نمایهٔ تِرای‌های سطلی

مهری جوانیان
دانشگاه زنجان، دانشکدهٔ علوم، گروه آمار

دریافت ۹۶/۰۵/۱۵ پذیرش ۹۷/۰۸/۰۷

چکیده

تِرای‌ها یکی از کاربردی‌ترین ساختمان داده‌ها با ساختار درختی در علوم کامپیوتر هستند. تِرای‌ها، داده‌های رشته‌ای را در برگ‌های درخت ذخیره می‌کنند. یک نسخهٔ تعمیم یافتهٔ تِرای، موسوم به تِرای سطلی است که در آن هر برگ یا سطل، ظرفیت ذخیرهٔ بیش از یک داده را دارد. تِرای تصادفی با تعریف یک قاعده رشد تصادفی برای تِرای حاصل می‌شود. تعداد گره‌های هم نوع که در فاصلهٔ یک‌سان از ریشهٔ یک درخت ریشه دار هستند را نمایه نامند. بررسی نمایهٔ یک درخت، اهمیت زیادی دارد. زیرا بسیاری از پارامترهای درخت ریشه‌دار را می‌توان برحسب نمایهٔ آن درخت بیان کرد. در این مقاله به بررسی مجانبی امیدریاضی، واریانس و توزیع حدی هر یک از دو نمایهٔ سطلی و داخلی (تعداد گره‌های سطلی یا برگ و تعداد گره‌های داخلی یا غیربرگ که در فاصلهٔ یک‌سان از ریشه هستند) در تِرای سطلی تصادفی می‌پردازیم، وقتی که تعداد داده‌های ذخیره شده در تِرای افزایش یابد. امید ریاضی و واریانس‌های هر دو نمایه شامل توابعی متناوب هستند و نشان می‌دهیم آن توابع متناوب ناصفرند که این نکته در مقاله مربوط به نمایه ترای معمولی، به اثبات نرسیده است. هم‌چنین به بررسی مقدار مجانبی نسبت امید ریاضی‌های دو نمایهٔ سطلی و داخلی می‌پردازیم. روش‌هایی که برای حصول نتایج به‌کار می‌بریم، براساس استفاده از پواسونی سازی، تبدیل ملین، معادلات بازگشتی، توابع مولد، تحلیل تکینی و روش نقطهٔ زینی است.

واژه‌های کلیدی: تِرای‌های سطلی، نمایه، پواسونی سازی، تبدیل ملین، معادلات بازگشتی، توابع مولد، تحلیل تکینی، روش نقطهٔ زینی.

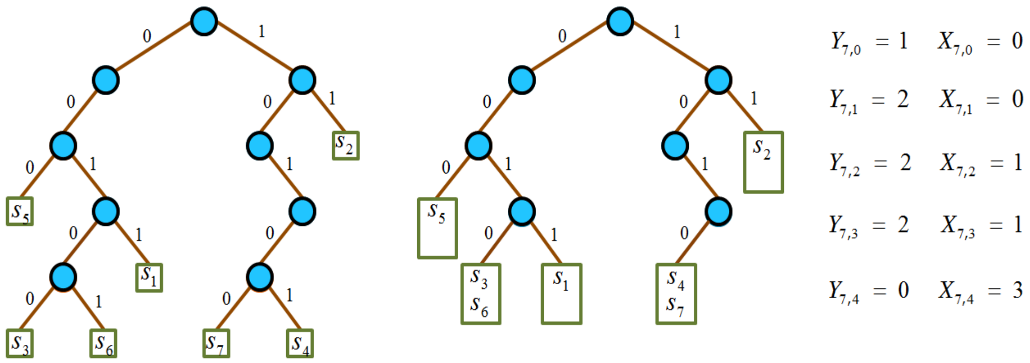
مقدمه

یکی از اساسی‌ترین الگوریتم‌ها برای ذخیره و جستجوی داده‌های رشته‌ای، ساختاری درختی به‌نام ترای است که کارایی بیش‌تری نسبت به درخت جستجوی دودویی دارد (برای جزئیات بیش‌تر [۷] را ببینید). برای سادگی، رشته‌های دودویی (دنباله‌هایی از ''0'' و ''1'') را درنظر می گیریم. در ترای، رشته‌ها در برگ‌ها (گره‌های خارجی) ذخیره می‌شوند. یک ترای، n رشتهٔ 1-0 ایی را بدین‌صورت در خود ذخیره می‌کند: اگر $n = 1$ آن‌گاه فقط یک رشته داریم و در ریشهٔ ترای ذخیره می‌شود که در این حالت ریشهٔ ترای یک برگ یا گره خارجی است و ترای شامل هیچ گرهٔ دیگری نیست؛ اگر $n \geq 2$ آن‌گاه ریشهٔ ترای یک گره داخلی (گره تهی) است که ذخیره‌کنندهٔ هیچ رشته‌ای نیست و رشته‌ها وابسته به این که اولین رقم‌شان ''0'' (''1'') باشد در زیردرخت چپِ (راستِ) ریشه ذخیره می‌شوند. با همین قاعده و با حذف کردن اولین رقم رشته‌ها، هر یک از زیردرخت‌های ریشه ساخته می‌شود. برای مثال ترای سمت چپ در شکل ۱ ملاحظه شود.

برای افزایش سرعت پردازش اطلاعات و عمل جستجو در الگوریتم ترای، وابسته به اختیار پردازشگر حداکثر تعداد مشخصی از داده‌ها را در هر برگ ذخیره می کنند و نه فقط یک داده در یک برگ (ترای معمولی). در این مقاله به بررسی نمایه این نوع ترای‌ها می‌پردازیم.

*نویسنده مسئول javanian@znu.ac.ir

[DOR: 20.1001.1.25882546.1399.6.2.4.5]
[DOI: 10.52547/mmr.6.2.169]



شکل ۱. نمایش یک ترای (درخت سمت چپ) و ترای سطلی (درخت سمت راست) ساخته شده از ۷ رشته ($\mathcal{S}_1 = \mathbf{0011}\ldots, \mathcal{S}_2 = \mathbf{11}\ldots, \mathcal{S}_3 = \mathbf{0010}\ldots, \cdots$) با $b = 2$. هم چنین گره‌های داخلی با دایره‌ها (در هر دو ترای)، گره‌های خارجی در ترای با مربع‌ها و در ترای سطلی با مستطیل‌ها (سطل‌ها) نشان داده شده است.

هرگاه در برگ‌های یک ترای بتوان به تعداد $b \geq 1$ رشته ذخیره کرد، آن‌گاه ترای ساخته شده را ترای سطلی نامند. برای مثال ترای سطلی با $b = 2$ (ترای سمت راست) در شکل ۱ ملاحظه شود.

رشتهٔ 1-0 ایی که دنباله‌ای از متغیرهای تصادفی برنولی با احتمال پیروزی $0 < p < 1$ (احتمال رخ دادن ''1'') و احتمال شکست $q := 1 - p \leq p$ باشد را رشتهٔ تصادفی می‌نامیم.

ترای ساخته شده از رشته‌های تصادفی را ترای تصادفی گویند. پارامترهای ترای تصادفی به‌طور گسترده بررسی شده است (برای اطلاعات بیش‌تر [۱]، [۵]، [۶] و [۱۰] را ببینید).

تعداد گره‌های هم‌نوع که در فاصلهٔ یک‌سان از ریشهٔ یک درخت قرار دارند را نمایهٔ آن درخت می‌نامند. چون از نمایهٔ درخت می‌توان مستقیماً نتایجی دربارهٔ پارامترهای دیگر درخت مثل ارتفاع، کوتاه‌ترین مسیر، عمق و طول مسیر کل به‌دست آورد؛ بنابراین بررسی نمایهٔ درخت اهمیت زیادی دارد. به طور مثال نمایهٔ درخت جستجوی دودویی و بازگشتی، درخت جستجوی رقمی، ترای و پاتریشیا ترای، به ترتیب در [۴]، [۲]، [۱۰] و [۸] بررسی شده‌اند.

در این مقاله به تعمیم نتایج به‌دست آمده از نمایهٔ ترای تصادفی در [۱۰] برای ترای سطلی تصادفی می‌پردازیم. یعنی به تحلیل رفتار دو نوع نمایه در ترای سطلی تصادفی می‌پردازیم که به‌وسیلهٔ n رشته بنا شده و ظرفیت هر سطل حداکثر $b \geq 1$ رشتهٔ تصادفی است (تعمیم حالت $b = 1$ در [۱۰]): نمایهٔ سطلی $X_{n,k}$ (تعداد گره‌های خارجی یا سطل‌ها در فاصلهٔ k از ریشه)؛ نمایهٔ داخلی $Y_{n,k}$ (تعداد گره‌های داخلی در فاصلهٔ k از ریشه). به‌طور مثال مقادیر این دو نوع نمایه را که برای ترای سطلی در شکل ۱ محاسبه شده است، ملاحظه شود.

ترتیب مطالب ارایه شده در این مقاله بدین‌صورت است. در بخش بعد نشان می‌دهیم که توابع مولد احتمال $X_{n,k}$ و $Y_{n,k}$ به‌عنوان تابعی از u ، در معادلهٔ بازگشتی

$$z_{n,k}(u) = \sum_{0 \leq j \leq n} \binom{n}{j} p^j q^{n-j} z_{j,k-1}(u) z_{n-j,k-1}(u), \qquad n \geq b+1, \ \ k \geq 1,$$

صدق می‌کنند. بنابراین امیدریاضی و واریانس $X_{n,k}$ و $Y_{n,k}$ در معادلهٔ بازگشتی

$$t_{n,k} = \sum_{0 \leq j \leq n} \binom{n}{j} p^j q^{n-j} (t_{j,k-1} + t_{n-j,k-1}), \qquad n \geq b+1, \ \ k \geq 1, \tag{۱}$$

صدق می‌کنند. در بخش‌های سوم و چهارم به‌منظور به‌دست آوردن امیدریاضی و واریانس مجانبی $X_{n,k}$ و $Y_{n,k}$، از روشی استاندارد استفاده می‌کنیم: ابتدا تابع مولد پواسن $f_k(z) := \sum_{n \geq 0} t_{n,k} z^n / n!\, e^{-z}$ را در نظر می‌گیریم. از (۱) نتیجه گرفته می‌شود که $f_k(z)$ در معادلهٔ بازگشتی

$$f_k(z) = f_{k-1}(pz) + f_{k-1}(qz),$$

[DOR: 20.1001.1.25882546.1399.6.2.4.5]
[DOI: 10.52547/mmr.6.2.169]

صدق می‌کند. سپس جواب دقیق و جواب مجانبی این معادله به‌ترتیب، با روش تکرار و تبدیل ملین حاصل می‌شوند (برای آشنایی با این روش [۳] و [۱۱] را ببینید). مرحلهٔ آخر برای برآورده شدن هدفمان، مرحلهٔ برگشت از پواسنی‌سازی است. یعنی از بسط مجانبی $f_k(z)$ (هرگاه $z \to \infty$)؛ بسط مجانبی $t_{n,k}$ (هرگاه $n \to \infty$)، را به‌دست می‌آوریم. در این مرحله به‌کمک تحلیل مختلط و روش نقطهٔ زینی نشان داده می‌شود که $t_{n,k} \sim f_k(n)$ (هرگاه $n \to \infty$).

به‌وسیلهٔ این روش نشان می‌دهیم هرگاه $\varepsilon > 0$ و $(\alpha_1 + \varepsilon)\log n \le k \le (\alpha_2 - \varepsilon)\log n$ که در آن

$$\alpha_1 := \frac{1}{\log(1/q)} \le \alpha_0 := \frac{2}{\log(1/p) + \log(1/q)} \le \alpha_2 := \frac{p^2 + q^2}{p^2\log(1/p) + q^2\log(1/q)},$$

آن‌گاه نتایج مجانبی به‌دست آمده برای $\mathbb{E}(X_{n,k})$ و $\mathbb{E}(Y_{n,k})$، شامل توابعی متناوب است. این پدیده به‌دلیل ظاهر شدن بی‌نهایت نقطهٔ زینی در انتگرال معکوس تبدیل ملین به‌وجود می‌آید که اولین بار در [۹] مشاهده شده است. هم‌چنین وقتی $(\alpha_0 + \varepsilon)\log n \le k \le (\alpha_2 - \varepsilon)\log n$، به‌بررسی مقادیر مجانبی نسبت $\mathbb{E}(X_{n,k})/\mathbb{E}(Y_{n,k})$ می‌پردازیم زیرا در بازهٔ $(\alpha_1 + \varepsilon)\log n \le k \le (\alpha_0 - \varepsilon)\log n$، $\mathbb{E}(Y_{n,k}) = 2^k + \mathcal{O}\left(\mathbb{E}(X_{n,k})\right)$ و با افزایش k، رشد سریع‌تری نسبت به $\mathbb{E}(X_{n,k})$ دارد و در نتیجه نسبت مذکور به صفر میل می‌کند. سپس به‌ازای k در بازهٔ $(\alpha_1 + \varepsilon)\log n \le k \le (\alpha_2 - \varepsilon)\log n$ به محاسبه واریانس‌های دو نمایه، $\mathbb{V}(X_{n,k})$ و $\mathbb{V}(Y_{n,k})$، می‌پردازیم که شامل توابعی متناوب و دارای مرتبه‌های یک‌سان با میانگین‌های دو نمایه هستند.

در بخش آخر ثابت می‌کنیم که هرگاه $\mathbb{E}(X_{n,k}) \to \infty$ و $\mathbb{E}(Y_{n,k}) \to \infty$ آن‌گاه توزیع‌های حدی $X_{n,k}$ و $Y_{n,k}$ نرمال هستند وقتی k در بازهٔ $(\alpha_1 + \varepsilon)\log n \le k \le (\alpha_2 - \varepsilon)\log n$ باشد.

در این مقاله، برای هر k بازهٔ $(\alpha_1 + \varepsilon)\log n \le k \le (\alpha_2 - \varepsilon)\log n$ را در نظر می‌گیریم. بنابراین بررسی نمایه‌ها را فقط در ترای غیرمتقارن ($p \neq q$) بررسی می‌کنیم. زیرا اگر $p = q$ آن‌گاه $\alpha_1 = \alpha_0 = \alpha_2$ و بازهٔ شامل نقاط زینی تهی است و بنابراین هیچ تابع متناوبی در میانگین‌ها و واریانس‌های دو نمایه ظاهر نمی‌شود.

معادلهٔ بازگشتی برحسب تابع مولد احتمال نمایه‌ها

در یک ترای سطلی با n رشته، $X_{n,k}$ هم توزیع با مجموع تعداد گره‌های خارجی در زیردرخت‌های چپ و راست ریشهٔ ترای سطلی است که در فاصلهٔ $k-1$ از ریشهٔ آن زیردرخت‌ها قرار دارند. یعنی:

$$X_{n,k} \underline{\underline{d}} X_{B_n,k-1} + X^*_{n-B_n,k-1}, \qquad n \ge b+1, \quad k \ge 1,$$

که در آن $B_n \underline{\underline{d}} Binomial(n,p), X_{n,k} \underline{\underline{d}} X^*_{n,k}$ و $X_{n,k}$، $X^*_{n,k}$ و B_n مستقل هستند. هم‌چنین به‌ازای $k \ge 0$، $X_{0,k} = 0$؛ به‌ازای $1 \le n \le b$، $X_{n,k} = 0$ ($k \ge 1$) و $X_{n,0} = 1$؛ به‌ازای $n \ge b+1$، $X_{n,0} = 0$.

معادلهٔ بالا برای $Y_{n,k}$ با شرایط اولیه متفاوت: به‌ازای $k \ge 0$، $Y_{0,k} = 0$؛ به‌ازای $1 \le n \le b$ و $k \ge 0$، $Y_{n,k} = 0$؛ و به‌ازای $n \ge b+1$، $Y_{n,0} = 1$ برقرار است.

فرض کنید که $P^{[X]}_{n,k}(u)$ و $P^{[Y]}_{n,k}(u)$ به‌ترتیب توابع مولد احتمال $X_{n,k}$ و $Y_{n,k}$ باشند. از معادلهٔ بالا نتیجه گرفته می‌شود که $P^{[X]}_{n,k}(u)$ در معادله بازگشتی

$$P^{[X]}_{n,k}(u) = \sum_{0 \le j \le n} \binom{n}{j} p^j q^{n-j} P^{[X]}_{j,k-1}(u) P^{[X]}_{n-j,k-1}(u), \qquad n \ge b+1, \ \ k \ge 1, \tag{۲}$$

$$P^{[X]}_{n,k}(u) = \begin{cases} u, & \text{اگر } 1 \le n \le b, k = 0 \\ 1. & \text{اگر } (n = 0, k \ge 0) \text{ یا } (n \ge b+1, k = 0) \text{ یا } (1 \le n \le b, k \ge 1) \end{cases}$$

[DOR: 20.1001.1.25882546.1399.6.2.4.5]

[DOI: 10.52547/mmr.6.2.169]

زیر صدق می‌کند. معادلهٔ (۲) برای نمایه داخلی $Y_{n,k}$ با شرایط اولیه متفاوت

$$P_{n,k}^{[Y]}(u) = \begin{cases} u, & \text{اگر } n \geq b+1, k=0 \\ 1. & \text{اگر } (n=0, k\geq 0) \text{ یا } (1\leq n\leq b, k\geq 0) \end{cases} \tag{۳}$$

برقرار است. در سراسر مقاله از نمادگذاری‌های زیر برای چند کمیت و تابع خاص استفاده می‌شود. به‌ازای عدد مختلط s، $T(s) := p^{-s} + q^{-s}$ و با حل معادلهٔ ($\alpha \in \mathbb{R}$)

$$\alpha = \frac{p^{-\rho} + q^{-\rho}}{p^{-\rho}\log\frac{1}{p} + q^{-\rho}\log\frac{1}{q}},$$

نسبت به ρ (چون $p > q$، آن‌گاه $\left(\log\frac{1}{q}\right)^{-1} < \alpha < \left(\log\frac{1}{p}\right)^{-1}$ است)، جواب (۴) به‌دست می‌آید:

$$\rho = \rho(\alpha) = \frac{1}{\log(p/q)}\log\frac{1-\alpha\log(1/p)}{\alpha\log(1/q)-1}. \tag{۴}$$

امیدریاضی نمایه‌های سطلی، داخلی و نسبت آن‌ها

در این بخش تقریب‌هایی مجانبی برای امیدریاضی $X_{n,k}$ و $Y_{n,k}$ به‌دست می‌آوریم و نشان می‌دهیم که هر دو غیرصفر هستند. هم‌چنین تحت فرض‌هایی به بررسی مقادیر مجانبی $\mathbb{E}(X_{n,k})/\mathbb{E}(Y_{n,k})$ می پردازیم.

تعریف می‌کنیم $\mu_{n,k}^{[X]} := \mathbb{E}(X_{n,k})$ و $\mu_{n,k}^{[Y]} := \mathbb{E}(Y_{n,k})$. سپس از (۲) نتیجه گرفته می‌شود که $\mu_{n,k}^{[X]}$ در معادلهٔ بازگشتی و شرایط اولیه زیر صدق می‌کند:

$$\mu_{n,k}^{[X]} = \sum_{0\leq j\leq n}\binom{n}{j}p^j q^{n-j}\left(\mu_{j,k-1}^{[X]} + \mu_{n-j,k-1}^{[X]}\right), \qquad n\geq b+1,\ \ k\geq 1,$$

$$\mu_{n,k}^{[X]} = \begin{cases} 1, & \text{اگر } 1\leq n\leq b, k=0 \\ 0, & \text{اگر } (n=0,k\geq 0) \text{ یا } (n\geq b+1, k=0) \text{ یا } (1\leq n\leq b, k\geq 1) \end{cases}$$

معادلهٔ بالا برای $\mu_{n,k}^{[Y]}$ نیز برقرار است ولی با شرایط اولیه متفاوت

$$\mu_{n,k}^{[Y]} = \begin{cases} 1, & \text{اگر } n\geq b+1, k=0 \\ 0, & \text{اگر } 0\leq n\leq b, k\geq 0 \end{cases}$$

در مرحلهٔ پواسونی سازی، از روابط مذکور در ملاحظه می‌شود که تبدیل‌های پواسونی $\mu_{n,k}^{[X]}$ و $\mu_{n,k}^{[Y]}$،

$$M_k^{[X]}(z) := \sum_{n\geq 0}\mu_{n,k}^{[X]}\frac{z^n}{n!}e^{-z}, \qquad M_k^{[Y]}(z) := \sum_{n\geq 0}\mu_{n,k}^{[Y]}\frac{z^n}{n!}e^{-z},$$

در معادلات و شرایط اولیهٔ (۵) و (۶) صدق می‌کند:

$$M_k^{[X]}(z) = \textstyle\sum_{j=0}^{k-1}\binom{k-1}{j}M_1^{[X]}\left(p^j q^{k-1-j}z\right), \qquad k\geq 1, \tag{۵}$$

$$M_k^{[Y]}(z) = \textstyle\sum_{j=0}^{k}\binom{k}{j}M_0^{[Y]}\left(p^j q^{k-j}z\right) = \sum_{j=0}^{k}\binom{k}{j}\left(1 - M_0^{[\hat{Y}]}\left(p^j q^{k-j}z\right)\right), \quad k\geq 0, \tag{۶}$$

$$M_1^{[X]}(z) = \sum_{n=1}^{b}\frac{(pz)^n}{n!}e^{-pz} + \sum_{n=1}^{b}\frac{(qz)^n}{n!}e^{-qz} - \sum_{n=1}^{b}\frac{(2-p^n-q^n)z^n}{n!}e^{-z},$$

$$M_0^{[Y]}(z) = 1 - \sum_{n=0}^{b}\frac{z^n}{n!}e^{-z} = 1 - M_0^{[\hat{Y}]}(z), \qquad \hat{Y}_{n,k} := 2^k - Y_{n,k}.$$

مرحلهٔ بعد به گرفتن تبدیل ملین از (۵) و (۶) تخصیص دارد. تبدیل‌های ملین $M_k^{[X]}(z)$ و $M_k^{[\hat{Y}]}(z)$ را با

$$M_k^{*[X]}(s) := \int_0^\infty M_k^{[X]}(z)\, z^{s-1} dz, \qquad M_k^{*[\hat{Y}]}(s) := \int_0^\infty M_k^{[\hat{Y}]}(z)\, z^{s-1} dz,$$

تعریف می‌کنیم که به‌ترتیب به‌ازای $\Re(s) > -2$ و $\Re(s) > 0$ وجود دارند (توجه: $M_k^{*[Y]}(s)$ به‌ازای $-2 < \Re(s) < 0$ وجود دارد). سپس (۵) و (۶) بدین‌صورت بازنویسی می‌شوند:

$$M_k^{*[X]}(s) = (p^{-s} + q^{-s})^{k-1} M_1^{*[X]}(s), \qquad M_k^{*[\hat{Y}]}(s) = (p^{-s} + q^{-s})^{k} M_0^{*[\hat{Y}]}(s),$$

با شرایط اولیهٔ

$$M_1^{*[X]}(s) = \sum_{n=1}^{b} \frac{\Gamma(n+s)}{\Gamma(n+1)} (p^{-s} + q^{-s} - 2 + p^n + q^n), \qquad M_0^{*[\hat{Y}]}(s) = \sum_{n=0}^{b} \frac{\Gamma(n+s)}{\Gamma(n+1)}.$$

بنابراین با استفاده از معکوس تبدیل ملین داریم:

$$M_k^{[X]}(z) = \frac{1}{2\pi i} \int_{\rho - i\infty}^{\rho + i\infty} g_b^{[X]}(s) (p^{-s} + q^{-s})^k z^{-s} ds, \qquad \rho > -2, \tag{۷}$$

$$M_k^{[Y]}(z) = 2^k - \frac{1}{2\pi i} \int_{\rho - i\infty}^{\rho + i\infty} \hat{g}_b^{[Y]}(s) (p^{-s} + q^{-s})^k z^{-s} ds, \qquad \rho > 0, \tag{۸}$$

که در آن $\rho := \Re(s)$ و

$$g_b^{[X]}(s) = \sum_{n=1}^{b} \frac{\Gamma(n+s)}{\Gamma(n+1)} \left(1 - \frac{2 - p^n - q^n}{p^{-s} + q^{-s}}\right), \qquad \hat{g}_b^{[Y]}(s) = \sum_{n=0}^{b} \frac{\Gamma(n+s)}{\Gamma(n+1)} := -g_b^{[Y]}(s).$$

اساساً به‌دنبال رفتار مجانبی $M_k^{[X]}(z)$ و $M_k^{[Y]}(z)$ به‌ازای $z = n$ هستیم؛ چون طبق تحلیلِ برگشت از پواسونی‌سازی انتظار داریم که $\mathbb{E}(X_{n,k}) \sim M_k^{[X]}(n)$ و $\mathbb{E}(Y_{n,k}) \sim M_k^{[Y]}(n)$.

در این‌جا به محاسبهٔ انتگرال‌های (۷) و (۸) به‌وسیلهٔ روش نقطه زینی می‌پردازیم ([۱۱] را ببینید). پس باید $\rho = \rho_{n,k}$ را به‌عنوان نقطه زینی تابعِ

$$T(s)^k n^{-s} = e^{k \log T(s) - s \log n},$$

انتخاب کنیم که جواب معادله $\frac{\partial}{\partial s}(k \log T(s) - s \log n) = 0$ است. به‌طور معادل ρ از معادلهٔ (۹) به‌دست می‌آید.

$$\frac{k}{\log n} = \frac{p^{-\rho} + q^{-\rho}}{p^{-\rho} \log(1/p) + q^{-\rho} \log(1/q)}, \tag{۹}$$

یعنی $\rho = \rho_{n,k} = \rho\left(\frac{k}{\log n}\right)$ تنها نقطه زینی حقیقی مقدارِ تابع مذکور است ([۱۰] را ببینید). تابع $T(s)^k n^{-s}$ هم‌چنین دارای بی‌نهایت نقطهٔ زینی مختلط مقدار $s_j = \rho + 2\pi i j / (\log p/q)$ ($j = \pm 1, \pm 2, \dots$) است. زیرا با برقراری واقعیت زیر

$$T(\rho + it) = p^{-\rho - it} \left(1 + \left(\frac{q}{p}\right)^{-\rho - it}\right) = p^{-\rho} \cdot e^{-it \log p} \left(1 + \left(\frac{q}{p}\right)^{-\rho} \cdot e^{it \log \frac{p}{q}}\right),$$

و قراردادن $t = 2\pi j / (\log p/q)$ در آن داریم:

$$T\left(\rho + 2\pi i j / (\log p/q)\right) = p^{-\rho} \cdot e^{-2\pi i j (\log p)/(\log p/q)} \left(1 + \left(\frac{q}{p}\right)^{-\rho} \cdot e^{2\pi i j}\right) = e^{-2\pi i j (\log p)/(\log p/q)} T(\rho).$$

در نتیجه رفتار تابع $T(s)^k n^{-s}$ در اطراف نقاط $s = s_j$ و نقطهٔ $s = \rho$ یک‌سان است. این باعث می‌شود که یک تابع متناوب در بسط‌های مجانبی $M_k^{[X]}(z)$ و $M_k^{[Y]}(z)$ ظاهر شود و هم‌چنین در بسط‌های مجانبی $\mu_{n,k}^{[X]} = \mathbb{E}(X_{n,k})$ و $\mu_{n,k}^{[Y]} = \mathbb{E}(Y_{n,k})$.

[DOR: 20.1001.1.25882546.1399.6.2.4.5]
[DOI: 10.52547/mmr.6.2.169]

قضیهٔ ۱. به‌ازای $\varepsilon > 0$، اگر $\beta(\rho) = (p^{-\rho}q^{-\rho}\ \log(p/q)^2)/(p^{-\rho}+q^{-\rho})^2$، $\rho_{n,k} = \rho(k/\log n)$، $j \in \mathbb{Z}$ و $t_j = 2\pi j/(\log p/q)$ باشد، آن‌گاه داریم:

۱. هرگاه $\alpha_1 + \varepsilon \le \frac{k}{\log n} \le \alpha_0 - \varepsilon$ یا به طور معادل $\rho := \rho_{n,k} > 0$، دراین صورت

$$\mu_{n,k}^{[Y]} = 2^k - \hat{G}_b^{[Y]}(\rho_{n,k}, \log_{p/q} p^k n)\frac{(p^{-\rho_{n,k}}+q^{-\rho_{n,k}})^k n^{-\rho_{n,k}}}{\sqrt{2\pi\beta(\rho_{n,k})k}}\left(1+\mathcal{O}\left(k^{-1/2}\right)\right).$$

که در آن $\hat{G}_b^{[Y]}(\rho,x) = \sum_{j\in\mathbb{Z}} \hat{g}_b^{[Y]}(\rho+it_j)e^{-2\pi ijx}$ یک تابع متناوب و غیرصفر است.

۲. هرگاه $\alpha_0 + \varepsilon \le \frac{k}{\log n} \le \alpha_2 - \varepsilon$ یا به طور معادل $-2 < \rho := \rho_{n,k} < 0$، دراین صورت

$$\mu_{n,k}^{[Y]} = G_b^{[Y]}(\rho_{n,k}, \log_{p/q} p^k n)\frac{(p^{-\rho_{n,k}}+q^{-\rho_{n,k}})^k n^{-\rho_{n,k}}}{\sqrt{2\pi\beta(\rho_{n,k})k}}\left(1+\mathcal{O}\left(k^{-1/2}\right)\right),$$

که در آن $G_b^{[Y]}(\rho,x) = \sum_{j\in\mathbb{Z}} g_b^{[Y]}(\rho+it_j)e^{-2\pi ijx}$ یک تابع متناوب و غیر صفر است.

۳. هرگاه $\alpha_1 + \varepsilon \le \frac{k}{\log n} \le \alpha_2 - \varepsilon$ یا به‌طور معادل $\rho := \rho_{n,k} > -2$، دراین صورت

$$\mu_{n,k}^{[X]} = G_b^{[X]}(\rho_{n,k}, \log_{p/q} p^k n)\frac{(p^{-\rho_{n,k}}+q^{-\rho_{n,k}})^k n^{-\rho_{n,k}}}{\sqrt{2\pi\beta(\rho_{n,k})k}}\left(1+\mathcal{O}\left(k^{-1/2}\right)\right).$$

که در آن $G_b^{[X]}(\rho,x) = \sum_{j\in\mathbb{Z}} g_b^{[X]}(\rho+it_j)e^{-2\pi ijx}$ یک تابع متناوب و غیرصفر است.

برهان. با استفاده از روش نقطه زینی و همانند برهان قضیهٔ ۲ در [۱۰] ولی با توابع جدید $g_b^{[Y]}(s)$، $\hat{g}_b^{[Y]}(s)$ و $g_b^{[X]}(s)$، انتگرال‌های (۷) و (۸) را به‌طور مجانبی محاسبه می‌کنیم. سپس تحلیلِ برگشت از پواسونی‌سازی، نتایج را به‌دست می‌دهد. حال به اثبات غیرصفر بودن توابع $G_b^{[X]}(\rho,x)$ و $G_b^{[Y]}(\rho,x)$ ($G_b^{[Y]} = -\hat{G}_b^{[Y]}$) می‌پردازیم. یعنی باید نشان دهیم که به‌ازای $x\in[0,1]$، $G_b^{[X]}(\rho,x) > 0$ و $\hat{G}_b^{[Y]}(\rho,x) > 0$.

ابتدا حالتی را در نظرمی‌گیریم که $\rho \to \infty$: در این حالت هر دو تابع مذکور غیرصفر است. زیرا داریم:

$$G_b^{[X]}(\rho,x) \sim \sum_{j\in\mathbb{Z}}\left(\sum_{n=1}^{b}\frac{\Gamma(n+\rho+it_j)}{\Gamma(n+1)}\right)e^{-2\pi ijx},$$

$$\hat{G}_b^{[Y]}(\rho,x) \sim \sum_{j\in\mathbb{Z}}\left(\sum_{n=0}^{b}\frac{\Gamma(n+\rho+it_j)}{\Gamma(n+1)}\right)e^{-2\pi ijx},$$

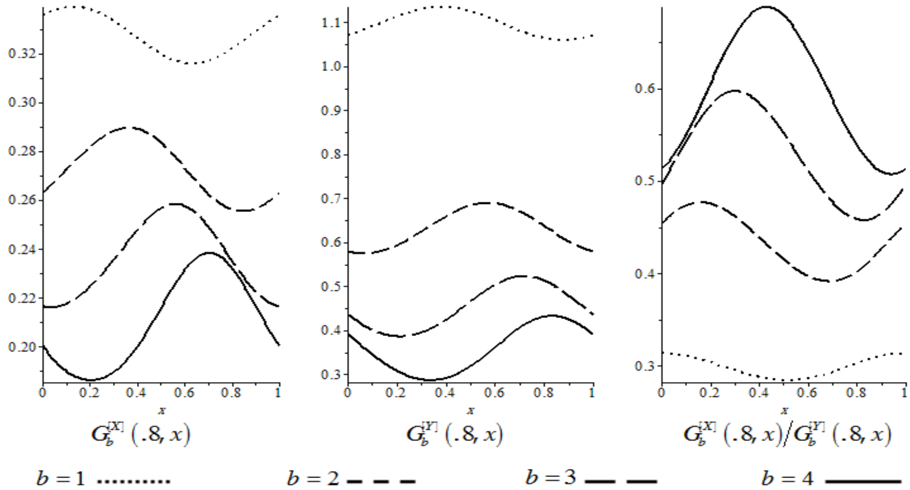
و هر دو تابع کراندار نیستند. چون با افزایش ρ و شرط $t = o(\rho)$،

$$\left|\frac{1}{\Gamma(\rho+b)}\sum_{n=1}^{b}\frac{\Gamma(n+\rho+it_j)}{\Gamma(n+1)}\right| = \mathcal{O}\left(e^{-{t_j}^2/\rho}\right) = \left|\frac{1}{\Gamma(\rho+b)}\sum_{n=0}^{b}\frac{\Gamma(n+\rho+it_j)}{\Gamma(n+1)}\right|.$$

از این برآورد نتیجه می‌گیریم که

$$G_b^{[X]}(\rho,x) = \mathcal{O}\left(\sum_{j\in\mathbb{Z}}\left|\sum_{n=1}^{b}\frac{\Gamma(n+\rho+it_j)}{\Gamma(n+1)}\right|\right) = \mathcal{O}\left(\rho^{1/2}\Gamma(\rho+b)\right),$$

و $\hat{G}_b^{[Y]}(\rho,x) = \mathcal{O}\left(\rho^{1/2}\Gamma(\rho+b)\right)$. این برآوردها حاکی از این است که اگر حتی توابع $G_b^{[X]}(\rho,x)$ و $\hat{G}_b^{[Y]}(\rho,x)$ را به $\Gamma(\rho+b)$ تقسیم کنیم، مرتبهٔ آنها $\mathcal{O}\left(\rho^{1/2}\right)$ است و با افزایش ρ به بی نهایت میل می‌کند.



شکل ۲. نمایش منحنی‌های متناوب توابع $G_b^{[X]}(.8,x)$، $G_b^{[Y]}(.8,x)$ در یک دورهٔ تناوب، یعنی $x\in[0,1]$، و نسبت‌شان برای ۴ مقدار مختلف b و $p=.9$

یعنی این توابع کراندار نیستند. پس $M>0$ و $-2\ge\rho_0$ ای وجود دارند به‌طوری که به‌ازای $\rho\ge\rho_0$ و $x\in[0,1]$، $G_b^{[X]}(\rho,x)>M>0$ و $\hat{G}_b^{[Y]}(\rho,x)>M>0$.

حال فرض کنید $\rho\le\rho_0$. در این حالت کافی است که نشان دهیم مشتق هر دو تابع $G_b^{[X]}(\rho,x)$ و $\hat{G}_b^{[Y]}(\rho,x)$ نسبت به x کراندار است. چون t^* ای وجود دارد به‌طوری که به‌ازای $|t|>t^*$، داریم $|\Gamma(y+it)|<\varepsilon$ و y ثابت است؛ آن‌گاه به‌ازای $j_0:=\max\{|j|\in\mathbb{Z}:|t_j|\le t^*\}$، $\sum_{j\in\mathbb{Z}:|t_j|\ge t^*} j=0$، $\rho\le\rho_0$ و $x\in[0,1]$،

$$\left|\frac{d}{dx}\left(G_b^{[X]}(\rho,x)\right)\right|\le\sum_{j\in\mathbb{Z}}\sum_{n=1}^{b}\left|\frac{\Gamma(n+\rho+it_j)}{\Gamma(n+1)}\right|\left|1-\frac{2-p^n-q^n}{p^{-\rho-it_j}+q^{-\rho-it_j}}\right|\cdot 2\pi|j|,$$

$$\le 4\pi j_0 b|\Gamma(b+\rho_0)||1+q^{\rho_0+1}|+2\pi b\varepsilon|1+q^{\rho_0+1}|\left|\sum_{j\in\mathbb{Z}:|t_j|\ge t^*} j\right|$$

$$=4\pi b j_0|\Gamma(b+\rho_0)||1+q^{\rho_0+1}|:=M^*<\infty.$$

به‌همین ترتیب ثابت می‌شود که مشتق تابع $\hat{G}_b^{[Y]}(\rho,x)$ نیز نسبت به x کراندار است.

برای ادامهٔ برهان دنبالهٔ $x_i\in[0,1]$ها را طوری انتخاب می‌نماییم که $G_b^{[X]}(\rho_0,x_i)>G_b^{[X]}(\rho,x_i)>\eta>0$ و $|x_{i+1}-x_i|\le\eta/2M^*$. در نتیجه طبق قضیهٔ مقدار میانگین داریم

$$G_b^{[X]}(\rho,x)\ge\eta-(\eta/2M^*)\cdot M^*=\eta>0,$$

و به همین ترتیب ثابت می‌شود $\hat{G}_b^{[Y]}(\rho,x)>0$.

قضیهٔ ۲. فرض کنید $\alpha_{n,k}:=\frac{k}{\log n}$. اگر $p\to\frac{1}{2}^+$ آن‌گاه

$$\rho^*(\alpha_{n,k}):=\lim_{p\to\frac{1}{2}^+}\rho(\alpha_{n,k})=\frac{\alpha_{n,k}}{1-\alpha_{n,k}\log 2},$$

$$\frac{\mu_{n,k}^{[X]}}{\mu_{n,k}^{[Y]}}\to\frac{\sum_{j=1}^{b}\frac{\Gamma\left(j+\rho^*(\alpha_{n,k})\right)}{\Gamma(j+1)}\left(2^{-\rho^*(\alpha_{n,k})}(1-2^{-j})-1\right)}{\sum_{j=0}^{b}\frac{\Gamma\left(j+\rho^*(\alpha_{n,k})\right)}{\Gamma(j+1)}},$$

$$\alpha_0-\varepsilon<\alpha_{n,k}<\alpha_2+\varepsilon.$$

برهان. طبق (۴) به سادگی ملاحظه می‌شود که $\rho^*(\alpha) = \frac{\alpha}{1-\alpha\log 2}$. در نتیجه هرگاه $p \to \frac{1}{2}^+$، از قضیهٔ ۱ داریم

$$\frac{\mu_{n,k}^{[X]}}{\mu_{n,k}^{[Y]}} \to \frac{G_b^{[X]}(\rho_{n,k}, \log_{p/q} p^k n)}{G_b^{[Y]}(\rho_{n,k}, \log_{p/q} p^k n)} \to \left.\frac{g_b^{[X]}(s)}{g_b^{[Y]}(s)}\right|_{s=\rho^*(\alpha_{n,k})}.$$

با جای‌گزین کردن $g_b^{[X]}(s)$ و $g_b^{[Y]}(s)$ برهان کامل می‌شود.

واریانس‌های نمایه‌های سطلی و داخلی

در این بخش، تقریب‌های مجانبی $\sigma_{n,k}^{[X]^2} := \mathbb{V}(X_{n,k})$، $\sigma_{n,k}^{[Y]^2} := \mathbb{V}(Y_{n,k})$ و $\gamma_{n,k} := Cov(X_{n,k}, Y_{n,k})$ را به‌دست می‌آوریم.

فرض نمایید $N_k^{[X]}(z) := \sum_{n\geq 0} \mathbb{E}(X_{n,k}^2) \frac{z^n}{n!} e^{-z}$ و $N_k^{[\hat{Y}]}(z) := \sum_{n\geq 0} \mathbb{E}(\hat{Y}_{n,k}^2) \frac{z^n}{n!} e^{-z}$. حال تبدیل‌های پواسونی واریانس‌ها، $V_k^{[X]}(z) := N_k^{[X]}(z) - M_k^{[X]^2}(z)$ و $V_k^{[Y]}(z) := N_k^{[\hat{Y}]}(z) - M_k^{[\hat{Y}]^2}(z)$ را تعریف می‌کنیم (چون $\mathbb{V}(Y_{n,k}) = \mathbb{V}(\hat{Y}_{n,k})$). سپس از (۲)، (۳) و مشابه با (۵) و (۶) به‌دست می‌آید.

$$V_k^{[X]}(z) = \sum_{j=0}^{k-1} \binom{k-1}{j} V_1^{[X]}(p^j q^{k-1-j} z), \qquad V_k^{[Y]}(z) = \sum_{j=0}^{k} \binom{k}{j} V_0^{[Y]}(p^j q^{k-j} z), \tag{۱۰}$$

با شرایط اولیهٔ

$$V_1^{[X]}(z) = M_1^{[X]}(z) + 2 \sum_{\substack{1\leq n\leq b\\ 1\leq m\leq b}} p^n q^m \frac{z^{n+m}}{n!\,m!} e^{-z} - 2\sum_{n=2}^{b} \frac{z^n}{n!} e^{-z}(1-p^n-q^n) - M_1^{[X]^2}(z),$$

$$V_0^{[Y]}(z) = \sum_{n=0}^{b} \frac{z^n}{n!} e^{-z} - \left(\sum_{n=0}^{b} \frac{z^n}{n!} e^{-z}\right)^2.$$

بنابراین به‌ازای $\rho > -2$ داریم

$$V_k^{[X]}(z) = \frac{1}{2\pi i} \int_{\rho-i\infty}^{\rho+i\infty} g_{V,b}^{[X]}(s)(p^{-s}+q^{-s})^k z^{-s} ds, \tag{۱۱}$$

$$V_k^{[Y]}(z) = \frac{1}{2\pi i} \int_{\rho-i\infty}^{\rho+i\infty} g_{V,b}^{[Y]}(s)(p^{-s}+q^{-s})^k z^{-s} ds, \tag{۱۲}$$

که در آن

$$g_{V,b}^{[X]}(s) = \sum_{n=1}^{b} \frac{\Gamma(n+s)}{\Gamma(n+1)} - \sum_{n=1}^{b} \frac{\Gamma(n+s)}{\Gamma(n+1)} \frac{2-p^n-q^n}{p^{-s}+q^{-s}} - \sum_{n=1}^{b} \frac{\Gamma(2n+s)}{\Gamma(n+1)^2} 2^{-2n-s}$$
$$-2 \sum_{1\leq n<m\leq b} \frac{\Gamma(n+m+s)}{\Gamma(n+1)\Gamma(m+1)} 2^{-n-m-s} - \sum_{n=1}^{b} \frac{\Gamma(2n+s)}{\Gamma(n+1)^2} \frac{(2-p^n-q^n)^2}{p^{-s}+q^{-s}} 2^{-2n-s}$$
$$+2 \sum_{\substack{1\leq n\leq b\\ 1\leq m\leq b}} \frac{\Gamma(n+m+s)}{\Gamma(n+1)\Gamma(m+1)}$$
$$\cdot \frac{(2-p^n-q^n)(p^m(1+p)^{-m-n-s} + q^m(1+q)^{-m-n-s})}{p^{-s}+q^{-s}},$$

$$g_{V,b}^{[Y]}(s) = \sum_{n=0}^{b} \frac{\Gamma(n+s)}{\Gamma(n+1)} - \sum_{n=0}^{b} \frac{\Gamma(2n+s) 2^{-2n-s}}{\Gamma(n+1)^2} - 2 \sum_{0\leq n<m\leq b} \frac{\Gamma(n+m+s) 2^{-n-m-s}}{\Gamma(n+1)\Gamma(m+1)}.$$

قضیهٔ ۳. به‌ازای $\varepsilon > 0$، هرگاه $\alpha_1 + \varepsilon \leq \frac{k}{\log n} \leq \alpha_2 - \varepsilon$ در این صورت

$$
{\sigma_{n,k}^{[X]}}^2 = G_{V,b}^{[X]}\left(\rho_{n,k}, \log_{p/q} p^k n\right) \frac{\left(p^{-\rho_{n,k}} + q^{-\rho_{n,k}}\right)^k n^{-\rho_{n,k}}}{\sqrt{2\pi\beta\left(\rho_{n,k}\right)k}} \left(1 + \mathcal{O}\left(k^{-1/2}\right)\right),
$$

$$
{\sigma_{n,k}^{[Y]}}^2 = G_{V,b}^{[Y]}\left(\rho_{n,k}, \log_{p/q} p^k n\right) \frac{\left(p^{-\rho_{n,k}} + q^{-\rho_{n,k}}\right)^k n^{-\rho_{n,k}}}{\sqrt{2\pi\beta\left(\rho_{n,k}\right)k}} \left(1 + \mathcal{O}\left(k^{-1/2}\right)\right),
$$

که در آن $\rho_{n,k} = \rho(k/\log n) > -2$ در معادلهٔ (۹) صدق می‌کند و

$$
G_{V,b}^{[X]}(\rho, x) = \sum\nolimits_{j\in\mathbb{Z}} g_{V,b}^{[X]}\left(\rho + it_j\right) e^{-2\pi ijx},
$$

$$
G_{V,b}^{[Y]}(\rho, x) = \sum\nolimits_{j\in\mathbb{Z}} g_{V,b}^{[Y]}\left(\rho + it_j\right) e^{-2\pi ijx}.
$$

برهان. با استفاده از روش نقطه زینی و همانند برهان قضیهٔ ۲ در [۱۰] ولی با توابع جدید $g_{V,b}^{[X]}(s)$ و $g_{V,b}^{[Y]}(s)$، انتگرال‌های (۱۱) و (۱۲) را محاسبه می‌کنیم. سپس طبق تحلیلِ برگشت از پواسونی‌سازی، نتایج به‌دست می‌آیند. از لم۱ و برهان قضیهٔ۱، غیرصفر بودن توابع متناوب $G_{V,b}^{[X]}$، $G_{V,b}^{[Y]}$ و درنتیجه واریانس‌ها نیز ثابت می‌شود.

لم ۱. ${\sigma_{n,k}^{[X]}}^2 = \Theta\left(\mu_{n,k}^{[X]}\right)$ و ${\sigma_{n,k}^{[Y]}}^2 = \Theta\left(\mu_{n,k}^{[X]}\right)$.

برهان. طبق لم ۸ در [۱۰]، رابطهٔ (۱۰) و برآورد زیر

$$
V_1^{[X]}(z) = \begin{cases} |z|^2, & z \to 0 \text{ اگر} \\ |z| e^{-q\Re(z)}, & z \to \infty, |z| \le \varepsilon \text{ اگر} \end{cases} = V_0^{[Y]}(z),
$$

داریم $V_k^{[X]}(z) = \Theta\left(M_k^{[X]}(z)\right)$ و $V_k^{[Y]}(z) = \Theta\left(M_k^{[X]}(z)\right)$. هم‌چنین طبق تحلیلِ برگشت از پواسونی‌سازی، داریم $\mu_{n,k}^{[X]} \sim M_k^{[X]}(n)$، ${\sigma_{n,k}^{[X]}}^2 \sim V_k^{[X]}(n)$ و ${\sigma_{n,k}^{[Y]}}^2 \sim V_k^{[Y]}(n)$. پس به این صورت برهان کامل می‌شود.

توزیع‌های حدی نمایه‌های سطلی و داخلی

در این بخش نشان می‌دهیم که اگر $\mu_{n,k}^{[X]} \to \infty$، آن‌گاه توزیع‌های حدی $X_{n,k}$ و $Y_{n,k}$ نرمال هستند. فرض کنید $F_k(z,u) := \sum_{n\ge 0} P_{n,k}^{[X]}(u) z^n/n!$ تابع مولد نمایی $P_{n,k}^{[X]}(u)$ باشد. براساس (۲) داریم

$$
F_k(z,u) = F_{k-1}(pz,u) F_{k-1}(qz,u), \qquad k \ge 2,
$$

با شرط اولیه

$$
F_1(z,u) = e^z + \sum_{n=1}^{b} \frac{z^n}{n!}\left(p^n\left(e^{qz}-1\right) + q^n\left(e^{pz}-1\right)\right)(u-1)
$$

$$
+ \sum_{\substack{1\le n\le b\\ 1\le m\le b}} p^n q^m \frac{z^{n+m}}{n!\,m!}(u-1)^2 - \sum_{n=2}^{b} \frac{z^n}{n!}\left(1 - p^n - q^n\right)\left(u^2-1\right).
$$

با حل معادله بازگشتی بالا به‌دست می‌آوریم:

$$
F_k(z,u) = \prod\nolimits_{0\le j\le k-1} F_1\left(p^j q^{k-1-j} z, u\right)^{\binom{k-1}{j}}, \qquad k \ge 1. \tag{۱۳}
$$

برای برگشت از پواسونی‌سازی در برهان قضیهٔ ۴ نیاز به کران بالای مذکور در قضیهٔ فرعی ۱ داریم. هم‌چنین در برهان قضیهٔ فرعی ۱، از لم زیر استفاده می‌کنیم.

لم ۲. اگر $z = re^{i\theta}$، که در آن $r \ge 0$ و $|\theta| \le \pi$، آن‌گاه

[DOR: 20.1001.1.25882546.1399.6.2.4.5]
[DOI: 10.52547/mmr.6.2.169]

$$\left|e^z-1-\sum_{n=1}^{b}\frac{z^n}{n!}\right|\le\left(e^r-1-\sum_{n=1}^{b}\frac{r^n}{n!}\right)e^{-c_1^{[b]}r\theta^2},\qquad (۱۴)$$

که در آن $c_1^{[b]}:=2/\left((b+2)\pi^2\right)$. هم‌چنین اگر r_b ریشهٔ معادلهٔ $1+e^{r/(b+2)}=e^r-\sum_{n=1}^{b}\frac{r^n}{n!}$ باشد، آن‌گاه به‌ازای $r\ge r_b$ داریم

$$\left|e^z-\sum_{n=1}^{b}\frac{z^n}{n!}\right|\le\left(e^r-\sum_{n=1}^{b}\frac{r^n}{n!}\right)e^{-c_1^{[b]}r\theta^2/2},\qquad |\theta|\le\pi.\qquad (۱۵)$$

برهان . نامساوی (۱۴)، حالت خاصی از نامساوی پیتل است ([۱۰] را ببینید). برای اثبات نامساوی (۱۵) داریم

$$\begin{aligned}\left|e^z-\sum_{n=1}^{b}\frac{z^n}{n!}\right|&\le\left|e^z-1-\sum_{n=1}^{b}\frac{z^n}{n!}\right|+1\\&\le\left(e^r-1-\sum_{n=1}^{b}\frac{r^n}{n!}\right)e^{-c_1^{[b]}r\theta^2}+1\\&\le\left(e^r-\sum_{n=1}^{b}\frac{r^n}{n!}\right)e^{-c_1^{[b]}r\theta^2/2},\end{aligned}$$

که در آن نامساوی آخر معادل با

$$1-e^{-c_1^{[b]}r\theta^2}\le\left(e^r-\sum_{n=1}^{b}\frac{r^n}{n!}\right)e^{-c_1^{[b]}r\theta^2/2}\left(1-e^{-c_1^{[b]}r\theta^2/2}\right),$$

یا $1+e^{c_1^{[b]}r\theta^2/2}\le e^r-\sum_{n=1}^{b}\frac{r^n}{n!}$ است. چون $1+e^{c_1^{[b]}r\theta^2/2}\le 1+e^{r/(b+2)}\le e^r-\sum_{n=1}^{b}\frac{r^n}{n!}$ به‌ازای $r\ge r_b$ برقرار است، بنابراین نامساوی $1+e^{c_1^{[b]}r\theta^2/2}\le e^r-\sum_{n=1}^{b}\frac{r^n}{n!}$ حاصل می‌شود. از طرفی اگر r_b ریشهٔ معادلهٔ $1+e^{r/(b+2)}=e^r-\sum_{n=1}^{b}\frac{r^n}{n!}$ باشد، آن‌گاه نامساوی مذکور به‌ازای هر $b\ge 1$ برقرار می‌شود ، هرگاه $r\ge r_b$.

قضیهٔ فرعی ۱. به‌ازای $k\ge 1$، $r\ge 0$، $|\theta|\le\pi$، و $|u|=1$، عدد ثابت $c^{[b]}>0$ ای مستقل از k، r و θ وجود دارد به‌طوری‌که

$$\left|F_k\left(re^{i\theta},u\right)\right|\le e^{r-c^{[b]}r\theta^2}.\qquad (۱۶)$$

برهان . فرض کنید که r_b^* ریشهٔ معادلهٔ زیر است:

$$\left(2\sum_{n=1}^{b}\frac{(r/2)^n}{n!}+\sum_{\substack{1\le n\le b\\1\le m\le b}}\frac{(r/2)^{n+m}}{n!\,m!}\right)\left(e^{r/\left(2(b+2)\right)}+1\right)=e^r,$$

(به‌طور مثال $r_1^*=0.99$ و $r_2^*=4.4$ است). در این صورت ارائهٔ استدلال برای (۱۶) را به دو حالت تقسیم می‌کنیم: $r\le r_b^*$ و $r\ge r_b^*$. در حالت اول، از بسط

$$\begin{aligned}F_1(z,u)&=\sum_{j=0}^{b}\frac{z^j}{j!}+\sum_{\substack{1\le j\le b\\1\le n\le b}}p^jq^n\frac{z^{n+j}}{j!\,n!}(u-1)^2\\&+\sum_{j\ge b+1}\frac{z^j}{j!}\left(1-\sum_{n=1}^{b}\left(p^nq^{j-n}+p^{j-n}q^n\right)\binom{j}{n}(1-u)\right)\end{aligned}$$

$$+\sum_{j=2}^{b}\frac{z^j}{j!}\left(\sum_{n=1}^{b}\left(p^nq^{j-n}+p^{j-n}q^n\right)\binom{j}{n}(u-1)-(1-p^n-q^n)(u^2-1)\right)$$

استفاده می‌کنیم که از آن (۱۷) به‌دست می‌آید:

$$\left|F_1\left(re^{i\theta},e^{i\varphi}\right)\right|\le\left|\sum_{j=0}^{b}\frac{r^je^{ij\theta}}{j!}\right|+\sum_{j\ge b+1}\frac{r^j}{j!}\le e^{r-c_2^{[b]}r\theta^2}, \qquad (۱۷)$$

به‌ازای $0\le r\le r_b^*$ ،$|\theta|\le\pi$ و $c_2^{[b]}:=\left(2\sum_{j=0}^{b-1}\frac{z^j}{(j+1)!}\right)/\left(\pi^2e^{r_b^*}\sum_{j=0}^{b}\frac{z^j}{j!}\right)$.

حال فرض کنید $r\ge r_b^*$. عبارت به‌دست آمده برای $F_1(z,u)$ را می‌توان به‌صورت

$$F_1(z,u)=a_1(pz)a_1(qz)+\sum_{n=1}^{b}\frac{z^n}{n!}(p^n+q^n)+\sum_{n=1}^{b}\frac{z^n}{n!}\left(p^na_2(qz)+q^na_2(pz)\right)u$$
$$+\sum_{\substack{1\le n\le b\\1\le m\le b}}p^nq^m\frac{z^{n+m}}{n!\,m!}u^2-\sum_{n=2}^{b}\frac{z^n}{n!}(1-p^n-q^n)(u^2-1),$$

بازنویسی شود که در آن $a_1(z):=e^z-\sum_{n=1}^{b}\frac{z^n}{n!}$ و $a_2(z):=e^z-1-\sum_{n=1}^{b}\frac{z^n}{n!}$.

چون $r\ge r_b^*\ge r_b$ است، در این صورت با به‌کاربردن نابرابری‌های (۱۴) و (۱۵) داریم:

$$\left|F_1\left(re^{i\theta},e^{i\varphi}\right)\right|\le a_1(pr)a_1(qr)e^{-c_1^{[b]}r\theta^2/2}+\sum_{n=1}^{b}\frac{r^n}{n!}(p^n+q^n)+\sum_{\substack{1\le n\le b\\1\le m\le b}}p^nq^m\frac{r^{n+m}}{n!\,m!}$$
$$+\sum_{n=1}^{b}\frac{r^n}{n!}p^na_2(qr)e^{-c_1^{[b]}qr\theta^2}+\sum_{n=1}^{b}\frac{r^n}{n!}q^na_2(pr)e^{-c_1^{[b]}pr\theta^2}$$
$$\le\left(e^r-\sum_{n=1}^{b}\frac{r^n}{n!}(p^n+q^n)-\sum_{\substack{1\le n\le b\\1\le m\le b}}p^nq^m\frac{r^{n+m}}{n!\,m!}\right)e^{-c_1^{[b]}qr\theta^2}$$
$$+\sum_{n=1}^{b}\frac{r^n}{n!}(p^n+q^n)+\sum_{\substack{1\le n\le b\\1\le m\le b}}p^nq^m\frac{r^{n+m}}{n!m!}\le e^{r-c_1^{[b]}qr\theta^2/2}, \qquad (۱۸)$$

که نامساوی آخر از نامساوی زیر نتیجه گرفته می‌شود:

$$\left(\sum_{n=1}^{b}\frac{r^n}{n!}(p^n+q^n)+\sum_{\substack{1\le n\le b\\1\le m\le b}}p^nq^m\frac{r^{n+m}}{n!\,m!}\right)\left(e^{c_1^{[b]}qr\theta^2/2}+1\right)\le e^r.$$

طرف چپ نامساوی بالا کم‌تر از عبارت $\left(2\sum_{n=1}^{b}\frac{(r/2)^n}{n!}+\sum_{\substack{1\le n\le b\\1\le m\le b}}\frac{(r/2)^{n+m}}{n!m!}\right)\left(e^{r/(2(b+2))}+1\right)$ است و این عبارت هم به‌ازای $r\ge r_b^*$، از e^r کم‌تر است.

از نامساوی های (۱۷) و (۱۸) داریم:

$$\left|F_1\left(re^{i\theta},e^{i\varphi}\right)\right|\le e^{r-c^{[b]}r\theta^2},\qquad \left(c^{[b]}:=\min\left\{c_1^{[b]},c_2^{[b]}\right\}\right),$$

به‌ازای $r\ge 0$ و $|\theta|\le\pi$. بنابراین طبق (۱۳)، نامساوی (۱۶) حاصل می شود.

حال به تعریف تابع $Q(z,u)$ و برهان خوش‌تعریفی آن در لم ۳ که در برهان قضیهٔ ۴ استفاده می‌شود، می‌پردازیم.

$$Q_k(z,u):=\sum_{j=0}^{k-1}\binom{k-1}{j}Q\left(p^jq^{k-1-j}z,u\right)=\log e^{-z}F_k(z,u), \qquad (۱۹)$$
$$Q(z,u):=\log e^{-z}F_1(z,u)$$
$$=\log\left(1+a_3(z)(u-1)+a_4(z)(u-1)^2-a_5(z)(u^2-1)\right),$$

که در آن

$$a_3(z) := \sum_{n=1}^{b} \frac{z^n}{n!}(p^n e^{-pz} + q^n e^{-qz}) - \sum_{n=1}^{b} \frac{z^n}{n!} e^{-z}(p^n + q^n),$$

$$a_4(z) := \sum_{\substack{1\le n\le b\\ 1\le m\le b}} p^n q^m \frac{z^{n+m}}{n!\,m!} e^{-z}, \qquad a_5(z) := \sum_{n=2}^{b} \frac{z^n}{n!} e^{-z}(1 - p^n - q^n).$$

لم ۳. به‌ازای $r \ge 0$ ، $|\theta| \le \varepsilon$ و $|u| = 1$، تابع $Q\left(re^{\theta}, u\right)$ خوش‌تعریف است.

برهان . برای خوش‌تعریف بودن $Q\left(re^{\theta}, u\right)$ باید ابتدا نشان دهیم که به‌ازای $r \ge 0$ و $|u| = 1$،

$$\left|a_3(r)\left(e^{i\varphi} - 1\right) + a_4(r)\left(e^{i\varphi} - 1\right)^2 - a_5(r)\left(e^{2i\varphi} - 1\right)\right| < 1.$$

با محاسبهٔ قدرمطلق عدد مختلط مذکور در بالا داریم

$$\left|a_3(r)\left(e^{i\varphi} - 1\right) + a_4(r)\left(e^{i\varphi} - 1\right)^2 - a_5(r)\left(e^{2i\varphi} - 1\right)\right|^2$$
$$= a_3(r)^2 v - a_4(r)\left(a_3(r) - a_4(r)\right)v^2 - a_5(r)\left(1 - a_5(r)\right)v(2 - v/2),$$

که در آن $v := 2(1 - \cos\varphi)$. از طرفی چون

$$a_3(r) - a_4(r) \ge a_3(r) - 2a_4(r) = e^{-r}$$
$$\times \left(\sum_{n=1}^{b} \frac{(pr)^n}{n!}\left(e^{rq} - 1 - \sum_{n=1}^{b} \frac{(qr)^n}{n!} e^{-r}\right) + \sum_{n=1}^{b} \frac{(qr)^n}{n!}\left(e^{rp} - 1 - \sum_{n=1}^{b} \frac{(pr)^n}{n!} e^{-r}\right)\right) \ge 0,$$

آن‌گاه داریم

$$\left|a_3(r)\left(e^{i\varphi} - 1\right) + a_4(r)\left(e^{i\varphi} - 1\right)^2 - a_5(r)\left(e^{2i\varphi} - 1\right)\right| \le \sqrt{v}\,a_3(r) \le \sqrt{2}\,a_3(r).$$

با محاسبه‌ای ساده به‌دست می‌آوریم که

$$\sqrt{2}a_3(r) \le 2\sqrt{2}\max_{r\ge 0} \sum_{n=1}^{b} \frac{(r/2)^n}{n!} e^{-r/2}\left(1 - e^{-r/2}\right) \le \frac{\sqrt{2}}{2}\cdot\left(e^{\frac{1.35725}{2}} - 1\right) = 0.68675 < 1.$$

بنابراین اثبات لم برای $z = r$ کامل شده است و ادعای لم برای $z = re^{\theta}$ از تحلیلی بودن نتیجه می‌شود.

ملاحظهٔ ۱. از لم ۴ در [۱۰] نتیجه می‌گیریم که اگر

$$\rho_0 := \begin{cases} \rho, & \text{اگر } \rho \ge 1, k \ge \alpha_1 \log n \\ 1, & \text{اگر } \rho \le 1 \end{cases}$$

که در آن ρ در (۹) صدق می‌کند، آن‌گاه

$$\frac{d^l}{d\theta^l}\left(M_k^{[X]}\left(ne^{\theta}\right)\right) = \mathcal{O}\left({\rho_0}^l n^{-l} M_k^{[X]}(n)\right), \qquad \frac{d^l}{d\theta^l}\left(M_k^{[Y]}\left(ne^{\theta}\right)\right) = \mathcal{O}\left({\rho_0}^l n^{-l} M_k^{[Y]}(n)\right),$$
$$\frac{d^l}{d\theta^l}\left(V_k^{[X]}\left(ne^{\theta}\right)\right) = \mathcal{O}\left({\rho_0}^l n^{-l} V_k^{[X]}(n)\right), \qquad \frac{d^l}{d\theta^l}\left(V_k^{[Y]}\left(ne^{\theta}\right)\right) = \mathcal{O}\left({\rho_0}^l n^{-l} V_k^{[Y]}(n)\right).$$

قضیهٔ ۴. فرض نمایید $\mu_{n,k}^{[X]} \to \infty$. در این صورت

$$\frac{X_{n,k} - \mu_{n,k}^{[X]}}{\sigma_{n,k}^{[X]}} \xrightarrow{d} \mathcal{N}(0,1), \qquad \frac{Y_{n,k} - \mu_{n,k}^{[Y]}}{\sigma_{n,k}^{[Y]}} \xrightarrow{d} \mathcal{N}(0,1),$$

که در آن $\xrightarrow{d}$ و $\mathcal{N}(0,1)$ به‌ترتیب همگرایی در توزیع و متغیر تصادفی نرمال استاندارد را نشان می‌دهند.

برهان . با توجه به بسط $\log(1+x) = x - \frac{x^2}{2} + o(x^3)$، هرگاه $|x| < 1$، آن‌گاه

[DOR: 20.1001.1.25882546.1399.6.2.4.5]
[DOI: 10.52547/mmr.6.2.169]

$$Q(z,u)=\big(a_3(z)-2a_5(z)\big)(u-1)$$
$$+\frac{1}{2}\Big(2\big(a_4(z)-a_5(z)\big)-\big(a_3(z)-2a_5(z)\big)^2\Big)(u-1)^2+\bar{Q}(z,u)(u-1)^3,$$

که در آن $\bar{Q}(z,u)$ از فرمول باقی‌ماندهٔ تیلور به‌دست می‌آید و در این‌جا اهمیت کم‌تری دارد. هم‌چنین

$$\bar{Q}(z,u)=\begin{cases}|z|^2, & \text{اگر } z\to 0\\ |z|e^{-q\Re(z)}, & \text{اگر } z\to\infty, |z|\le\varepsilon\end{cases} \tag{۲۰}$$

به‌دست می‌آید که مورد نیاز است. از (۱۹) و بسط $Q(z,u)$ در بالا حاصل می‌شود

$$Q_k(z,u)=M_k^{[X]}(z)(u-1)+\frac{1}{2}\Big(V_k^{[X]}(z)-M_k^{[X]}(z)\Big)(u-1)^2+\bar{Q}_k(z,u)(u-1)^3,$$
$$\bar{Q}_k(z,u)=\sum_{j=0}^{k-1}\binom{k-1}{j}\bar{Q}\big(p^jq^{k-1-j}z,u\big).$$

از لم ۸ در [۱۰]، (۱۹) و (۲۰) نتیجه می‌شود که هرگاه $|\theta|\le\varepsilon$ و $|u-1|=o(1)$، داریم

$$\bar{Q}_k(z,u)(u-1)^3=\mathcal{O}\left(|u-1|^3\left|M_k^{[X]}\big(ne^{\theta}\big)\right|\right).$$

براساس لم ۱ ($\sigma_{n,k}^{[X]^2}=\Theta\left(\mu_{n,k}^{[X]}\right)\to\infty$) ملاحظهٔ ۱ و با استفاده از بسط‌های $M_k^{[X]}\big(ne^{\theta}\big)$ و $V_k^{[X]}\big(ne^{\theta}\big)$ در $\theta=0$ و $\varphi=o\left(\sigma_{n,k}^{[X]^{-2/3}}\right)$ حاصل می‌شود:

$$F_k\big(ne^{i\theta},e^{i\varphi}\big)=\exp\left(n-\frac{n}{2}\theta^2+M_k^{[X]}(n)i\varphi-nM_k^{[X]'}\big(ne^{\theta}\big)\varphi\theta-\frac{V_k^{[X]}(n)}{2}\varphi^2+\mathcal{O}(E)\right),$$
$$E:=n|\theta|^3+{\rho_0}^2\sigma_{n,k}^{[X]^2}|\varphi|\theta^2+\rho_0\sigma_{n,k}^{[X]^2}\varphi^2|\theta|+\sigma_{n,k}^{[X]^2}|\varphi|^3.$$

با استفاده از فرمول کوشی، تقریب استرلینگ، (۱۶) و بسط $F_k\big(ne^{i\theta},e^{i\varphi}\big)$ در بالا،

$$\mathbb{E}(e^{X_{n,k}})=\frac{n!}{2\pi i}\oint z^{-n-1}F_k\big(z,e^{i\varphi}\big)dz=\frac{n!\,n^{-n}}{2\pi}\int_{-\pi}^{\pi}e^{-in\theta}F_k\big(ne^{i\theta},e^{i\varphi}\big)d\theta$$
$$=\frac{n!\,n^{-n}}{2\pi}\int_{-n^{-\frac{2}{5}}}^{n^{-\frac{2}{5}}}e^{-in\theta}F_k\big(ne^{i\theta},e^{i\varphi}\big)d\theta+\mathcal{O}\left(n^{-1/10}e^{-c^{[b]}n^{1/5}}\right)$$
$$\sim\frac{e^{M_k^{[X]}(n)i\varphi-\sigma_{n,k}^{[X]^2}\varphi^2/2}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\left(\theta+\sqrt{n}M_k^{[X]'}(n)\varphi\right)^2/2}d\theta.$$

که در آن عبارت خطا بدین‌صورت محاسبه می‌شود:

$$\left|\frac{n!\,n^{-n}}{2\pi}\left(\int_{-\pi}^{-n^{-\frac{2}{5}}}e^{-in\theta}F_k\big(ne^{i\theta},e^{i\varphi}\big)d\theta+\int_{n^{-\frac{2}{5}}}^{\pi}e^{-in\theta}F_k\big(ne^{i\theta},e^{i\varphi}\big)d\theta\right)\right|$$
$$\le\frac{n!\,n^{-n}}{2\pi e^{-n}}F_k\big(n,e^{i\varphi}\big)\left(\int_{-\pi}^{-n^{-\frac{2}{5}}}e^{-c^{[b]}n\theta^2}d\theta+\int_{n^{-\frac{2}{5}}}^{\pi}e^{-c^{[b]}n\theta^2}d\theta\right)=\mathcal{O}\left(F_k\big(n,e^{i\varphi}\big)e^{-c^{[b]}n\theta^2}\right).$$

پس توزیع حدی $X_{n,k}$ و به طور مشابه توزیع حدی $Y_{n,k}$ به‌دست می‌آید.

بحث و نتیجه‌گیری

در این مقاله به تعمیم مهم‌ترین قسمت از نتایج به‌دست آمده در [۱۰] پرداختیم و هم‌چنین برای نکته‌ای اثبات

[DOR: 20.1001.1.25882546.1399.6.2.4.5]
[DOI: 10.52547/mmr.6.2.169]

نشده در آن نتایج، برهانی ارائهٔ کردیم. به‌طور دقیق‌تر، هدف این مقاله بررسی نمایهٔ داخلی (تعداد گره‌های غیربرگ یا داخلی در سطح k) و نمایهٔ سطلی (تعداد گره‌های برگ یا سطل در سطح k) در یک ساختمان دادهٔ مهم به نام ترای سطلی تصادفی (ترای تصادفی بنا شده از n داده و حداکثر ظرفیت $b \geq 1$ داده در هر برگ) است (حالت $b = 1$ در [۱۰] بررسی شده است). با روش‌هایی در تحلیل مختلط نشان داده‌ایم که به‌ازای $\varepsilon > 0$ و اعداد ثابت و مشخص α_1 و α_2، اگر k در بازهٔ $(\alpha_1 + \varepsilon)\log n \leq k \leq (\alpha_2 - \varepsilon)\log n$ باشد، آن‌گاه امیدریاضی و واریانس‌های هر دو نمایه شامل توابعی متناوب وغیرصفر هستند (غیرصفر بودن در [۱۰] ثابت نشده است). ترسیمی از توابع متناوب در فرمول‌های امیدریاضی نمایه‌ها برای $b = 1, 2, 3, 4$ و به‌کمک MAPLE ارائه کرده‌ایم. سپس نسبت دو نمایه را به‌ازای $b \geq 1$ بررسی کردیم. به‌طور مثال در حالت $b = 1$، متوسط تعداد گره‌های داخلی در یک سطح، 1.56 برابر متوسط تعداد گره‌های داخلی در همان سطح و در حالت $b = 4$ است. در حالی‌که وقتی $b = 1$، متوسط تعداد برگ‌ها در یک سطح، سه برابر متوسط تعداد برگ‌ها در همان سطح و در حالت $b = 4$ است. بالاخره با یافتن بسط مجانبی تابع مولد پواسن برای دنبالهٔ توابع مولد احتمال نمایه‌ها و فرمول انتگرال کوشی، بسط مجانبی توابع مولد احتمال برای نمایه‌ها به‌دست می‌آید که حاکی از نرمال بودن توزیع حدی نمایه‌ها است.

منابع

1. Drmota M., "Random trees: An interplay between combinatorics and probabilityو" Springer Wien New York, Vienna (2009).
2. Drmota M., Fuchs M., Hwang H.-K., Neininger R., "External profile of symmetric digital search trees", Proceedings of the 14th Workshop on Analytic Algorithmics and Combinatorics (ANALCO17) (2017) 124-130.
3. Flajolet P., Gourdon X., Dumas P., "Mellin transforms and asymptotics: harmonic sums", Theoret. Comput. Sci., 144 (1995) 3-58.
4. Fuchs M., Hwang H.-K., Neininger R., "Profiles of random trees: Limit theorems for random recursive trees and binary search trees", Algorithmica, 46:3-4 (2006) 367-407.
5. Fuchs M., Lee C.-K., "A general central limit theorem for shape parameters of m-ary tries and patricia tries", Electron. J. Combin., 21 (2014) 1-26.
6. Fuchs M., Hwang H.-K., Zacharovas V., "An analytic approach to the asymptotic variance of trie statistics and related structures", Theoret. Comput. Sci., 527 (2014) 1-36.
7. Knuth D., "The art of computer programming. Sorting and searching", Addison-Wesley, Reading, MA, New York (1998).
8. Magner A., Szpankowski W., "Profile of PATRICIA Tries", Algorithmica, 76 (2016) 1-67.
9. Nicodème P., "Average profiles, from tries to suffix-trees", In Discrete Math. Theor. Comput. Sci., Proc., (2005) 257-266.
10. Park G., Hwang H.-K., Nicodème P., Szpankowski W., "Profiles of tries", SIAM J. Computing, 38 (2009) 1821-1880.
11. Szpankowski W., "Average case analysis of algorithms on sequences", Wiley Inter science, New York (2001).

[DOR: 20.1001.1.25882546.1399.6.2.4.5]
[DOI: 10.52547/mmr.6.2.169]