t}} \xrightarrow{d} N\left(0, \frac{\sigma^2}{\mu^3}\right).$$

另外，若  $\mathbb{E}X_0^2 < \infty$ ，則

$$\text{Var}(\bar{v}(t)) = \frac{\sigma^2}{\mu^3}t + o(t);$$

$$\mathbb{E} \left( \bar{v}(t) - \frac{t}{\mu} \right)^2 = \frac{\sigma^2}{\mu^3}t + o(t).$$

另外我們之後有些地方允許  $X_0$  的分佈是跟  $t$  有關的，所以如果要使用之前的定理則需要加上一些條件。

### Remark 2.20.

(1) 若  $X_0 = X_0^{(t)}$ ，則當  $t \rightarrow \infty$  時，

$$\frac{\bar{v}(t)}{t} \xrightarrow{P} \frac{1}{\mu}$$

(2) 若  $X_0 = X_0^{(t)}$ ，設  $\mathbb{E} |X_0^{(t)}| < \infty$ ，則  $\lim_{t \rightarrow \infty} \frac{\bar{U}(t)}{t} = \frac{1}{\mu}$

(3) 若  $X_0 = X_0^{(t)}$  且  $X_0^{(t)}$  是均勻可積的 (uniformly integrable)，則有 Remark 2.19(3) 的性質。

(4) 若  $X_0 = X_0^{(t)}$ ，則我們依然有 Remark 2.19(4) 的結果。而證明可以參考 S. Janson (2012). *Renewal theory in the analysis of tries and strings*, *Theor. Comput. Sci.*, **416**, 33-54.

我們將使用第二章的結果當工具來看更新理論在資料結構 *trie* 上的應用。

## Chapter 3

### 更新理論在 trie 上的應用

在進入第三章之前，我們先複習一些跟機率有關的分析背景，首先我們考慮所有的資料 (data) 都是隨機由  $\{0, 1\}$  所組成的，我們設資料為  $\Xi = \xi_1 \xi_2 \dots$ ，而之後要討論的隨機模型都如同第一章最後給出的，另外我定義隨機變數為

$$X_i := -\ln \mathbb{P}(\xi_i) = -\ln(p^{\xi_i} q^{1-\xi_i}) = \begin{cases} -\ln q, & \text{if } \xi_i = 0; \\ -\ln p, & \text{if } \xi_i = 1. \end{cases}$$

因為  $\xi_i$  是 i.i.d，所以  $X_i$  之間也是 i.i.d，最後我們給出隨機變數的動差為

$$\begin{cases} \mathbb{E}X_i = -p \ln p - q \ln q =: H, \\ \mathbb{E}X_i^2 = p \ln^2 p + q \ln^2 q =: H_2, \\ \text{Var}(X_i) = H_2 - H^2 = pq(\ln p - \ln q)^2 = pq \ln^2\left(\frac{p}{q}\right). \end{cases}$$

我們先在第一章介紹了資料結構中 trie 的排列方式，這章我們將關心如何用更新理論來看節點分佈的期望值或是相關的問題。

首先，trie 可以看成是一個無窮完滿二元樹 (complete infinite tree) 的有限子樹，其中每個節點都可以被有限的字串  $\alpha = \alpha_1 \alpha_2 \dots \alpha_k$  所標記，我們可以看出如果一個節點  $\alpha_1 \dots \alpha_k$  是 trie 的中間節點，則它的充要條件是至少有二個資料由  $\alpha_1 \dots \alpha_k$  為開頭，所以如果一個節點為終結點的話，它的充要條件是恰好只有一個資料由  $\alpha_1 \dots \alpha_k$  為開頭，且至少有二個資料由  $\alpha_1 \dots \alpha_{k-1}$  開頭。有了這兩個充要條件，我們就可以來看第一個要討論的主題 — 深度 (depth)。

### Definition 3.1.

若有  $n$  個隨機資料 (random data)，則 trie 的深度  $D_n$  表示一個 trie 中的隨機資料到根的路徑長度。

我們先來關心若有  $n$  個隨機資料，那深度的期望值為何？我們先令一個資料  $\Xi = \xi_1 \xi_2 \cdots$ ，若這項資料的深度  $D_n$  最多是  $k$  則充要條件為沒有其它的資料由  $\xi_1 \xi_2 \cdots \xi_k$  開頭，所以  $D_n$  最多為  $k$  的條件機率為

$$\begin{aligned}\mathbb{P}(D_n \leq k \mid \Xi) &= (1 - \mathbb{P}(\xi_1 \xi_2 \cdots \xi_k))^{n-1} \\ &= \left(1 - \prod_{i=1}^k \mathbb{P}(\xi_i)\right)^{n-1} \\ &= (1 - e^{-X_1 - X_2 - \cdots - X_k})^{n-1} \\ &= (1 - e^{-S_k})^{n-1}, \text{ 其中我們令 } S_k = \sum_{i=1}^k X_i.\end{aligned}$$

現在我們令  $X_0^{(n)}$  是個獨立於  $\Xi$  之外的隨機變數，且以下的分佈。

$$\mathbb{P}(X_0^{(n)} > x) = \left(1 - \frac{e^x}{n}\right)^{n-1} = (1 - e^{x - \ln n})^{n-1}, \quad x \in (-\infty, \infty).$$

所以當  $n \rightarrow \infty$  時會收斂到  $e^{-e^x}$ ，則  $X_0^{(n)} \xrightarrow{d} X_0^*$ ，其中  $-X_0^*$  有 Gumbel distribution。

接下來，我們來看延遲更新過程 (delayed renewal process)，因為我們允許第一項  $X_0^{(n)}$  可以與後項不同分佈，所以我們先在這定義一些要使用的符號：

$$\begin{aligned}\bar{S}_n &:= \sum_{i=0}^n X_i = X_0^{(n)} + \sum_{i=1}^n X_i = X_0^* + S_n, \\ \bar{v}(t) &:= \min\{n \mid \bar{S}_n > t\}\end{aligned}$$

結合以上的觀點，我們得到了一個結果。

### Proposition 3.2.

$$D_n \stackrel{d}{=} \bar{v}(\ln n).$$

*Proof.* 直接結合上述二個式子，我們可以得到

$$\begin{aligned}\mathbb{P}(D_n \leq k \mid \Xi) &= (1 - e^{-S_k})^{n-1} \\ &= \mathbb{P}(X_0^{(n)} > \ln n - S_k) \\ &= \mathbb{P}(\bar{S}_k > \ln n) \\ &= \mathbb{P}(\bar{v}(\ln n) \leq k).\end{aligned}$$

因為對所有的  $k \geq 1$  都滿足，所以  $D_n \stackrel{d}{=} \bar{v}(\ln n)$ 。 ■

有了 Proposition 3.2 這個結果，我們就可以用更新理論裡的一些定理來看關於深度的性質。

### Theorem 3.3.

若  $p \in (0, 1)$  則  $\frac{D_n}{\ln n} \xrightarrow{P} \frac{1}{H}$ ，其中  $H = \mathbb{E}X_i = -p \ln p - q \ln q$ 。

*Proof.* 因為  $D_n \stackrel{d}{=} \bar{v}(\ln n)$ ，所以根據 Remark 2.20(1)，令  $t = \ln n$  則  $\frac{\bar{v}(t)}{t} = \frac{D_n}{\ln n} \xrightarrow{P} \frac{1}{H}$ 。 ■

更進一步應用 Remark 2.20，我們就有以下深度的期望值和分佈的定理。

在接下去之前，我們先給一個判斷  $X_1$  是否為非算術型的方法，

- $d$ -算術型： $\frac{\ln p}{\ln q}$  是有理數且  $d = \gcd(\ln p, \ln q)$ ，則  $X_1$  為  $d$ -算術型。
- 非算術型： $\frac{\ln p}{\ln q}$  是無理數，則  $X_1$  為非算術型。

### Theorem 3.4.

(1) 若  $\frac{\ln p}{\ln q}$  為無理數，當  $n \rightarrow \infty$ ，則

$$\mathbb{E}D_n = \frac{\ln n}{H} + \frac{H_2}{2H^2} + \frac{\gamma}{H} + o(1).$$

(2) 若  $\frac{\ln p}{\ln q}$  為有理數，且  $d = \gcd(\ln p, \ln q)$  當  $n \rightarrow \infty$ ，則

$$\mathbb{E}D_n = \frac{\ln n}{H} + \frac{H_2}{2H^2} - \frac{1}{H} \sum_{k \neq 0} \Gamma\left(\frac{-2\pi ik}{d}\right) e^{\frac{2\pi ik \ln n}{d}} + \frac{\gamma}{H} + o(1).$$

*Proof.* 根據 Theorem 3.3 和 Remark 2.20，把  $X_0^{(n)}$  換成  $X_0^*$ ，我們就能得到

$$(1) \mathbb{E}D_n = \frac{\ln n}{H} + \frac{H_2}{2H^2} - \frac{\mathbb{E}X_0^*}{H} + o(1).$$

$$(2) \mathbb{E}D_n = \frac{\ln n}{H} + \frac{H_2}{2H^2} + \frac{d}{H} \left( \frac{1}{2} - \mathbb{E} \left\{ \frac{\ln n - X_0^*}{d} \right\} \right) - \frac{\mathbb{E}X_0^*}{H} + o(1).$$

因為 Gumbel distribution  $-X_0^*$  的特徵函數 (characteristic function) 為  $\mathbb{E}e^{-itX_0^*} = \Gamma(1 - it)$  所以我們得到  $\mathbb{E}X_0^* = \Gamma'(1) = -\gamma$ ，而 (2) 中的

$$\begin{aligned} \mathbb{E} \left\{ \frac{\ln n}{d} - \frac{X_0^*}{d} \right\} &= \frac{1}{2} - \sum_{k \neq 0} \frac{\Gamma \left( 1 - \frac{2\pi ik}{d} \right)}{2\pi ki} e^{\frac{2\pi ik \ln n}{d}} \\ &= \frac{1}{2} + \frac{1}{d} \sum_{k \neq 0} \Gamma \left( -\frac{2\pi ik}{d} \right) e^{\frac{2\pi ik \ln n}{d}}. \end{aligned}$$

詳細的證明在 S. Janson (2006). Rounding of continuous random variables and oscillatory asymptotics, *Ann. Probab.*, **34:5**, 1807-1826. ■

### Theorem 3.5.

設  $p \in (0, 1)$ ，當  $n \rightarrow \infty$ ，則

$$\frac{D_n - \frac{\ln n}{H}}{\sqrt{\ln n}} \xrightarrow{d} N\left(0, \frac{\sigma^2}{H^3}\right),$$

其中  $\sigma^2 = H_2 - H^2 = pq(\ln p - \ln q)^2$ 。

在接下去之前，我們先來看上述的每個定理，由於  $X_0^{(n)}$  是由  $n$  來決定的，雖然根據 Remark 2.20，有些更新理論中的定理還是可以應用，只是計算比較麻煩，所以我們使用另一種方法來計算，就是把隨機的資料數量波松化 (Poissonization)。

我們假設有  $1 + \text{Po}(\lambda)$  個資料，其中  $\text{Po}(\lambda)$  為波松數 (Poisson random variable)，所有  $\xi_1 \cdots \xi_k$  開頭的資料數量有波松分佈  $\text{Po}(\lambda \mathbb{P}(\xi_1 \cdots \xi_k))$ ，我們令  $\tilde{D}_\lambda = D_{1+\text{Po}(\lambda)}$ ，設 trie 的深度為  $\tilde{D}_\lambda$ ，則

$$\begin{aligned} \mathbb{P}(\tilde{D}_\lambda \leq k \mid \Xi) &= e^{-\lambda \mathbb{P}(\xi_1 \cdots \xi_k)} = e^{-\lambda e^{-S_k}} = e^{-e^{-(S_k - \ln \lambda)}} = \mathbb{P}(-X_0^* < S_k - \ln \lambda \mid \Xi) \\ &= \mathbb{P}(S_k + X_0^* > \ln \lambda \mid \Xi), \end{aligned}$$



所以我們有  $\mathbb{P}(\tilde{D}_\lambda \leq k) = \mathbb{P}(S_k + X_0^* > \ln \lambda) = \mathbb{P}(\bar{v}(\ln \lambda) \leq k)$ ，其中  $X_0 := X_0^*$ ，不再是由  $n$  決定的了，我們得到  $\tilde{D}_\lambda \stackrel{d}{=} \bar{v}(\ln \lambda)$ 。

所以前述的定理和之後要介紹的定理都可以波松化來看，而只要使  $\lambda$  等於適當的  $n$  就能作逆波松化 (depoissonization)，舉例來說根據定理 2.17 我們有

$$\frac{\tilde{D}_\lambda - \frac{\ln \lambda}{H}}{\sqrt{\ln \lambda}} \xrightarrow{d} N\left(0, \frac{\sigma^2}{H^3}\right),$$

現在我們取  $\lambda = n \pm n^{\frac{2}{3}}$ ，因為  $\ln \lambda - \ln n = O(n^{-\frac{1}{3}})$ ，所以

$$\frac{\tilde{D}_{n \pm n^{\frac{2}{3}}} - \frac{\ln n}{H}}{\sqrt{\ln n}} \xrightarrow{d} N\left(0, \frac{\sigma^2}{H^3}\right).$$

又因為當  $n \rightarrow \infty$  時，

$$\begin{cases} \mathbb{P}(\text{Po}(n + n^{\frac{2}{3}}) > n) \rightarrow 1, \\ \mathbb{P}(\text{Po}(n - n^{\frac{2}{3}}) < n) \rightarrow 1. \end{cases}$$

有了以上的結果，我們就能在機率為  $1 - o(1)$  的同一個機率空間中比較，且  $\tilde{D}_{n - n^{\frac{2}{3}}} \leq D_n \leq \tilde{D}_{n + n^{\frac{2}{3}}}$ ，所以我們得到  $\frac{D_n - \frac{\ln n}{H}}{\sqrt{\ln n}} \xrightarrow{d} N\left(0, \frac{\sigma^2}{H^3}\right)$ ，接下來要討論的理論，我們都假設有  $1 + \text{Po}(\lambda)$  個資料來證明，再取適當的  $n$  來逆波松化。

現在我們要來討論 trie 的大小之期望值，一個有  $n$  個資料的 trie，根據 trie 的生成方式我們知道會剛好有  $n$  個終節點 (external node)，但是資料是隨機的，所以中間節點 (internal node) 的數量也是隨機的，所以 trie 的大小之期望值也就是中間節點數量的期望值。

#### Note

若有  $n$  個資料，則中間節點的數量為  $W_n$ ，若有  $1 + \text{Po}(\lambda)$  個資料，則中間節點的數量為  $\tilde{W}_\lambda$ 。

在找出  $\mathbb{E}W_n$  之前，我們先給兩個要用到的工具，第一個定理是用來判斷函數是否為直接 Riemann 可積函數。

### Theorem 3.6.

設  $f$  是一個非負函數定義在  $(-\infty, \infty)$ ，若

- (1)  $f$  是有界 (bounded) 且幾乎處處連續，
- (2) 存在一個在  $(-\infty, -A)$  上非遞減，在  $(A, \infty)$  非遞增的可積函數  $F$  使得  $0 \leq f \leq F$ ，其中  $A$  為常數，則  $f$  為直接 Riemann 可積。

*Proof.* 對所有的整數  $k$ 、 $h$ ，設  $x \in [(k-1)h, kh)$ ，

我們令  $f_{n^-}(x) := \inf_{[(k-1)h, kh]} f$ ，

$$f_{n^+}(x) := \sup_{[(k-1)h, kh]} f，$$

則當  $h \rightarrow 0$  時，對每個在  $f$  連續的點，我們有

$$f_{n^-}(x) \rightarrow f(x)，$$

$$f_{n^+}(x) \rightarrow f(x)。$$

若我們定義

$$g(x) := \begin{cases} F(x+1), & x < -A-1; \\ F(x-1), & x > A+1. \end{cases}$$

當  $|x| \leq A+1$  時， $g(x) := \sup f < \infty$ ，則因為  $F$  是可積的，所以  $g$  是可積的，且  $0 < h < 1$  時， $0 \leq f_{n^-}(x) \leq f_{n^+}(x) \leq g(x)$ ，根據控制收斂定理 (dominated convergence theorem)，當  $h \rightarrow 0$  時，

$$\int_{-\infty}^{\infty} f_{n^+} \rightarrow \int_{-\infty}^{\infty} f, \quad \int_{-\infty}^{\infty} f_{n^-} \rightarrow \int_{-\infty}^{\infty} f,$$

故完成了證明。 ■

接著我們要使用 Theorem 3.6 和關鍵更新定理來證明以下的定理。

### Theorem 3.7.

設  $f$  為非負且 a.e continuous 的函數定義在  $(0, \infty)$ ，並滿足

$$\begin{cases} f(x) = O(x^2), & 0 < x < 1, \\ f(x) = O(1), & 1 < x < \infty. \end{cases}$$

另  $F(\lambda) = \sum_{\alpha} f(\lambda \mathbb{P}(\alpha))$ ，其中  $\alpha$  為所有由  $(0, 1)$  組成的字串，則

(1) 若  $\frac{\ln p}{\ln q}$  是非算術型的，當  $\lambda \rightarrow \infty$  時，

$$\frac{F(\lambda)}{\lambda} \rightarrow \frac{1}{H} \int_0^\infty \frac{f(x)}{x^2} dx.$$

(2) 若  $\frac{\ln p}{\ln q}$  是  $d$ -算術型的，當  $\lambda \rightarrow \infty$  時，

$$\frac{F(\lambda)}{\lambda} = \frac{1}{H} \psi(\ln \lambda) + o(1),$$

其中  $\psi(t)$  可由基本的傅立葉分析 (Fourier analysis) 得到

$$\psi(t) \sim \sum_{m=-\infty}^{\infty} \widehat{\psi}(m) e^{\frac{2\pi i m t}{d}},$$

$$\widehat{\psi}(t) = \int_0^\infty f(x) x^{-2 - \frac{2\pi i m}{d}} dx.$$

*Proof.* 對所有的非負函數  $h(\alpha)$ ，若  $k \geq 0$  則

$$\begin{aligned} \sum_{\alpha_1, \alpha_2, \dots, \alpha_k} h(\alpha_1 \alpha_2 \cdots \alpha_k) &= \sum_{\alpha_1, \alpha_2, \dots, \alpha_k} \frac{h(\alpha_1 \cdots \alpha_k)}{\mathbb{P}(\alpha_1 \cdots \alpha_k)} \mathbb{P}(\alpha_1 \cdots \alpha_k) \\ &= \mathbb{E} \frac{h(\xi_1 \cdots \xi_k)}{\mathbb{P}(\xi_1 \cdots \xi_k)} = \mathbb{E}(e^{S_k} h(\xi_1 \cdots \xi_k)). \end{aligned}$$

我們另  $h(\alpha) = f(\lambda \mathbb{P}(\alpha))$ ，則

$$F(\lambda) = \sum_{\alpha} f(\lambda \mathbb{P}(\alpha)) = \sum_{\alpha} h(\alpha) = \sum_{k=0}^{\infty} \mathbb{E}(e^{S_k} f(\lambda e^{-S_k})) = \int_0^\infty f(\lambda e^{-x}) e^x dU(x).$$

又另

$g(t) = e^t f(e^{-t})$  則

$$F(\lambda) = \lambda \int_0^\infty g(x - \ln \lambda) dU(x).$$

又因為  $f$  是 a.e continuous，所以  $g$  也是，因為

$$\begin{cases} f(x) = O(x^2), & 0 < x < 1, \\ f(x) = O(1), & 1 < x < \infty, \end{cases}$$

所以  $g(t) \leq M e^{-|t|}$ ，其中  $M$  是個常數，有了以上的條件，根據定理 3.8 我們得到  $g$  是直接 Riemann 可積定義在  $(-\infty, \infty)$ ，所以應用關鍵更新定理並另  $x = e^{-t}$ ，我們得到若是非算術型，則

$$\frac{F(\lambda)}{\lambda} \rightarrow \frac{1}{H} \int_{-\infty}^{\infty} g(t) dt = \frac{1}{H} \int_{-\infty}^{\infty} e^t f(e^t) dt = \frac{1}{H} \int_0^\infty f(x) x^{-2} dx,$$

而  $d$ -算術型 根據關鍵更新定理和

$$\widehat{\psi}(m) = \int_{-\infty}^{\infty} e^{-\frac{2\pi imt}{d}} e^t f(e^{-t}) = \int_0^{\infty} f(x) x^{-2-\frac{2\pi im}{d}} dx,$$

我們能得到證明的結果。 ■

有了上面的定理，我們就可以來算出中間節點數量的期望值。

### Theorem 3.8.

(1) 若  $\frac{\ln p}{\ln q}$  是非算術型的，當  $\lambda \rightarrow \infty$  時，

$$\frac{\mathbb{E}\widetilde{W}_\lambda}{\lambda} \rightarrow \frac{1}{H}.$$

(2) 若  $\frac{\ln p}{\ln q}$  是  $d$ -算術型的，當  $\lambda \rightarrow \infty$  時，

$$\frac{\mathbb{E}\widetilde{W}_\lambda}{\lambda} = \frac{1}{H} + \frac{1}{H} \psi_W(\ln \lambda) + o(1),$$

其中

$$\psi_W(t) = \sum_{m \neq 0} -\chi_m \Gamma(-1 + \chi_m) e^{-\chi_m t}, \text{ 其中 } \chi_m = -\frac{2\pi im}{d}.$$

*Proof.* 若  $\alpha = \alpha_1 \cdots \alpha_k$  為一個有限的字串，我們訂一個指示 (indicator) 函數，

$$I(\alpha) = \begin{cases} 1, & \alpha \text{ 為中間節點,} \\ 0, & \alpha \text{ 非中間節點.} \end{cases}$$

我們知道  $\alpha$  為中間節點的充要條件為：至少有二個資料由  $\alpha$  開頭，而在波松模型裡，由  $\alpha$  開頭的資料有波松分佈  $\text{Po}(\lambda P(\alpha))$ ，所以

$$\mathbb{E}\widetilde{W}_\lambda = \sum_{\alpha} \mathbb{E}I(\alpha) = \sum_{\alpha} \mathbb{P}(\text{Po}(\lambda P(\alpha)) \geq 2).$$

我們另  $f(x) = \mathbb{P}(\text{Po}(x) \geq 2) = 1 - (1+x)e^{-x}$ ，所以

$$\mathbb{E}\widetilde{W}_\lambda = \sum_{\alpha} f(\lambda P(\alpha)).$$

可以輕易看出  $f$  在  $(0, \infty)$  是非負的且  $\begin{cases} f(x) = O(x^2), & 0 < x < 1, \\ f(x) = O(1), & 1 < x < \infty. \end{cases}$

接著利用 Theorem 3.7 和分部積分 (integration by parts)

$$\int_0^{\infty} f(x)x^{-2}dx = \int_0^{\infty} f'(x)x^{-1}dx = \int_0^{\infty} xe^{-x}x^{-1}dx = \int_0^{\infty} e^{-x}dx = 1.$$

所以根據 Theorem 3.7

$$\frac{\mathbb{E}\widetilde{W}_\lambda}{\lambda} \rightarrow \frac{1}{H} \int_0^{\infty} f(x)x^{-2}dx = \frac{1}{H}.$$

而  $d$ -算術型 也是一樣的證明手法且

$$\begin{aligned} \int_0^{\infty} f(x)x^{-2-\frac{2\pi ik}{d}}dx &= \left(1 + \frac{2\pi ik}{d}\right)^{-1} \int_0^{\infty} f'(x)x^{-1-\frac{2\pi ik}{d}}dx \\ &= \frac{\Gamma(1 - \frac{2\pi ik}{d})}{1 + \frac{2\pi ik}{d}} \\ &= \frac{2\pi ik}{d} \Gamma\left(-1 - \frac{2\pi ik}{d}\right). \end{aligned}$$

故完成了證明。 ■

現在我們可以使用之前用過的手法逆波松化來得到一般情況的中間節點數量。

### Theorem 3.9.

(1) 若  $\frac{\ln p}{\ln q}$  是非算術型的，當  $n \rightarrow \infty$  時，

$$\frac{\mathbb{E}W_n}{n} \rightarrow \frac{1}{H}.$$

(2) 若  $\frac{\ln p}{\ln q}$  是  $d$ -算術型的，當  $n \rightarrow \infty$  時，

$$\frac{\mathbb{E}W_n}{n} = \frac{1}{H} + \frac{1}{H}\psi_W(\ln n) + o(1),$$

*Proof.* 首先我們要證明  $\mathbb{P}(\text{Po}(2n) \geq n) \geq \frac{1}{2}$ 。

根據 Chebyshev's inequality: 對所有實數  $k$ ， $\mathbb{P}(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$ 。

我們令  $\mu = 2n$ ， $\sigma^2 = 2n$ ， $k\sigma = n$ ，

所以

$$\mathbb{P}(|\text{Po}(2n) - 2n| \geq n) \leq \frac{2}{n}.$$

由此推論

$$\begin{aligned} \frac{2}{n} &\geq \mathbb{P}(|\text{Po}(2n) - 2n| \geq n) = \mathbb{P}(\{\text{Po}(2n) \geq 3n\} \cup \{\text{Po}(2n) \leq n\}) \\ &\geq \mathbb{P}(\text{Po}(2n) \leq n). \end{aligned}$$

所以當  $n$  夠大時，

$$\mathbb{P}(\text{Po}(2n) \geq n) = 1 - \mathbb{P}(\text{Po}(2n) < n) \geq 1 - \frac{2}{n} \geq \frac{1}{2}.$$

再來我們要證明  $\mathbb{E}W_n \leq 2\mathbb{E}\widetilde{W}_{2n} = O(n)$ ，因為  $\mathbb{E}W_n$  是隨著  $n$  遞增而遞增，所以當  $n$  夠大時

$$\begin{aligned} \mathbb{E}\widetilde{W}_{2n} &= \sum_k \mathbb{E}W_k \mathbb{P}(\text{Po}(2n) = k) \geq \mathbb{E}W_n \sum_{k \geq n} \mathbb{P}(\text{Po}(2n) = k) \\ &= \mathbb{E}W_n \mathbb{P}(\text{Po}(2n) \geq n) \geq \frac{1}{2} \mathbb{E}W_n. \end{aligned}$$

最後我們要證明當  $\lambda_{\pm} = n \pm n^{\frac{2}{3}}$  時，

$$\mathbb{E}\widetilde{W}_{\lambda_-} + o(n) \leq \mathbb{E}W_n \leq \mathbb{E}\widetilde{W}_{\lambda_+} + o(n),$$

根據 Chebyshev's inequality，

$$\begin{aligned} (n^{-\frac{1}{3}} + n^{-\frac{2}{3}}) &\geq \mathbb{P}(|\text{Po}(n + n^{\frac{2}{3}}) - n - n^{\frac{2}{3}}| \geq n^{\frac{2}{3}}) \\ &= \mathbb{P}(\{\text{Po}(n + n^{\frac{2}{3}}) \geq n + n^{\frac{4}{3}}\} \cup \{\text{Po}(n + n^{\frac{2}{3}}) \leq n\}) \\ &\geq \mathbb{P}(\text{Po}(n + n^{\frac{2}{3}}) \leq n) \text{———(1)}, \end{aligned}$$

另外

$$1 - \mathbb{P}(\text{Po}(n + n^{\frac{2}{3}}) \leq n) = \mathbb{P}(\text{Po}(n + n^{\frac{2}{3}}) > n) \text{———(2)},$$

結合 (1)、(2) 我們得到

$$\mathbb{P}(\text{Po}(n + n^{\frac{2}{3}}) > n) \geq 1 - (n^{-\frac{1}{3}} + n^{-\frac{2}{3}}) = 1 + o(1).$$

所以

$$\begin{aligned} \mathbb{E}\widetilde{W}_{\lambda_+} &= \sum_k \mathbb{E}_k \mathbb{P}(\text{Po}(n + n^{\frac{2}{3}}) = k) \geq \mathbb{E}W_n \sum_{k \geq n+1} \mathbb{P}(\text{Po}(n + n^{\frac{2}{3}}) = k) \\ &= \mathbb{E}W_n \mathbb{P}(\text{Po}(n + n^{\frac{2}{3}}) > n) \geq \mathbb{E}W_n (1 + o(1)) = \mathbb{E}W_n + o(n). \end{aligned}$$

同樣的作法我們也能得到

$$\mathbb{E}W_n \geq \mathbb{E}\widetilde{W}_{\lambda_-} + o(n).$$

因為

$$\mathbb{E}\widetilde{W}_{\lambda_-} + o(n) \leq \mathbb{E}W_n \leq \mathbb{E}\widetilde{W}_{\lambda^+} + o(n).$$

所以

$$\frac{\mathbb{E}\widetilde{W}_{\lambda_-} + o(n)}{n} \leq \frac{\mathbb{E}W_n}{n} \leq \frac{\mathbb{E}\widetilde{W}_{\lambda^+} + o(n)}{n}.$$

根據這個不等式和 Theorem 3.8，我們就得到了期望值的結果。 ▮

接下來我們要來關心 trie 中一些常常需要用到的節點的期望值，我們先定義一些符號方便後面的介紹。

#### Note

設有  $n$  個資料，則

$L_n$ : 葉子的數量；

$O_n$ : 只有一個孩子之節點的數量；

$T_n$ : 洽有二個孩子之節點的數量；

$R_n$ : 沒有左邊孩子之節點的數量。

根據 trie 的生成規則，我們知道葉子的數量就是資料的數量，所以  $L_n = n$ ，而根據第一章提到的 patricia trie 的生成規則，我們知道 trie 的中間節點數量減去 patricia trie 的中間節點數量就是  $O_n$ ，最後  $T_n$  就是所有的節點減去  $O_n$  和  $n$ ，而圖論中也有用別的方法得到結果  $T_n = n - 1$ ，所以我們接下來就用類似前面的方法來找出  $R_n$  的期望值。

#### **Theorem 3.10.**

(1) 若  $\frac{\ln p}{\ln q}$  是非算術型的，當  $n \rightarrow \infty$  時，

$$\mathbb{E}R_n = \frac{n}{H}(p + q \ln q) + o(n).$$

(2) 若  $\frac{\ln p}{\ln q}$  是  $d$ -算術型的，當  $n \rightarrow \infty$  時，

$$\mathbb{E}R_n = \frac{n}{H}(p + q \ln q) + \frac{n}{H}\psi_R(\ln n) + o(n),$$

其中

$$\psi_R(t) = \sum_{m \neq 0} -p\Gamma(-1 + \chi_m)\chi_m e^{-\chi_m t}, \quad \chi_m = -\frac{2\pi im}{d}.$$

*Proof.* 我們先證明波松化的 case，首先，一個節點  $\alpha$  為沒有左邊孩子的節點之充要條件是：至少有兩項資料是由  $\alpha 1$  開頭的且沒有  $\alpha 0$  開頭的。所以類似前面定理的方法，我們先找出  $f$ 。

$$\begin{aligned} \text{令 } f(x) &:= \mathbb{P}(\text{Po}(px) \geq 2)\mathbb{P}(\text{Po}(qx) = 0) \\ &= (1 - e^{-px} - pxe^{-px})e^{-qx}. \end{aligned}$$

且

$$\begin{cases} f(x) = O(x^2), & 0 < x < 1, \\ f(x) = O(1), & 1 < x < \infty. \end{cases}$$

我們先計算一般解，

$$\begin{aligned} \int_0^\infty f(x)x^{s-1}dx &= \int_0^\infty (1 - e^{-px} - pxe^{-px})e^{-qx}x^{s-1}dx \\ &= \int_0^\infty (e^{-qx} - e^{-x} - pxe^{-x})x^{s-1}dx \\ &= q^{-s}\Gamma(s) - \Gamma(s) - p\Gamma(s+1) \\ &= \Gamma(s)(q^{-s} - 1 - ps). \end{aligned}$$

而  $s = -1$  是非算術型的， $s = -1 - \frac{2\pi im}{d} = -1 + \chi_m$  是  $d$ -算術型的。當  $s = -1 + \chi_m$  時  $m \neq 0$ ，

$$\begin{aligned} \Gamma(s)(q^{-s} - 1 - ps) &= \Gamma(-1 + \chi_m)(q^{-(1+\chi_m)} - 1 - p(-1 + \chi_m)) \\ &= \Gamma(-1 + \chi_m)(q - 1 - p(-1 + \chi_m)) \\ &= -p\Gamma(-1 + \chi_m)\chi_m. \end{aligned}$$

當  $s = -1$  時，因為  $\Gamma(-1)$  是發散的，所以根據 Gamma function 的泰勒展開式

$$\lim_{s \rightarrow -1} \Gamma(s)(q^{-s} - 1 - ps) = \lim_{s \rightarrow -1} \frac{q^{-s} - 1 - ps}{s + 1} = -(-q^{\ln q} - p) = p + q \ln q,$$

所以根據 Theorem 3.7 我們得到非算數型的結果為

$$\frac{\mathbb{E}\tilde{R}_\lambda}{\lambda} \rightarrow \frac{1}{H} \int_0^\infty \frac{f(x)}{x^2} dx = \frac{1}{H}(p + q \ln q).$$

而  $d$ -算術型的結果為

$$\frac{\mathbb{E}\tilde{R}_\lambda}{\lambda} = \frac{1}{H}(p + q \ln q) + \frac{1}{H}\psi_R(\ln n) + o(1),$$



其中

$$\begin{aligned}\psi_R(t) &= \sum_{m \neq 0} \int_0^\infty f(x) x^{-2 - \frac{2\pi im}{d}} dx e^{-\chi_m t} \\ &= \sum_{m \neq 0} -p \Gamma(-1 + \chi_m) \chi_m e^{-\chi_m t}, \quad \chi_m = \frac{-2\pi im}{d}.\end{aligned}$$

故完成了證明，而且我們可以用類似 Theorem 3.9 的方式來作逆波松化使資料量為  $n$  時有一樣的結果。 ■



# Chapter 4

## 2-Protected Nodes 和 3-Protected Nodes

這個章節主要是要來討論 2-protected nodes 和 3-protected nodes，由於 trie 的特殊生成方式，雖然在搜尋資料時能達到最快的速度，但卻容易佔據大量的儲存空間，這是因為每個中間節點都是一個指標而被儲存起來，例如  $A: \{a = 10000001 \dots, b = 10000000 \dots\}$ ，則圖形為



圖八

中間有一長串的指標，儲存起來很占空間，所以透過研究一些特殊的中間節點，可以看出 trie 的走向或是哪些節點的數量過多來刪除以節省空間。我們先定義何謂  $k$ -protected nodes。

### Definition 4.1.

一個中間節點與每個終結點的距離都大於或等於  $k$ ，就稱此中間節點為  $k$ -protected nodes。

Note

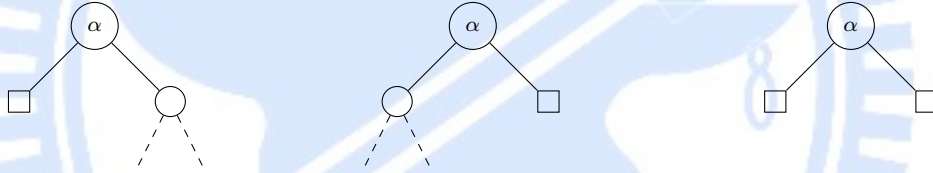
- (1) 設資料的數是為  $n$ ，則  $K_n^{(k)}$ :  $k$ -protected nodes 的數量。  
 而在波松分佈下，設資料的數量為  $1 + \text{Po}(\lambda)$ ，則  $\widetilde{K}_\lambda^{(k)} = K_{\text{Po}(\lambda)}^{(k)}$ 。  
 (2) 若  $k_1 < k_2$ ，則  $K_n^{(k_1)} \geq K_n^{(k_2)}$ 。  
 (3) 設資料的數量為  $n$ ，則

$M_n^{(k)}$ : 是  $k$ -protected nodes 但不是  $k+1$  protected nodes 的數量。

另外若資料的數量為  $1 + \text{Po}(\lambda)$ ，則

$$\widetilde{M}_\lambda^{(k)} = M_{\text{Po}(\lambda)}^{(k)}.$$

有了這些定義與符號，我們就可以進入這章的重點了，首先我們討論有  $1 + \text{Po}(\lambda)$  個資料時， $k$ -protected nodes 數量的期望值，我們發現若要直接給出  $k$ -protected nodes 的充要條件會相當複雜，所以我們用的手法為全部的中間節點減去不是  $k$ -protected nodes 的中間節點，並使用排容原理 (principle of inclusion and exclusion) 來幫助計算，而算 2-protected nodes 時要減去的節點就像是下圖的  $\alpha$ ，



圖九

而  $\alpha$  為這些點的充要條件為”至少有一個點是  $\alpha 1$  開頭和恰好只有一個點是  $\alpha 0$  開頭”或”至少一個點是  $\alpha 0$  開頭和恰好一個點是  $\alpha 1$  開頭”。

所以我們的目標是算出下列的式子，

$$\mathbb{E}\widetilde{K}_\lambda^{(2)} = \mathbb{E}\widetilde{W}_\lambda - \mathbb{E}\widetilde{M}_\lambda^{(1)}.$$

**Theorem 4.2.**

- (1) 若  $\frac{\ln p}{\ln q}$  為非算術型，當  $\lambda \rightarrow \infty$  時，

$$\frac{\mathbb{E}\widetilde{M}_\lambda^{(1)}}{\lambda} \rightarrow 1 - \frac{pq}{H},$$

- (2) 若  $\frac{\ln p}{\ln q}$  為  $d$ -算術型，當  $\lambda \rightarrow \infty$  時，

$$\frac{\mathbb{E}\widetilde{M}_\lambda^{(1)}}{\lambda} = 1 - \frac{pq}{H} + \frac{1}{H}\psi_{M^{(1)}}(t) + o(1),$$

其中

$$\psi_{M^{(1)}}(t) = \sum_{m \neq 0} \Gamma(\chi_m)(qp - qp\chi_m)e^{-\chi_m t}, \quad \chi_m = -\frac{2\pi im}{d}.$$

*Proof.* 根據對圖九中  $\alpha$  的充要條件，我們得到  $\mathbb{E}\widetilde{M}_\lambda^{(1)} = \sum_\alpha I(\alpha) = \sum_\alpha f(\lambda\mathbb{P}(\alpha))$ ，其中

$$f(x) := \mathbb{P}(\text{Po}(px) \geq 1)\mathbb{P}(\text{Po}(qx) = 1) + \mathbb{P}(\text{Po}(px) = 1)\mathbb{P}(\text{Po}(qx) \geq 1) \\ - \mathbb{P}(\text{Po}(px) = 1)\mathbb{P}(\text{Po}(qx) = 1).$$

所以

$$f(x) = (1 - e^{-px})e^{-qx}qx + (1 - e^{-qx})e^{-px}px - e^{-qx}qx e^{-px}px.$$

因為  $f$  是非負的且  $\begin{cases} f(x) = O(x^2), & 0 < x < 1, \\ f(x) = O(1), & 1 < x < \infty, \end{cases}$

所以根據 Theorem 3.7，非算術型的解為

$$\begin{aligned} \frac{\mathbb{E}\widetilde{M}_\lambda^{(1)}}{\lambda} &\rightarrow \frac{1}{H} \int_0^\infty f(x)x^{-2}dx = \frac{q}{H} \int_0^\infty \frac{e^{-qx} - e^{-x}}{x} dx \\ &\quad + \frac{p}{H} \int_0^\infty \frac{e^{-px} - e^{-x}}{x} dx - \frac{pq}{H} \int_0^\infty e^{-x} dx \\ &= \frac{-q \ln q}{H} + \frac{-p \ln p}{H} - \frac{pq}{H} \\ &= 1 - \frac{pq}{H}. \end{aligned}$$

而  $d$ -算術型的解法我們先算出一般解，

$$\begin{aligned} \int_0^\infty f(x)x^{s-1}dx &= \int_0^\infty (-e^{-x}x + e^{-qx}qx + e^{-px}px - e^{-x}qpx^2)x^{s-1}dx \\ &= -\Gamma(s+1) + q^{-s}\Gamma(s+1) + p^{-s}\Gamma(s+1) - qps\Gamma(s+1) \\ &= \Gamma(s+1)(q^{-s} + p^{-s} - 1 - qps). \end{aligned}$$

我們另  $s = -1 - \frac{2\pi im}{d} = -1 + \chi_m$ ，因為  $q^{-s}$  和  $p^{-s} = 1$ ，所以代入  $s$  並根據 Theorem 3.7 即可得到結果

$$\frac{\mathbb{E}\widetilde{M}_\lambda^{(1)}}{\lambda} = 1 - \frac{pq}{H} + \frac{1}{H}\psi_{M^{(1)}}(\ln n) + o(1).$$

故完成了證明。 ■

有了上述的結果，就能算出 2-protected nodes 的期望值。

**Theorem 4.3.**

(1) 若  $\frac{\ln p}{\ln q}$  為非算術型，當  $n \rightarrow \infty$  時，

$$\mathbb{E}K_n^{(2)} = \frac{n}{H}(1 - H + pq) + o(n).$$

(2) 若  $\frac{\ln p}{\ln q}$  為  $d$ -算術型，當  $n \rightarrow \infty$  時，

$$\mathbb{E}K_n^{(2)} = n \left( \frac{1 - H + pq}{H} \right) + \frac{n}{H} (\psi_W(\ln n) - \psi_{M^{(1)}}(\ln n)) + o(n).$$

其中  $\psi_{M^{(1)}}(t) = \sum_{m \neq 0} \Gamma(\chi_m)(qp - qp\chi_m)e^{-\chi_m t}$

*Proof.* 我們根據 2-protected nodes 的定義知道

$$\tilde{K}_\lambda^{(2)} = \tilde{W}_\lambda - \tilde{M}_\lambda^{(1)},$$

所以

$$\mathbb{E}\tilde{K}_\lambda^{(2)} = \mathbb{E}\tilde{W}_\lambda - \mathbb{E}M_\lambda^{(1)}.$$

結合 Theorem 3.8 和 Theorem 4.2 的非算術型，我們得到

$$\begin{aligned} \mathbb{E}\tilde{K}_\lambda^{(2)} &= \frac{\lambda}{H} - \lambda \left(1 - \frac{pq}{H}\right) + o(\lambda) \\ &= \lambda \left( \frac{1 - H + pq}{H} \right) + o(\lambda). \end{aligned}$$

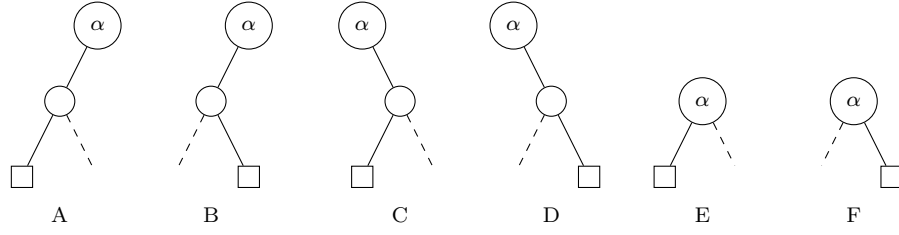
而我們可以使用 Theorem 3.9 證明裡的手法來逆波松化，

所以得到

$$\mathbb{E}K_n^{(2)} = n \left( \frac{1 - H + pq}{H} \right) + o(n).$$

$d$ -算術型也使用一樣的證明手法，故完成了證明。 ■

接著我們要來看更為複雜的 3-protected nodes 的期望值，我們用和前面相同的方法，先找出一個節點為 3-protected nodes 的充要條件；所有的中間節點減去  $\tilde{M}_\lambda^{(1)}$  和  $\tilde{M}_\lambda^{(2)}$ ，所以我們先把要減去的節點列出來，分別是



圖十

而這 6 種節點在波松分佈下的機率分別為

$$\mathbb{P}(A) := \mathbb{P}(\text{Po}(\alpha q^2) = 1)\mathbb{P}(\text{Po}(\alpha qp) \geq 1) = e^{-\alpha^2 x} q^2 x (1 - e^{-\alpha p x}),$$

$$\mathbb{P}(B) := \mathbb{P}(\text{Po}(\alpha qp) = 1)\mathbb{P}(\text{Po}(\alpha q^2) \geq 1) = e^{-\alpha p x} q p x (1 - e^{-\alpha^2 x}),$$

$$\mathbb{P}(C) := \mathbb{P}(\text{Po}(\alpha qp) = 1)\mathbb{P}(\text{Po}(\alpha p^2) \geq 1) = e^{-\alpha p x} q p x (1 - e^{-\alpha^2 x}),$$

$$\mathbb{P}(D) := \mathbb{P}(\text{Po}(\alpha p^2) = 1)\mathbb{P}(\text{Po}(\alpha qp) \geq 1) = e^{-\alpha^2 x} p^2 x (1 - e^{-\alpha p x}),$$

$$\mathbb{P}(E) := \mathbb{P}(\text{Po}(\alpha q) = 1)\mathbb{P}(\text{Po}(\alpha p) \geq 1) = e^{-\alpha x} q x (1 - e^{-\alpha p x}),$$

$$\mathbb{P}(F) := \mathbb{P}(\text{Po}(\alpha p) = 1)\mathbb{P}(\text{Po}(\alpha q) \geq 1) = e^{-\alpha p x} p x (1 - e^{-\alpha x}).$$

接著我們利用排容原理找出  $f$ ,

$$\begin{aligned} f(x) := & \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) + \mathbb{P}(D) + \mathbb{P}(E) + \mathbb{P}(F) \\ & - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(A \cap D) - \mathbb{P}(A \cap E) - \mathbb{P}(A \cap F) \\ & - \mathbb{P}(B \cap C) - \mathbb{P}(B \cap D) - \mathbb{P}(B \cap E) - \mathbb{P}(B \cap F) - \mathbb{P}(C \cap D) \\ & - \mathbb{P}(C \cap E) - \mathbb{P}(C \cap F) - \mathbb{P}(D \cap E) - \mathbb{P}(D \cap F) - \mathbb{P}(E \cap F) \\ & + \mathbb{P}(A \cap B \cap C) + \mathbb{P}(A \cap B \cap D) + \mathbb{P}(A \cap B \cap E) + \mathbb{P}(A \cap B \cap F) \\ & + \mathbb{P}(A \cap C \cap D) + \mathbb{P}(A \cap C \cap E) + \mathbb{P}(A \cap C \cap F) + \mathbb{P}(A \cap D \cap E) \\ & + \mathbb{P}(A \cap D \cap F) + \mathbb{P}(A \cap E \cap F) + \mathbb{P}(B \cap C \cap D) + \mathbb{P}(B \cap C \cap E) \\ & + \mathbb{P}(B \cap C \cap F) + \mathbb{P}(B \cap D \cap E) + \mathbb{P}(B \cap D \cap F) + \mathbb{P}(B \cap E \cap F) \\ & + \mathbb{P}(C \cap D \cap E) + \mathbb{P}(C \cap D \cap F) + \mathbb{P}(C \cap E \cap F) + \mathbb{P}(D \cap E \cap F) \\ & - \mathbb{P}(A \cap B \cap C \cap D) - \mathbb{P}(A \cap B \cap C \cap E) - \mathbb{P}(A \cap B \cap C \cap F) \\ & - \mathbb{P}(A \cap B \cap D \cap E) - \mathbb{P}(A \cap B \cap D \cap F) - \mathbb{P}(A \cap B \cap E \cap F) \\ & - \mathbb{P}(A \cap C \cap D \cap E) - \mathbb{P}(A \cap C \cap D \cap F) - \mathbb{P}(A \cap C \cap E \cap F) \\ & - \mathbb{P}(A \cap D \cap E \cap F) - \mathbb{P}(B \cap C \cap D \cap E) - \mathbb{P}(B \cap C \cap D \cap F) \\ & - \mathbb{P}(B \cap C \cap E \cap F) - \mathbb{P}(B \cap D \cap E \cap F) - \mathbb{P}(C \cap D \cap E \cap F) \\ & + \mathbb{P}(A \cap B \cap C \cap D \cap E) + \mathbb{P}(A \cap B \cap C \cap D \cap F) + \mathbb{P}(A \cap B \cap C \cap E \cap F) \\ & + \mathbb{P}(A \cap B \cap D \cap E \cap F) + \mathbb{P}(A \cap C \cap D \cap E \cap F) + \mathbb{P}(B \cap C \cap D \cap E \cap F) \\ & - \mathbb{P}(A \cap B \cap C \cap D \cap E \cap F). \end{aligned}$$

有了這個  $f$ ，我們就能算出需要減掉的節點之期望值，而用所有中間節點的期望值減去這個期望值就是 3-protected nodes 的期望值，其中  $A \cap E$ 、 $B \cap E$ 、 $C \cap F$  和  $D \cap F$  都是空集合，我們列出一些不易化簡的機率來看。

$$\mathbb{P}(A \cap B) := \mathbb{P}(\text{Po}(\alpha q^2) = 1)\mathbb{P}(\text{Po}(\alpha pq) = 1) = e^{-q^2 x} q^2 x e^{-\alpha pq x} q p x,$$

$$\begin{aligned} \mathbb{P}(A \cap F) &:= \mathbb{P}(\text{Po}(\alpha q^2) = 1)\mathbb{P}(\text{Po}(\alpha qp) \geq 1)\mathbb{P}(\text{Po}(\alpha p) = 1) \\ &= e^{-q^2 x} q^2 x (1 - e^{-qp x}) e^{-p x} p x, \end{aligned}$$

$$\begin{aligned} \mathbb{P}(B \cap F) &:= \mathbb{P}(\text{Po}(\alpha qp) = 1)\mathbb{P}(\text{Po}(\alpha q^2) \geq 1)\mathbb{P}(\text{Po}(\alpha p) = 1) \\ &= e^{-pq x} p q x (1 - e^{-q^2 x}) (e^{-p x} p x), \end{aligned}$$

$$\mathbb{P}(C \cap D) := \mathbb{P}(\text{Po}(\alpha p^2) = 1)\mathbb{P}(\text{Po}(\alpha pq) = 1) = e^{-p^2 x} p^2 x e^{-\alpha pq x} p q x,$$

$$\begin{aligned} \mathbb{P}(C \cap E) &:= \mathbb{P}(\text{Po}(\alpha q) = 1)\mathbb{P}(\text{Po}(\alpha qp) = 1)\mathbb{P}(\text{Po}(\alpha p^2) \geq 1) \\ &= e^{-q x} q x e^{-qp x} q p x (1 - e^{-p^2 x}), \end{aligned}$$

$$\begin{aligned} \mathbb{P}(D \cap E) &:= \mathbb{P}(\text{Po}(\alpha q) = 1)\mathbb{P}(\text{Po}(\alpha qp) \geq 1)\mathbb{P}(\text{Po}(\alpha p^2) = 1) \\ &= e^{-q x} q x (1 - e^{-qp x}) e^{-p^2 x} p^2 x, \end{aligned}$$

$$\mathbb{P}(E \cap F) := \mathbb{P}(\text{Po}(\alpha q) = 1)\mathbb{P}(\text{Po}(\alpha p) = 1) = e^{-p x} p x e^{-q x} q x.$$

另外三項以上的交集就不列出來了，依此類推並結合 maple 的計算就能得到。

#### Theorem 4.4.

(1) 若  $\frac{\ln p}{\ln q}$  為非算術型，當  $n \rightarrow \infty$  時，

$$\mathbb{E}K_n^{(3)} = \frac{n}{H} (1 - C(p)) + o(n).$$

$$\begin{aligned} \text{其中 } C(p) &= \frac{2p^4 q^4}{1 - 3pq + 2p^2 q^2} - \frac{1}{2} \frac{q^2 p^2 (1 + 2pq)}{9pq(q^2 + p^2) + 2[(q^2 + p^2) + 5p^2 q^2]} + qp \\ &+ \frac{p^4 q^4 (2 - pq)}{(1 - pq)^2} + \frac{2pq}{(p^2 + q^2)^2 + 4pq(1 - pq)} + q^2 \ln q + p^2 \ln p \\ &+ pq \ln q + pq \ln p - q \ln \frac{1}{q} - p \ln \frac{1}{p} + q^2 p + p^2 q \\ &- \frac{p^4 (p + 2q) q^2}{(q + 2p)(1 - pq)} - \frac{q^4 (q + 2p) p^2}{(p + 2q)(1 - pq)} + \frac{q^3 p}{(q + 1)(q^2 + pq + p)} \\ &+ \frac{p^3 q}{(p + 1)(p^2 + pq + q)} - \frac{p^3 q^2}{(q + pq + p^2)^2} - \frac{q^3 p}{(p + pq + q^2)^2} \\ &+ \frac{q^2 p^4 (1 + 2pq + q^2)}{(q + 2p)^2} + \frac{p^2 q^4 (1 + 2pq + p^2)}{(p + 2q)^2}. \end{aligned}$$

(2) 若  $\frac{\ln p}{\ln q}$  為  $d$ -算術型，當  $n \rightarrow \infty$  時，

$$\mathbb{E}K_n^{(3)} = \frac{n}{H}(1 - C(p)) + \frac{n}{H}(\psi_W(\ln n) - \psi_{M(1,2)}(\ln n)) + o(n),$$

其中  $C(p)$  與 (1) 相同，而

$$\psi_{M(1,2)}(t) = \sum_{m \neq 0} f^*(-1 + \chi_m) e^{-\chi_m t}, \quad \chi_m = -\frac{2\pi i m}{d}.$$

由於  $f^*$  所佔的篇幅過大，所以我們把計算結果放在附錄。





# Chapter 5

## 結論

我們此篇文章主要在證明更新理論的一些定理，並舉出一些更新理論在 trie 裡的應用，也補足了 Svante Janson 在他的文章裡提出的一些結果的證明，並用這些方法在最後一章算出了 2-protected nodes 和 3-protected nodes 的期望值。

2-protected nodes 的期望值已經有別的方式可以求出，此篇文章是提供另一個方式來求期望值，我們先把目標的節點之充要條件找出來，然後利用這充要條件給出一個機率式子，最後將這式子使用關鍵更新定理算出期望值，我們也用些方法更進一步的算出 3-protected nodes 的期望值，事實上也能用在  $k$ -protected nodes，只是計算量會相當地大。

而這個方法在求期望值的時候很方便，但目前還未發現此方法在求變異數上如何應用，也未證明更新理論無法求變異數，另外在求 3-protected nodes 的期望值時，我們也注意到在排容原理的計算部分，需要借助 maple 幫忙計算，如何推廣到  $k$ -protected nodes 也是未來可能的研究方向。

## 附錄

Theorem 4.4 的  $d$ -算術型 裡的計算結果

$$\begin{aligned}
 f^*(s) = & -\frac{1}{4} \frac{1}{(q+1)^2(q-2)^2} \Gamma(s+1) (16 - q^{-s+6} s 2^{-s} + q^{-s+4} s 2^{-s} - q^6 s 2^{-s} + q^4 s 2^{-s} - \\
 & 8 q^8 e^{-(q-1)^2} (-q^2+1+q)^{-3-s} + 32 q^7 e^{-(q-1)^2} (-q^2+1+q)^{-3-s} - 16 q^6 e^{-(q-1)^2} (-q^2+ \\
 & 1+q)^{-3-s} + 16 (-1)^{s+1} q^2 (q-2)^{-s-2} + 16 q e^{-(q-1)^2} (1-q^2)^{-s} + 36 q^3 e^{-(q-1)^2} (1- \\
 & q^2)^{-s} - 8 q^4 e^{-(q-1)^2} (1-q^2)^{-s} + 8 q^4 s (1-q^2)^{-s} - 16 q s (1-q^2)^{-s} - 48 q^3 s (-q^2+1+ \\
 & 2q)^{-3-s} + 16 q^4 s^2 (-q^2+1+2q)^{-3-s} + 48 q^4 s (-q^2+1+2q)^{-3-s} - 32 q^4 e^{-(q-1)^2} (-q^2+ \\
 & 1+2q)^{-3-s} + 4 (-1)^s q^6 (q^2-1-q)^{-s-2} - 24 q^3 s (1-q)^{-s-2} + 120 q^4 s (1- \\
 & q)^{-s-2} - 64 q^2 s (1-q)^{-s-2} - 36 q^3 s (1-q^2)^{-s} - 48 q^3 e^{-(q-1)^2} s - 84 q^4 e^{-(q-1)^2} s - \\
 & 120 q^4 e^{-(q-1)^2} (1-q)^{-s-2} + 24 q^3 e^{-(q-1)^2} (1-q)^{-s-2} + 64 q^2 e^{-(q-1)^2} (1-q)^{-s-2} + \\
 & 32 (q+1)^{-3-s} q^4 e^{-(q-1)^2} - 84 (q+1)^{-3-s} q^5 s + 56 (q+1)^{-3-s} q^5 e^{-(q-1)^2} - 32 (q+ \\
 & 1)^{-3-s} q^3 e^{-(q-1)^2} + 4 (-1)^s q^7 (q-2)^{-s-2} + 40 (-1)^s q^7 (q-1)^{-s-2} + 60 q^5 s^2 (2- \\
 & q)^{-3-s} + 180 q^5 s (2-q)^{-3-s} - 12 q^4 s^2 (2-q)^{-3-s} - 36 q^4 s (2-q)^{-3-s} + 48 q^2 s (2- \\
 & q)^{-3-s} + 32 q^3 e^{-(q-1)^2} (-q^2+1+q)^{-3-s} - 64 q^5 e^{-(q-1)^2} (-q^2+1+q)^{-3-s} - \\
 & 32 q^2 e^{-(q-1)^2} (-q^2+1+q)^{-3-s} - 16 q^3 s^2 (-q^2+1+q)^{-3-s} - 48 q^3 s (-q^2+1+ \\
 & q)^{-3-s} - 28 q^4 s^2 (-q^2+1+q)^{-3-s} - 84 q^4 s (-q^2+1+q)^{-3-s} - 16 q^2 (q^2+1)^{-s-2} s + \\
 & 16 q e^{-(q-1)^2} (-q^2+1+q)^{-s-2} - 28 (q+1)^{-3-s} q^5 s^2 + 28 (-1)^s q^4 (q-2)^{-s-2} + \\
 & 28 (-1)^s q^3 (q^2-1-q)^{-s-2} - 16 q^3 s^2 (-q^2+1+2q)^{-3-s} + 96 q^3 s (2-q^2)^{-3-s} - \\
 & 64 q^3 e^{-(q-1)^2} (2-q^2)^{-3-s} + 32 q^3 s^2 (2-q^2)^{-3-s} - 16 q^2 s^2 (2-q^2)^{-3-s} - 48 q^2 s (2- \\
 & q^2)^{-3-s} - 16 q^3 e^{-(q-1)^2} s^2 + 4 (-1)^{s+1} q^5 (q-2)^{-s-2} + 12 (-1)^{s+1} q^5 (q^2-1-q)^{-s-2} + \\
 & 4 (-1)^{s+1} q^4 (q^2-1-q)^{-s-2} + 48 (-1)^{s+1} q^5 (q-1)^{-s-2} + 48 (-1)^{s+1} q^6 (q-1)^{-s-2} + \\
 & 48 q^2 s (-q^2+1+q)^{-3-s} + 3 q^{-s+5} s 2^{-s} - 7 q^{-s+3} s 2^{-s} + 4 q^{-s+6} e^{-(q-1)^2} s - 12 q^{-s+5} e^{-(q-1)^2} s - \\
 & 4 q^{-s+4} e^{-(q-1)^2} s + 28 q^{-s+3} e^{-(q-1)^2} s + 48 q^2 e^{-(q-1)^2} s - 60 q^5 s^2 (2-q^2)^{-3-s} + \\
 & 8 (-1)^{s+1} q^8 (q-1)^{-s-2} + 32 q^5 s^2 (-q^2+1+q)^{-3-s} + 96 q^5 s (-q^2+1+q)^{-3-s} - \\
 & 16 q^2 s (2-q)^{-s-2} - 180 q^5 s (2-q^2)^{-3-s} - 24 q^4 e^{-(q-1)^2} (2-q^2)^{-3-s} + 36 q^4 s (2- \\
 & q^2)^{-3-s} + 12 q^4 s^2 (2-q^2)^{-3-s} + 32 q^2 e^{-(q-1)^2} (2-q^2)^{-3-s} + 16 (q+1)^{-s-2} q^2 s - \\
 & 16 q^{-s+1} e^{-(q-1)^2} s + 4 q s 2^{-s} - 48 q^2 e^{-(q-1)^2} (1-q^2)^{-s} + 48 q^2 s (1-q^2)^{-s} + 32 q e^{-(q-1)^2} (-q^2+ \\
 & q)^{-s} - 4 e^{-(q-1)^2} s (-q(q-2))^{-s} + 28 (2+q^2-2q)^{-s-2} q^3 s + 28 (q+1)^{-s-2} q^3 s + 16 (q+ \\
 & 1)^{-3-s} q^3 s^2 + 32 q^2 s (q^2+q)^{-s} - 56 q^5 e^{-(q-1)^2} (-q^2+1+2q)^{-3-s} + 84 q^5 s (-q^2+ \\
 & 1+2q)^{-3-s} + 16 (2+q^2-2q)^{-s-2} q^2 s + 48 (q+1)^{-3-s} q^3 s - 48 (q+1)^{-3-s} q^4 s - \\
 & 16 (q+1)^{-3-s} q^4 s^2 - 16 q s (q^2+q)^{-s} - 4 q s (-q(q-2))^{-s} - 12 q^2 s (-q(q-2))^{-s} -
 \end{aligned}$$

$$\begin{aligned}
& 28q^4e^{-(q-1)^2}s^2 - 120q^5e^{-(q-1)^2}(2-q)^{-3-s} - 32q^2e^{-(q-1)^2}(2-q)^{-3-s} + 120q^5e^{-(q-1)^2}(2- \\
& q^2)^{-3-s} - 32q^3s^2(2-q)^{-3-s} + 56q^4e^{-(q-1)^2}(-q^2+1+q)^{-3-s} - 28(-q+1+ \\
& q^2)^{-s-2}q^3s + 16(-q+1+q^2)^{-s-2}qs + 32qs(1-q)^{-s-2} - 96q^3s(2-q)^{-3-s} + \\
& 64q^3e^{-(q-1)^2}(2-q)^{-3-s} + 24q^4e^{-(q-1)^2}(2-q)^{-3-s} + 8q^2e^{-(q-1)^2}(-q(q-2))^{-s} - \\
& 16(2+q^2-2q)^{-s-2}qs - 16q(q+1)^{-s-2}s + 4q^{-s+1}s2^{-s} + 16q^2e^{-(q-1)^2}s^2 + \\
& 16q^2s^2(2-q)^{-3-s} + 16q^2s^2(-q^2+1+q)^{-3-s} + 28q^5s^2(-q^2+1+2q)^{-3-s} + \\
& 32q^3e^{-(q-1)^2}(-q^2+1+2q)^{-3-s} - 4q^4e^{-(q-1)^2}(-q(q-2))^{-s} - 16qs(-q^2+1+ \\
& q)^{-s-2} + 3q^5s2^{-s} - 7q^3s2^{-s} - 4(-q+1+q^2)^{-s-2}q^6s + 12(-q+1+q^2)^{-s-2}q^5s + \\
& 4(-q+1+q^2)^{-s-2}q^4s - 48q^5s(1-q)^{-s-2} - 12q^9s(2-q)^{-3-s} + 60q^8s(2-q)^{-3-s} - \\
& 72q^7s(2-q)^{-3-s} - 72q^6s(2-q)^{-3-s} + 8q^9e^{-(q-1)^2}(2-q)^{-3-s} - 40q^8e^{-(q-1)^2}(2- \\
& q)^{-3-s} + 48q^7e^{-(q-1)^2}(2-q)^{-3-s} + 48q^6e^{-(q-1)^2}(2-q)^{-3-s} - 4(2+q^2-2q)^{-s-2}q^7s + \\
& 16(2+q^2-2q)^{-s-2}q^6s - 8(2+q^2-2q)^{-s-2}q^5s - 32(2+q^2-2q)^{-s-2}q^4s - 4q^7(q+ \\
& 1)^{-s-2}s + 16q^6(q+1)^{-s-2}s - 8q^5(q+1)^{-s-2}s - 32q^4(q+1)^{-s-2}s + 4q^8e^{-(q-1)^2}s^2 - \\
& 16q^7e^{-(q-1)^2}s^2 + 8q^6e^{-(q-1)^2}s^2 + 32q^5e^{-(q-1)^2}s^2 - 4q^9s^2(2-q)^{-3-s} + 20q^8s^2(2- \\
& q)^{-3-s} - 24q^7s^2(2-q)^{-3-s} - 24q^6s^2(2-q)^{-3-s} - 4q^9s^2(-q^2+1+2q)^{-3-s} + \\
& 16q^8s^2(-q^2+1+2q)^{-3-s} - 8q^7s^2(-q^2+1+2q)^{-3-s} - 32q^6s^2(-q^2+1+ \\
& 2q)^{-3-s} + 8q^9e^{-(q-1)^2}(-q^2+1+2q)^{-3-s} - 32q^8e^{-(q-1)^2}(-q^2+1+2q)^{-3-s} + \\
& 16q^7e^{-(q-1)^2}(-q^2+1+2q)^{-3-s} + 64q^6e^{-(q-1)^2}(-q^2+1+2q)^{-3-s} + 28q^3s(-q^2+ \\
& 1+q)^{-s-2} + 4q^7(q^2+1)^{-s-2}s - 12q^6(q^2+1)^{-s-2}s - 4q^5(q^2+1)^{-s-2}s + 28q^4(q^2+ \\
& 1)^{-s-2}s - 12q^9s(-q^2+1+2q)^{-3-s} + 48q^8s(-q^2+1+2q)^{-3-s} - 24q^7s(-q^2+ \\
& 1+2q)^{-3-s} - 96q^6s(-q^2+1+2q)^{-3-s} + 12q^8e^{-(q-1)^2}s - 48q^7e^{-(q-1)^2}s + \\
& 24q^6e^{-(q-1)^2}s + 96q^5e^{-(q-1)^2}s + 4q^4s(q^2+q)^{-s} - 20q^3s(q^2+q)^{-s} - 8(q+ \\
& 1)^{-3-s}q^9e^{-(q-1)^2} + 32(q+1)^{-3-s}q^8e^{-(q-1)^2} - 16(q+1)^{-3-s}q^7e^{-(q-1)^2} - 64(q+ \\
& 1)^{-3-s}q^6e^{-(q-1)^2} + 12(q+1)^{-3-s}q^9s - 48(q+1)^{-3-s}q^8s + 24(q+1)^{-3-s}q^7s + \\
& 96(q+1)^{-3-s}q^6s + 4(q+1)^{-3-s}q^9s^2 - 16(q+1)^{-3-s}q^8s^2 + 8(q+1)^{-3-s}q^7s^2 + \\
& 32(q+1)^{-3-s}q^6s^2 - 12q^6s(2-q)^{-s-2} + 12q^9s(2-q^2)^{-3-s} - 60q^8s(2-q^2)^{-3-s} + \\
& 72q^7s(2-q^2)^{-3-s} + 72q^6s(2-q^2)^{-3-s} - 8q^9e^{-(q-1)^2}(2-q^2)^{-3-s} + 40q^8e^{-(q-1)^2}(2- \\
& q^2)^{-3-s} - 48q^7e^{-(q-1)^2}(2-q^2)^{-3-s} - 48q^6e^{-(q-1)^2}(2-q^2)^{-3-s} + 4q^9s^2(2-q^2)^{-3-s} - \\
& 20q^8s^2(2-q^2)^{-3-s} + 24q^7s^2(2-q^2)^{-3-s} + 24q^6s^2(2-q^2)^{-3-s} + 8e^{-(q-1)^2}(-q^2+ \\
& q)^{-s}q^4 - 16e^{-(q-1)^2}(-q^2+q)^{-s}q^3 - 24e^{-(q-1)^2}(-q^2+q)^{-s}q^2 + 4q^3s(-q(q- \\
& 2))^{-s} + 8q^4s(-q(q-2))^{-s} + 4q^8s^2(-q^2+1+q)^{-3-s} - 16q^7s^2(-q^2+1+q)^{-3-s} + \\
& 8q^6s^2(-q^2+1+q)^{-3-s} + 12q^8s(-q^2+1+q)^{-3-s} - 48q^7s(-q^2+1+q)^{-3-s} - \\
& 4q^{-2s+4} + 8q^{-2s+3} + 12q^{-2s+2} - 16q^{-2s+1} + q^{-s+4}2^{-s} - q^{-s+6}2^{-s} - q^62^{-s} + \\
& q^42^{-s} + 24q^6s(-q^2+1+q)^{-3-s} - 16q^{-2s} - 16(1-q)^{-s} - 16q^{-s} + 32q^{-s+1} + 24q^3 + \\
& 28q^{-s+2} + 16(1-q)^{-1-s} - 44q^2 - 64q^5 + 60q^4 + 28q^{-s+5} - 4q^{-s+4} - 60q^{-s+3} - \\
& 8q^{-s+6} + 4(-q(q-2))^{-s} - 16(q-1)^{-2s-2} - 32(-q^2+q)^{-s} - 16q^6 + 32q^7 - 8q^8 + \\
& 16q + 24q^7e^{-(q-1)^2}s^2(2-q)^{-3-s} + 24q^6e^{-(q-1)^2}s^2(2-q)^{-3-s} - 4q^6e^{-(q-1)^2}s(-q^2+ \\
& 1+q)^{-s-2} + 4q^4e^{-(q-1)^2}s(-q^2+1+q)^{-s-2} - 28q^3e^{-(q-1)^2}s(-q^2+1+q)^{-s-2} - \\
& 60q^5e^{-(q-1)^2}s^2(2-q)^{-3-s} - 32q^5e^{-(q-1)^2}s^2(-q^2+1+q)^{-3-s} + 8(-1)^{s+1}q^8s(q- \\
& 1)^{-s-2} + 48(-1)^{s+1}q^6s(q-1)^{-s-2} + 40(-1)^{s+1}q^7e^{-(q-1)^2}(q-1)^{-s-2} - 48q^2e^{-(q-1)^2}s(1- \\
& q^2)^{-s} + 16qe^{-(q-1)^2}s(1-q^2)^{-s} - 8q^4e^{-(q-1)^2}s(1-q^2)^{-s} + 36q^3e^{-(q-1)^2}s(1-q^2)^{-s} +
\end{aligned}$$

$$\begin{aligned}
& 40(-1)^s q^7 s (q-1)^{-s-2} + 8(-1)^s q^8 e^{-(q-1)^2} (q-1)^{-s-2} + 48(-1)^s q^6 e^{-(q-1)^2} (q-1)^{-s-2} + 8q^2 e^{-(q-1)^2} s(-q(q-2))^{-s} + 32(-1)^{s+1} q e^{-(q-1)^2} (q-1)^{-s-2} + 28(-1)^s q^4 s (q-2)^{-s-2} + 48(-1)^s q^5 e^{-(q-1)^2} (q-1)^{-s-2} + 4(-1)^s q^7 s (q-2)^{-s-2} + 4(-1)^s q^6 s (q^2-1-q)^{-s-2} + 12(-1)^s q^5 e^{-(q-1)^2} (q^2-1-q)^{-s-2} + 4(-1)^s q^4 e^{-(q-1)^2} (q^2-1-q)^{-s-2} - 180q^5 e^{-(q-1)^2} s(2-q)^{-3-s} + 32q^3 e^{-(q-1)^2} s^2(2-q)^{-3-s} + 96q^3 e^{-(q-1)^2} s(2-q)^{-3-s} + 12q^4 e^{-(q-1)^2} s^2(2-q)^{-3-s} + 36q^4 e^{-(q-1)^2} s(2-q)^{-3-s} - 16q^2 e^{-(q-1)^2} s^2(-q^2+1+q)^{-3-s} - 48q^2 e^{-(q-1)^2} s(-q^2+1+q)^{-3-s} - 96q^5 e^{-(q-1)^2} s(-q^2+1+q)^{-3-s} + 16q^3 e^{-(q-1)^2} s^2(-q^2+1+q)^{-3-s} + 48q^3 e^{-(q-1)^2} s(-q^2+1+q)^{-3-s} + 28q^4 e^{-(q-1)^2} s^2(-q^2+1+q)^{-3-s} + 84q^4 e^{-(q-1)^2} s(-q^2+1+q)^{-3-s} - 32q e^{-(q-1)^2} s(1-q)^{-s-2} - 120q^4 e^{-(q-1)^2} s(1-q)^{-s-2} + 24q^3 e^{-(q-1)^2} s(1-q)^{-s-2} + 64q^2 e^{-(q-1)^2} s(1-q)^{-s-2} + 16q^2 e^{-(q-1)^2} s^2(2-q^2)^{-3-s} + 48(q+1)^{-3-s} q^4 e^{-(q-1)^2} s + 16(q+1)^{-3-s} q^4 e^{-(q-1)^2} s^2 - 48(q+1)^{-3-s} q^3 e^{-(q-1)^2} s - 96q^3 e^{-(q-1)^2} s(2-q^2)^{-3-s} + 12(-1)^{s+1} q^5 s (q^2-1-q)^{-s-2} + 4(-1)^{s+1} q^4 s (q^2-1-q)^{-s-2} + 4(-1)^{s+1} q^6 e^{-(q-1)^2} (q^2-1-q)^{-s-2} + 28(-1)^{s+1} q^3 e^{-(q-1)^2} (q^2-1-q)^{-s-2} + 84(q+1)^{-3-s} q^5 e^{-(q-1)^2} s + 48q^3 e^{-(q-1)^2} s(-q^2+1+2q)^{-3-s} + 16q^3 e^{-(q-1)^2} s^2(-q^2+1+2q)^{-3-s} - 16q^4 e^{-(q-1)^2} s^2(-q^2+1+2q)^{-3-s} + 4(-1)^{s+1} q^5 s (q-2)^{-s-2} - 16(q+1)^{-3-s} q^3 e^{-(q-1)^2} s^2 + 28(q+1)^{-3-s} q^5 e^{-(q-1)^2} s^2 - 4e^{-(q-1)^2} s(-q(q-2))^{-s} q^4 - 48q^4 e^{-(q-1)^2} s(-q^2+1+2q)^{-3-s} - 84q^5 e^{-(q-1)^2} s(-q^2+1+2q)^{-3-s} - 28q^5 e^{-(q-1)^2} s^2(-q^2+1+2q)^{-3-s} - 16q^2 e^{-(q-1)^2} s^2(2-q)^{-3-s} - 48q^2 e^{-(q-1)^2} s(2-q)^{-3-s} - 12q^8 e^{-(q-1)^2} s(-q^2+1+q)^{-3-s} + 48q^7 e^{-(q-1)^2} s(-q^2+1+q)^{-3-s} - 24q^6 e^{-(q-1)^2} s(-q^2+1+q)^{-3-s} - 4q^8 e^{-(q-1)^2} s^2(-q^2+1+q)^{-3-s} + 16q^7 e^{-(q-1)^2} s^2(-q^2+1+q)^{-3-s} - 8q^6 e^{-(q-1)^2} s^2(-q^2+1+q)^{-3-s} + 12q^9 e^{-(q-1)^2} s(2-q)^{-3-s} - 60q^8 e^{-(q-1)^2} s(2-q)^{-3-s} + 72q^7 e^{-(q-1)^2} s(2-q)^{-3-s} + 72q^6 e^{-(q-1)^2} s(2-q)^{-3-s} - 12(q+1)^{-3-s} q^9 e^{-(q-1)^2} s + 20q^8 e^{-(q-1)^2} s^2(2-q^2)^{-3-s} - 24q^7 e^{-(q-1)^2} s^2(2-q^2)^{-3-s} - 24q^6 e^{-(q-1)^2} s^2(2-q^2)^{-3-s} + 48(q+1)^{-3-s} q^8 e^{-(q-1)^2} s - 24(q+1)^{-3-s} q^7 e^{-(q-1)^2} s - 96(q+1)^{-3-s} q^6 e^{-(q-1)^2} s - 12q^9 e^{-(q-1)^2} s(2-q^2)^{-3-s} + 60q^8 e^{-(q-1)^2} s(2-q^2)^{-3-s} - 72q^7 e^{-(q-1)^2} s(2-q^2)^{-3-s} - 72q^6 e^{-(q-1)^2} s(2-q^2)^{-3-s} - 4q^9 e^{-(q-1)^2} s^2(2-q^2)^{-3-s} - 4(q+1)^{-3-s} q^9 e^{-(q-1)^2} s^2 + 16(q+1)^{-3-s} q^8 e^{-(q-1)^2} s^2 - 8(q+1)^{-3-s} q^7 e^{-(q-1)^2} s^2 - 32(q+1)^{-3-s} q^6 e^{-(q-1)^2} s^2 + 12q^9 e^{-(q-1)^2} s(-q^2+1+2q)^{-3-s} - 48q^8 e^{-(q-1)^2} s(-q^2+1+2q)^{-3-s} + 24q^7 e^{-(q-1)^2} s(-q^2+1+2q)^{-3-s} - 32q^3 e^{-(q-1)^2} s^2(2-q^2)^{-3-s} + 48q^2 e^{-(q-1)^2} s(2-q^2)^{-3-s} - 12q^4 e^{-(q-1)^2} s^2(2-q^2)^{-3-s} - 36q^4 e^{-(q-1)^2} s(2-q^2)^{-3-s} + 180q^5 e^{-(q-1)^2} s(2-q^2)^{-3-s} + 60q^5 e^{-(q-1)^2} s^2(2-q^2)^{-3-s} + 96q^6 e^{-(q-1)^2} s(-q^2+1+2q)^{-3-s} + 4q^9 e^{-(q-1)^2} s^2(-q^2+1+2q)^{-3-s} - 16q^8 e^{-(q-1)^2} s^2(-q^2+1+2q)^{-3-s} + 8q^7 e^{-(q-1)^2} s^2(-q^2+1+2q)^{-3-s} + 32q^6 e^{-(q-1)^2} s^2(-q^2+1+2q)^{-3-s} + 4q^9 e^{-(q-1)^2} s^2(2-q)^{-3-s} - 20q^8 e^{-(q-1)^2} s^2(2-q)^{-3-s} + 16(-1)^s q e^{-(q-1)^2} s(q^2-1-q)^{-s-2} + 40(-1)^{s+1} q^7 e^{-(q-1)^2} s(q-1)^{-s-2} + 8(-1)^s q^8 e^{-(q-1)^2} s(q-1)^{-s-2} + 48(-1)^s q^6 e^{-(q-1)^2} s(q-1)^{-s-2} + 48(-1)^s q^5 e^{-(q-1)^2} s(q-1)^{-s-2} + 12(-1)^s q^5 e^{-(q-1)^2} s(q^2-1-q)^{-s-2} + 32q^2 e^{-(q-1)^2} - 16q(1-q)^{-1-s} + 32q(1-q)^{-s-2} - 64q^2(1-q)^{-s-2} + 120q^4(1-q)^{-s-2} - 24q^3(1-q)^{-s-2} + 16q^3 s^2 - 32q^3 e^{-(q-1)^2} + 24q^{-s+5} s + 8q^{-s+4} s - 56q^{-s+3} s - 7q^{-s+3} 2^{-s} + 4q^{-s+6} e^{-(q-1)^2} - 12q^{-s+5} e^{-(q-1)^2} - 4q^{-s+4} e^{-(q-1)^2} + 28q^{-s+3} e^{-(q-1)^2} - 8q^{-s+6} s + 3q^{-s+5} 2^{-s} + 16(q+1)^{-s-2} q^2 - 16q(1-q^2)^{-s} - 4q(-q(q-2))^{-s} - 36q^3(1-q^2)^{-s} + 8q^4(1-q^2)^{-s} + 64q^3(2-q^2)^{-3-s} - 32q^2(2-q^2)^{-3-s} - 120q^5(2-q^2)^{-3-s} + 16(q-1)^{-2s-2} q - 32q^3(-q^2+1+2q)^{-3-s} +
\end{aligned}$$

$$\begin{aligned}
& 32q^4(-q^2+1+2q)^{-3-s} + 56q^5(-q^2+1+2q)^{-3-s} - 16q(q^2+q)^{-s} + 28(2+q^2-2q)^{-s-2}q^3 - 12q^2(-q(q-2))^{-s} + 32q^2(q^2+q)^{-s} + 4s(-q(q-2))^{-s} - \\
& 4e^{-(q-1)^2}(-q(q-2))^{-s} + 16(2+q^2-2q)^{-s-2}q^2 + 4q^3(-q(q-2))^{-s} - 4(1-q)^{-s}q^4 + 8(1-q)^{-s}q^3 + 12(1-q)^{-s}q^2 - 16(1-q)^{-s}q + 8q^4(-q(q-2))^{-s} + 24(-q^2+q)^{-s}q^2 + 3q^52^{-s} - 7q^32^{-s} - 4(q-1)^{-2s-2}q^6 + 16(q-1)^{-2s-2}q^5 - 8(q-1)^{-2s-2}q^4 - \\
& 32(q-1)^{-2s-2}q^3 - 8q^4(-q^2+q)^{-s} + 16q^3(-q^2+q)^{-s} - 4(q+1)^{-s-2}q^7 + 16(q+1)^{-s-2}q^6 - 8(q+1)^{-s-2}q^5 - 32(q+1)^{-s-2}q^4 - 12q^6(2-q)^{-s-2} + 8(q+1)^{-3-s}q^9 - \\
& 32(q+1)^{-3-s}q^8 + 16(q+1)^{-3-s}q^7 + 64(q+1)^{-3-s}q^6 + 8q^8e^{-(q-1)^2} - 32q^7e^{-(q-1)^2} + 16q^6e^{-(q-1)^2} + 64q^5e^{-(q-1)^2} + 4q^6(1-q)^{-1-s} - 16q^5(1-q)^{-1-s} + 8q^4(1-q)^{-1-s} + \\
& 32q^3(1-q)^{-1-s} + 4q^7(q^2+1)^{-s-2} - 12q^6(q^2+1)^{-s-2} - 4q^5(q^2+1)^{-s-2} + 28q^4(q^2+1)^{-s-2} + 8q^9(2-q^2)^{-3-s} - 40q^8(2-q^2)^{-3-s} + 48q^7(2-q^2)^{-3-s} + 48q^6(2-q^2)^{-3-s} - \\
& 8q^9(-q^2+1+2q)^{-3-s} + 32q^8(-q^2+1+2q)^{-3-s} - 16q^7(-q^2+1+2q)^{-3-s} - 64q^6(-q^2+1+2q)^{-3-s} - 4(2+q^2-2q)^{-s-2}q^7 + 16(2+q^2-2q)^{-s-2}q^6 - 8(2+q^2-2q)^{-s-2}q^5 - \\
& 32(2+q^2-2q)^{-s-2}q^4 - 4(-q+1+q^2)^{-s-2}q^6 + 12(-q+1+q^2)^{-s-2}q^5 + 4(-q+1+q^2)^{-s-2}q^4 + 4q^4(q^2+q)^{-s} - 20q^3(q^2+q)^{-s} - 8q^9(2-q)^{-3-s} + 40q^8(2-q)^{-3-s} - \\
& 48q^7(2-q)^{-3-s} - 48q^6(2-q)^{-3-s} + 8q^8(-q^2+1+q)^{-3-s} - 32q^7(-q^2+1+q)^{-3-s} + 16q^6(-q^2+1+q)^{-3-s} + 16(-q+1+q^2)^{-s-2}q - 28(-q+1+q^2)^{-s-2}q^3 - \\
& 32q(-q^2+q)^{-s} + 32e^{-(q-1)^2}(-q^2+q)^{-s} + 84q^4s - 16q^{-s+1}e^{-(q-1)^2} - 16(2+q^2-2q)^{-s-2}q - 12q^8s + 48q^7s - 24q^6s - 96q^5s - 4q^8s^2 + 16q^7s^2 - 8q^6s^2 - 32q^5s^2 - \\
& 16q^2s^2 + 48q^2(1-q^2)^{-s} + 24q^4(2-q^2)^{-3-s} - 16q(q+1)^{-s-2} + 32q^2(2-q)^{-3-s} - 64q^3(2-q)^{-3-s} + 120q^5(2-q)^{-3-s} - 24q^4(2-q)^{-3-s} - 32q^3(-q^2+1+q)^{-3-s} - \\
& 56q^4(-q^2+1+q)^{-3-s} + 32q^2(-q^2+1+q)^{-3-s} + 64q^5(-q^2+1+q)^{-3-s} - 48q^2s - 16q^2(q^2+1)^{-s-2} - 16q(-q^2+1+q)^{-s-2} + 28(q-1)^{-2s-2}q^2 + 28(q+1)^{-s-2}q^3 + \\
& 28q^4s^2 + 32q^{-s+1}s + 4q^{-s+1}2^{-s} + 4q2^{-s} - 56(q+1)^{-3-s}q^5 + 32(q+1)^{-3-s}q^3 - 32(q+1)^{-3-s}q^4 + 48q^3s - 56q^4e^{-(q-1)^2} - 28q^2(1-q)^{-1-s})/(q+1)^2(q-2)^2
\end{aligned}$$

# Bibliography

- [1] D. Blackwell (1948). A renewal theorem, *Duke Math. J.*, **15**, 145-151.
- [2] K.-L. Chung. *A Course in Probability Theory*, 2nd edition, Academic Press, 2000.
- [3] D. R. Cox. *Renewal Theory*, 1st edition, Methuen & Co., 1970.
- [4] L. Devroye (1991). Limit laws for local counters in random binary search trees, *Random Struct. Algor.*, **2:3**, 303-315.
- [5] W. Feller. *An Introduction to Probability Theory and Its Applications*, vol. 1, 3rd edition, Wiley, 1968.
- [6] A. Gut. *Stopped Random Walks: Limit Theorems and Applications*, 2nd edition, Springer, 2009.
- [7] S. Janson (2006). Rounding of continuous random variables and oscillatory asymptotics, *Ann. Probab.*, **34:5**, 1807-1826.
- [8] S. Janson (2012). Renewal theory in the analysis of tries and strings, *Theor. Comput. Sci.*, **416**, 33-54.
- [9] S. Karlin. *A First Course in Stochastic Processes*, 2nd edition, Academic Press, 1975.
- [10] H. M. Mahmoud. *Evolution of Random Search Trees*, 1st edition, Wiley, 1991.
- [11] S. Ross. *Stochastic Processes*, 2nd edition, Wiley, 1995.
- [12] 黃文璋 (1995)。《隨機過程》。臺北市: 華泰書局。
- [13] 程雋 (2003)。《應用機率》。臺北市: 文笙。
- [14] 溫惠美 (2002)。《資料結構理論與應用—使用 C 語言》。臺北市: 博碩文化。