سهم کولموگوروف در مبانی احتمال•

ولادیمیر ووک، گلن شیفر*

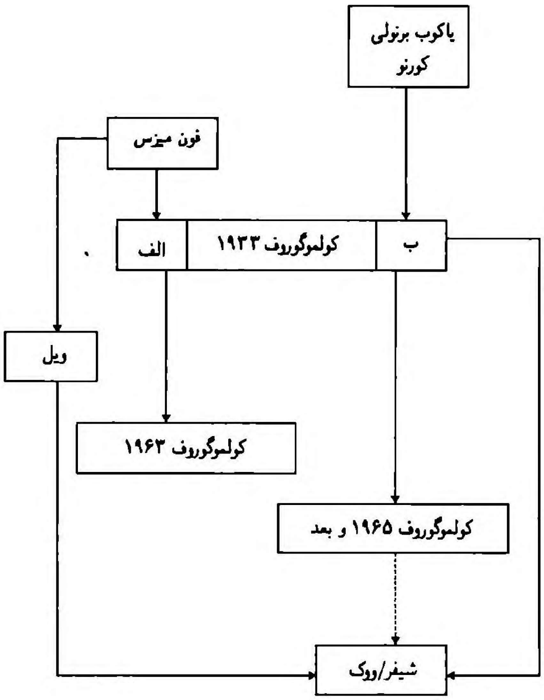
ترجمهٔ عبدالله حسنی جلیلیان

۱. مقدمه

این مقاله در قالب نمودار۱ ارائه می‌شود. در قالب این نمودار، اولین اقدام کولموگوروف برای صوری‌سازی مفهوم شهودی احتمال، در اثر مشهورش مبانی نظریهٔ احتمال، در ۱۹۳۳، قرار دارد [۱۰]. حاصل این صوری‌سازی، یعنی احتمال مبتنی بر نظریهٔ اندازه، به‌عنوان بنیان استاندارد نظریهٔ احتمال به‌کار رفته و هنوز هم به‌کار می‌رود؛ تقریباً در همهٔ کارهای ریاضی جاری دربارهٔ احتمال، رویکرد مبتنی بر نظریهٔ اندازه را به‌کار می‌گیرند. برای مرتبط کردن احتمال مبتنی بر نظریهٔ اندازه با واقعیت تجربی، کولموگوروف از دو اصل، که آنها را الف و ب نامید، استفاده کرد. اصل الف صورتی از شرط فون میزس[1] است حاکی از اینکه احتمالها باید فراوانیهای مشاهده شده باشند. اصل ب یک برداشت متناهیانه از «اصل کورنو» است که ریشه در فن حدس زدن (۱۷۱۳) اثر یاکوب برنولی [۴] دارد و در قرن هجدهم توسط آنتوان کورنو[2] در جامعهٔ علمی رواج یافت.

هدف تلاشهای بعدی کولموگوروف در صوری‌سازی احتمال، فراهم آوردن مبنای ریاضی بهتری برای کاربردها بود. در بخش ۳، نظریهٔ احتمال فراوانی‌گرای وی و در بخش ۴، نظریهٔ تصادفی بودنِ الگوریتمی‌اش را مورد بحث قرار می‌دهیم.

رشتهٔ دیگری از تحقیقات دربارهٔ احتمال، از ایده‌های فون میزس نشأت گرفته است. ژان ویل[3] در کتاب بررسی نقادانهٔ مفهوم گردایه[4] خود با استفاده از ایده‌های نظریهٔ بازیها، که از کارهای هموطن مشهورش بلز پاسکال سرچشمه می‌گیرد، اصلاحاتی در رویکرد فون میزس به عمل آورد. مفهوم مارتینگلِ منتج از آن، هرگز در بررسیهای کولموگوروف دربارهٔ مبانی احتمال مورد استفاده قرار نگرفت؛ او تعریف فون میزس را در جهانی کاملاً متفاوت با تحقیقات ویل گسترش داد. در بخش ۵ دربارهٔ این رشته از تحقیق، از جمله در مورد پیشنهاد اخیرمان [۲۹] برای مبتنی کردن مستقیم نظریهٔ ریاضی و تعبیر احتمال برمبنای مفهوم مارتینگل، بحث می‌کنیم.



نمودار۱ اقدامات مهم برای صورت‌بندی و تعبیر احتمال (به‌ترتیب تاریخی، از بالا به پایین)

مقاله را با بحث دربارهٔ سودمندی تصادفی بودنِ الگوریتمیِ کولموگوروف خاتمه می‌دهیم. استدلال ما این است که گرچه ممکن است سودمندی آن به‌عنوان چارچوبی برای بیان نتایج جدید دربارهٔ احتمال محدود باشد، اما به‌عنوان ابزاری برای کشف حقایق جدید، توانایی بالقوهٔ زیادی دارد. مثالی را از پژوهشهای خود به تفصیل شرح می‌دهیم.

هر دو جنبهٔ ریاضی و علمی کار کولموگوروف مورد بحث ما خواهد

---

1. Richard von Mises　2. Antoine Cournot　3. Jean Ville

4. *Étude critique de la notion de collectif*

بود. از آنجایی که جنبه‌های ریاضی غیرقابل مناقشه و عموماً شناخته شده هستند، توجه ما اغلب معطوف به جنبه‌های علمی است (نظریه‌های ریاضی تکوین‌یافته توسط کولموگوروف چگونه با واقعیت ارتباط پیدا می‌کند). اما بخش‌بندی بحث ما، مطابق نظریه‌های ریاضی احتمال است.

۲. احتمالِ مبتنی بر نظریهٔ اندازه

بنابر مبانی، نظریهٔ ریاضی احتمال، اندازه‌های احتمال یعنی اندازه‌هایی مانند P بر یک فضای اندازه‌پذیر $(\Omega, \mathcal{F})$ را که $P(\Omega) = ۱$، بررسی می‌کند. هر پیشامد صرفاً مجموعه‌ای مانند E، $E \in \mathcal{F}$، و احتمال آن $P(E)$ است. کولموگوروف نظریهٔ ریاضی پرباری را بر این پایهٔ ساده بنا کرد که پژوهشگران بسیاری آن را گسترش داده‌اند و امروز معروف است. در این بخش کوتاه توجه ما عمدتاً به روابط احتمالِ مبتنی بر نظریهٔ اندازه با واقعیت، از دید کولموگوروف، است.

کولموگوروف دو اصل برای تعبیر احتمال $P(E)$، که در آن E یک پیشامد است که در آزمایش C رخ می‌دهد یا رخ نمی‌دهد، مطرح کرد:

الف. می‌توان عملاً مطمئن بود که اگر آزمایش C به دفعات زیاد تکرار شود، فراوانی نسبی E تفاوت بسیار اندکی با $P(E)$ خواهد داشت.

ب. اگر $P(E)$ خیلی کوچک باشد، می‌توان عملاً مطمئن بود که وقتی آزمایش C تنها یک بار اجرا شود، پیشامد E اصلاً رخ نخواهد داد.

هرکدام از این دو اصل تاریخچه‌ای غنی دارند. در اصل الف، کولموگوروف ایده‌های فراوانی‌گرای ریچارد فون میزس (بویژه با اشاره به اثر او[۲۴]) را دنبال می‌کند. اصل ب به برنولی، کورنو، و لوی[1] برمی‌گردد (ر.ک. [۲۹]).

۳. احتمالِ فراوانی‌گرا

کولموگوروف طی سالهای ۱۹۳۸ـ۱۹۵۹ چند شرح غیررسمی از فلسفهٔ فراوانی‌گرای خود در باب احتمال را منتشر کرد [۱۱، ۱۲، ۱۳، ۱۴، ۲۰]. تنها کوشش وی برای صوری‌سازی این فلسفه در ۱۹۶۳، در مقالهٔ «در باب جدولهای اعداد تصادفی» [۱۵] به عمل آمده است. وی این مقاله را با بیان دلایل اینکه چرا پیش از این مبادرت به چنین کاری نکرده، شروع می‌کند:

۱. رویکرد فراوانی نامتناهیانه بر پایهٔ فراوانی حدی (وقتی که تعداد آزمایشها به بینهایت میل می‌کند)، نمی‌تواند اطلاعی در مورد کاربردهای واقعی، که در آنها همواره با تعداد متناهی آزمایش سروکار داریم، بدهد.

۲. رویکرد فراوانی‌گرا را در حالتی که تعداد آزمایشها زیاد ولی متناهی است نمی‌توان به‌صورت صرفاً ریاضی تدوین کرد.

کولموگوروف هرگز عقیده‌اش را در بارهٔ دلیل ۱ تغییر نداد. اما تفکراتش در باب پیچیدگی الگوریتمها تا سال ۱۹۶۳، وی را به تغییر عقیده‌اش در مورد دلیل ۲ سوق داد و متوجه شد که می‌تواند از این حقیقت که چند الگوریتم ساده برای تعریف صورت متناهیانه‌ای از گردایه‌های فون میزس وجود دارد، استفاده کند.

دنبالهٔ متناهی $(x_1, \ldots, x_N)$ متشکل از صفرها و یکها را در نظر بگیرید. این دنباله چه ویژگیهایی باید داشته باشد تا آن را تصادفی (یا به بیان شهودی، نتیجهٔ آزمایه‌های مستقل) تلقی کنیم؟ فون میزس تلاش کرده بود به این سؤال در مورد دنباله‌های نامتناهی پاسخ دهد؛ او گفته بود که دنبالهٔ نامتناهیْ تصادفی است (یا اینکه یک «گردایه») است، هرگاه در دو شرط زیر صدق کند

۱. فراوانی حدی یکها موجود باشد.

۲. اگر بدون آگاهی قبلی از برآمدها زیردنباله‌ی نامتناهی از این دنباله انتخاب کنیم، این فراوانی حدی تغییر نکند.

در حالت دنباله‌های متناهی، شرط اول بی‌معنی است و تنها باید نگران شرط دوم باشیم. البته به‌جای عبارت «تغییر نکند» باید «چندان تغییر نکند» را قرار داد، که در این‌صورت لازم است از دنباله‌های (N, ϵ)ـتصادفی به‌جای دنباله‌های تصادفی صحبت کنیم. در واقع، تعریف کولموگوروف در [۱۵] حتی پیچیده‌تر از این است، زیرا یک پارامتر اضافی $\mathcal{R}_N$ (مجموعهٔ قواعد انتخاب‌پذیرفتنی، که خود اشیای پیچیده‌ای هستند) وجود دارد که صریحاً در نمادگذاری وی منعکس نشده است. از نظر کولموگوروف، این عدم ظرافت ریاضی، احتمالاً با اهمیت فلسفی مفهوم فراوانی‌گرایی جبران شده است.

کولموگوروف در [۱۵] می‌گوید که به قواعد انتخاب ساده علاقه‌مند است. او سادگی را دقیقاً تعریف نمی‌کند، اما بر این فرض که قواعد انتخاب بسیار زیادی در $\mathcal{R}_N$ وجود ندارند تکیه می‌کند و در توجیه آن به این حقیقت که قواعد انتخاب سادهٔ زیادی نمی‌توانند موجود باشند، استناد می‌کند.

با اینکه قصد او از تعریف (N, ϵ)ـتصادفی بودن، به‌وضوح استفاده از آن به‌عنوان مبنایی برای یک صورت متناهیانه از نظریهٔ فون میزس بود، اما کولموگوروف کارِ ایجاد چنین نظریه‌ای را ادامه نداد. در واقع او این مفهوم را کلاً به نفع رویکرد ظریف و سرراست‌تری برای تعریف تصادفی بودن، که اکنون به آن می‌پردازیم، کنار گذاشت.

۴. تصادفی بودنِ الگوریتمی متناهیانه

همان‌طور که در بالا دیدیم، کولموگوروف در ۱۹۶۳ در بارهٔ استفاده از پیچیدگی الگوریتمی به‌عنوان نقطهٔ شروع تعریف فراوانی‌گرا از احتمال، به سبک فون میزس، فکر می‌کرد. ولی در آستانهٔ ۱۹۶۵، به تعریف مستقیمی از تصادفی بودن برحسب پیچیدگی الگوریتمی می‌اندیشید که بتوان با استفاده از آن، تصادفی بودن را به شکل مستقیم‌تر و بدون توسل به مفهوم فراوانی با کاربردها مربوط کرد. ما این رویکرد را نظریهٔ متناهیانهٔ تصادفی بودنِ الگوریتمی کولموگوروف می‌نامیم. وی آن را اولین بار به شکل چاپ شده در پاراگراف پایانی «سه رویکرد به تعریف مفهوم مقدار اطلاع» [۱۶] مطرح کرد. او بعدها این نظریه را در مقاله‌هایی که در ۱۹۶۸ [۱۷] و ۱۹۸۳ [۱۸، ۱۹] چاپ شد، گسترش داد. مفصل‌ترین توصیف او در این باره در «مبانی ترکیبیاتی نظریهٔ اطلاع و نظریهٔ احتمال» [۱۸] آمده است که در ۱۹۸۳ منتشر شد اما اولین بار در ۱۹۷۰ در ارتباط با سخنرانی‌اش در کنگرهٔ جهانی ریاضیات در نیس تهیه شده بود.

دنباله‌های برنولی. فرض کنید دنبالهٔ دودویی $(x_1, \ldots, x_N)$، k یک و $N-k$ صفر دارد. برای توصیف این دنباله، چون حداکثر $\binom{N}{k}$ تا از چنین دنباله‌هایی وجود دارد، $\log \binom{N}{k}$ بیت کافی است. به عقیدهٔ کولموگوروف، دنبالهٔ برنولی است اگر نتوان آن را با بیتهایی که تعدادشان در حد قابل ملاحظه‌ای کمتر باشد، توصیف کرد.

برای دقیق کردن این تعریف، به تعریف کوتاه‌ترین توصیف یک دنباله نیاز داریم. کشف مهمی که این امر را ممکن ساخت آن بود که یک روش عام

1. Lévy

توصیف وجود دارد که توصیفهایی تقریباً به کوتاهی توصیفهای فراهم آمده با هر روش جایگزینی را فراهم می‌آورد. (ری سولومونوف، پیشتر و مستقلاً [۳۱، ۳۲] به این مطلب پی برده بود. وجود یک روش عام توصیف، پیامدی از وجود یک الگوریتم عام است.)

پیچیدگی کولموگوروفِ x، $K(x)$، به‌عنوان طول کوتاه‌ترین توصیف x، در صورت استفاده از روش عام توصیف، تعریف می‌شود. با مجاز دانستن اینکه در روش عام توصیف از اطلاعات اضافی y استفاده شود، تعریف پیچیدگی شرطی کولموگوروف، $K(x|y)$، را به‌دست می‌آوریم. حال می‌توانیم بگوییم که دنبالهٔ $x = (x_1, \dots, x_N)$ با k تا یک، برنولی است اگر $K(x|N,k)$ به $\log\binom{N}{k}$ نزدیک باشد. کولموگوروف با به‌دست دادن یک عدد حقیقی m که این نزدیکی را اندازه می‌گیرد، این مفهوم را دقیق‌تر کرد: دنبالهٔ mـبرنولی است اگر در

$$K(x|N,k) \geq \log\binom{N}{k} - m$$

صدق کند.

فرض مربوط به تعبیر احتمال. دنبالهٔ برنولی تنها یک مثال از شئ تصادفی است. مفهوم کلی، که قبلاً در قالب تذکری کوتاه در مقالهٔ ۱۹۶۵ رخ نموده بود، مفهوم یک شئ تصادفی در مجموعه‌ای بزرگ و متناهی مانند A است که خود آن باید با توصیفی متناهی مشخص شود. با مفروض بودن توصیف A، پیچیدگی یک عنصر حداکثر برابر با لگاریتم تعداد عناصر A است. عنصر تصادفی x عنصری است که پیچیدگی‌اش نزدیک به این ماکسیمم باشد، یعنی آنکه در

$$K(x|A) \approx \log|A| \qquad (۱)$$

صدق کند. به‌طور رسمی، می‌توانیم تفاضل $\log|A| - K(x|A)$ را کاستی تصادفی بودن x در A بنامیم.

به خاطر آورید که احتمالِ مبتنی بر نظریهٔ اندازهٔ کولموگوروف با دو فرض برای تعبیر احتمال، اصلهای الف و ب، با جهان تجربی مرتبط شده بود. فرض تعبیری اصلی در نظریهٔ متناهیانهٔ تصادفی بودن الگوریتمی وی این است که انتظار داریم برآمد تحقق یافتهٔ x از آزمایشی که برآمدهایی در مجموعهٔ متناهی A دارد، به مفهوم (۱)، در A تصادفی باشد.

آمار با مدلهای پیچیدگی. اساس آمار متداول، مفهوم مدل آماری است؛ یعنی، خانواده‌ای از توزیعهای احتمال برای برآمدهای یک آزمایش. کولموگوروف پیشنهاد کرد به‌جای مدلهای آماری چیزی که ما در این مقاله مدلهای پیچیدگی می‌نامیم، قرار گیرد: رده‌هایی از مجموعه‌های مجزا که اجتماع آنها شامل تمام برآمدهای ممکن آزمایش باشد. (آزمایش ممکن است بسیار پیچیده باشد. ولی نوعاً متشکل از دنباله‌ای از آزمایه‌های یک آزمایش مقدماتی‌تر است.)

ما فرض تعبیری کولموگوروف را در مورد یک مدل پیچیدگی با این فرض به‌کار می‌بریم که برآمد آزمایش نسبت به مجموعه‌ای در مدل که شامل آن است، تصادفی خواهد بود. (چون مجموعه‌های مطرح در مدل مجزا هستند، دقیقاً یک مجموعه با این ویژگی موجود است.) به بیان دیگر، پذیرفتن یک مدل پیچیدگی به‌عنوان باور به این موضوع تعبیر می‌شود که کاستی تصادفی بودنِ برآمد واقعی در مجموعهٔ مدل که برآمد به آن تعلق دارد، کوچک خواهد بود.

کولموگوروف چندین مثال از مدلهای پیچیدگی ارائه کرده است، که ما در اینجا آنها را بازپردازی کرده‌ایم. فرض می‌شود که همهٔ دنباله‌ها در توصیفهای ما متناهی‌اند.

مثال ۱. دنبالهٔ دودویی x برنولی است هرگاه x به شرط

$$\begin{cases} N = \text{طول } x \\ k_0 = \text{تعداد ۰ها در } x \\ k_1 = \text{تعداد ۱ها در } x \end{cases}$$

N، k_0، k_1 تصادفی باشد.

همان‌طورکه پیشتر شرح دادیم، این بدین‌معنی است که پیچیدگی $K(x|N,k_1)$ به $\log\binom{N}{k_1}$ نزدیک است. مدل پیچیدگی برنولی به‌طور صوری متشکل از تمام رده‌های هم‌ارزی از دنباله‌های متناهی صفر و یکهاست که در آن، هم‌ارزی دو دنباله بدین‌معنی است که طول و تعداد یکها (و در نتیجه تعداد صفرها)ی آنها یکسان است.

مثال ۲. دنبالهٔ دودویی x مارکوف است هرگاه x به شرط

$$\begin{cases} N = \text{طول } x \\ s = \text{اولین عنصر } x \\ k_{00} = \text{تعداد ۰۰ها در } x \\ k_{01} = \text{تعداد ۰۱ها در } x \\ k_{10} = \text{تعداد ۱۰ها در } x \\ k_{11} = \text{تعداد ۱۱ها در } x \end{cases}$$

N، s، k_{00}، k_{01}، k_{10}، k_{11} تصادفی باشد.

در قیاس با مثال قبل، مدل پیچیدگی مارکوف متشکل از تمام رده‌های هم‌ارزی است که در آن، هم‌ارزی دو دنباله بدین‌معنی است که دارای طول برابر، عنصر اول یکسان و تعداد انتقالهای $i \to j$ یکسان به‌ازای هر $i, j \in \{0, 1\}$ هستند.

مثال ۳. دنباله‌های مارکوف x از مرتبهٔ d مشابه مثال قبل تعریف می‌شود.

مثال ۴. دنبالهٔ $x = (x_1, \dots, x_N)$ از اعداد حقیقی گاوسی است هرگاه x به شرط

$$\begin{cases} N = x \quad \text{طول} \\ m = \text{میانگین نمونه} = \dfrac{1}{N}\sum_{n=1}^{N} x_n \\ \sigma^2 = \text{واریانس نمونه} = \dfrac{1}{N}\sum_{n=1}^{N}(x_n - m)^2 \end{cases}$$

N، m، σ^2 تصادفی باشد.

البته برای دقیق کردن این مورد، باید خط حقیقی را به‌گونه‌ای گسسته ساخت.

مثال ۵. به طریقی مشابه می‌توان دنباله‌های پواسون را تعریف کرد.

برخی نتایج ریاضی در مورد مدلهای پیچیدگی در مقاله‌های یوگنی آسارین[۱، ۲] و رسالهٔ دکتری وی[۳] (به راهنمایی کولموگوروف) اثبات شده‌اند. یک نتیجهٔ نوعی به شکل زیر است: اگر x گاوسی (مثال ۴ را ببینید)، N بزرگ و $[a, b]$ بازه‌ای ثابت باشد، آنگاه

$$\frac{\#\{x_n \in [a,b]\}}{N} \approx \frac{1}{\sqrt{2\pi}\sigma}\int_a^b e^{-\frac{(t-m)^2}{2\sigma^2}}\, dt\,.$$

در ایدهٔ مدل پیچیدگی، احتمال به‌عنوان مفهوم اساسی برای مبانی آمار در نظر گرفته نمی‌شود. می‌توانیم با احتمال نامیدن برخی فراوانیهای نسبی، احتمال را دوباره معرفی کنیم. مثلاً، در مثال ۲، می‌توانیم احتمال شرطی این را که در دنبالهٔ مارکوف x، یک ۱ پس از یک ۰ بیاید با نسبت تعداد دفعات وقوع ۱۰ در x به تعداد کل دفعات وقوع ۰۰ و ۱۰ در x، تعریف کنیم. با توجه به انگیزهٔ اولیهٔ کولموگوروف، این گام بسیار طبیعی است، اما می‌توانیم مدلهای پیچیدگی را مستقیماً، بدون هیچ بحثی از احتمال، برای پیشگویی به‌کار بریم. مثلاً، فرض کنید مایل به اتخاذ این فرض هستیم که دنباله‌ای از اعداد حقیقی به طول N، که تنها نیمی از آن تاکنون معلوم است، گاوسی است. در این‌صورت می‌توانیم پیشگویی کنیم که میانگین اعداد واقع در نیمهٔ دوم به میانگین نیمهٔ اول نزدیک خواهد بود. روشن است که پیشگوییهای متعدد دیگری از این نوع را می‌توان برای مدل گاوسی و مدلهای دیگری که در نظر گرفته‌ایم انجام داد. در این کاربردها به هیچ احتمالی نیاز نیست.

دنباله‌های استوکاستیک. اگر دنبالهٔ $(x_1, \ldots, x_N)$ برنولی باشد، آنگاه در یک مجموعهٔ ساده (مجموعهٔ همهٔ دنباله‌های با طول برابر و تعداد یکهای برابر را می‌توان با استفاده از تقریباً ۲ $\log N$ بیت اطلاع توصیف کرد) تصادفی است. این مطلب برای سایر مدلهایی هم که ذکر کردیم صادق است و به نظر می‌رسد که منبع توان پیشگویی این مدلها نیز همین باشد. در یک سمینار در سال ۱۹۸۲ در دانشگاه مسکو، کولموگوروف یک شئ متناهی را (α, β)-استوکاستیک (که α و β نوعاً اعداد صحیح مثبت کوچکی هستند) تعریف کرد هرگاه یک مجموعهٔ متناهی A وجود داشته باشد به‌طوری که

$$x \in A, \quad K(A) = \leq \alpha,. \quad K(x|A) \geq \log|A| - \beta.$$

شِن[۳۰] و ویوگین[۳۷] نخستین کسانی بودند که مفهوم استوکاستیک‌بودن[1] کولموگوروف را بررسی کردند؛ آنها به این پرسش که چه تعداد دنبالهٔ استوکاستیک (به معانی مختلف) وجود دارد پاسخ گفتند. پرسشهای ریاضی جالب دیگری، هم پاسخ داده شده و هم بی‌پاسخ مانده، برخاسته از مفهوم استوکاستیک بودن کولموگوروف موجودند؛ به‌عنوان مثال، ر. ک. [۳۸، ۲۲] زیربخش ۲.۲.۲، [۹، ۳۳].

آیا نظریهٔ جدید فراوانی‌گراست؟ از بین مدلهای پیچیدگی که ذکر کردیم، تنها مثال ۱ برای تعبیر فراوانی‌گرای احتمال مستقیماً مورد توجه است. سایر مثالها گسترشهایی طبیعی از موضوع هستند، اما روشن است که از فراوانی‌گرایی فراتر می‌روند. به‌علاوه همان‌طورکه قبلاً متذکر شده‌ایم، وقتی که از مدلهای پیچیدگی برای پیشگویی استفاده می‌کنیم، نیازی به مفاهیم احتمال و فراوانی نیست.

کولموگوروف هرگز فراوانی‌گرایی را نفی نکرد، اما روشن است که نظریهٔ جدید به معنای مورد نظر فون میزس فراوانی‌گرا نیست. مشخصه‌های اساسی این نظریه به شرح زیر است

۱. این نظریه اکیداً متناهیانه است: تنها دنباله‌ها و مجموعه‌های اشیای ساختنی متناهی در نظر گرفته می‌شوند.

۲. این نظریه فرضی مشابه با اصل ب کولموگوروف را که می‌گفت پیشامدی با احتمال بسیار کوچک رخ نخواهد داد، در نظر می‌گیرد. حال می‌گوییم که یک پیشامد خاص، این پیشامد که برآمد تحقق‌یافته بسیار ساده است، رخ نخواهد داد.

۳. اصل الف، اصل فراوانی، نقشی اساسی ایفا نمی‌کند. توزیعهای احتمال نیز هیچ نقشی ایفا نمی‌کنند؛ به‌جای آنها مجموعه‌های متناهی قرار داده می‌شود.

1. stochasticity

دستاورد مارتین-لوف. در ۱۹۶۶، مارتین-لوف[۲۳] توجیه دیگری برای مفهوم کاستی تصادفی بودن کولموگوروف ارائه دادند که نشان می‌داد این مفهوم، یک آزمون آماری عام است. توسیع مفهوم آزمون آماری عام به حالت دنباله‌های نامتناهی ساده بود. بعدها، لوین[۲۱] و اشنور[۲۸]، نشان دادند که یک دنبالهٔ تصادفی مارتین-لوف است اگر و تنها اگر کاستی تصادفی بودن پاره‌های آغازین آن کراندار باشد (با این حال کاستی تصادفی بودن باید برحسب صورتی دیگر از پیچیدگی کولموگوروف، مانند پیچیدگی «یکنوا» یا «پیشوندی» تعریف شود).

۵. مارتینگل

رویکرد فراوانی‌گرای کولموگوروف که در [۱۵] مطرح شد، مبتنی بر ایده‌های فون‌میزس بود. ولی ژان ویل درکتاب ۱۹۳۹ خود[۳۴] به این نتیجه رسید که تعریف فون میزس از گردایه، به شکلی صوری که والد [۳۹] عرضه کرده بود، یک نقص جدی دارد: گردایه‌هایی وجود دارند که قانون لگاریتم مکرر را نقض می‌کنند؛ این امر در مورد صوری‌سازی چرچ در ۱۹۴۰ نیز صادق است. کولموگوروف به دنباله‌های متناهی علاقه‌مند بود، اما مثال ویل برای دنباله‌های طویل متناهی هم پیامدهای ناخوشایندی دارد. در[۱۵] کولموگوروف این مشکل را با مجاز دانستن قواعد گزینش زیردنباله‌ها برای تقطیع دنباله با ترتیب دلخواه، حل کرد، اما معلوم شد که راه حل ارائه‌شده توسط ویل خود کشفی متحول‌کننده بوده، و سرچشمهٔ عرصهٔ جدید و پرباری در نظریهٔ احتمال است. توجیه مطرح‌شده توسط فون میزس برای مفهوم گردایه‌اش، «اصل سیستم قماربازی با طرد» بود؛ ویل به طریقی بسیار طبیعی مفهوم فون میزس از سیستم قمار را بسط داد، مفهوم مارتینگل را معرفی، و مفهوم گردایه را اصلاح کرد. نکته‌ای اساسی که ویل آن را اثبات کرد، این بود که مثال او، یا مثالی مشابه بر پایهٔ هر قضیهٔ حدی قوی دیگری، برای گردایه‌های اصلاح‌شده غیرممکن است. (پیشتر از یک ویژگی عام مشابه یاد کردیم که مارتین-لوف[۲۳] آن را برای تعریف دوم کولموگوروف، مبتنی بر پیچیدگی کولموگوروف، ثابت کرد).

مقارن با ۱۹۷۱، اشنور[۲۶، ۲۷] چندین تعریف از تصادفی بودن به کمک مارتینگلها را مطرح کرد. نظریه‌ای مشابه، اما بدون استفادهٔ صریح از مارتینگلها، توسط لوین عرضه شد (ر. ک. تعریف لوین در [۲۱] از تصادفی بودن برحسب یک احتمال پیشین).

دوب[۸] ایده‌های ویل را در قالب احتمال مبتنی بر نظریهٔ اندازه شرح و گسترش داد و هم‌اکنون مارتینگلها در این نظریه نقشی اساسی دارند. با این حال، بدیهی است که مفهوم مارتینگل بیشتر مبتنی بر نظریهٔ بازیهاست تا نظریهٔ اندازه. ما درکتاب خود[۲۹] تاریخچه‌ای از ایده‌های مبتنی بر نظریهٔ بازیها در مبانی احتمال، که پیشینه‌اش به پاسکال برمی‌گردد، آورده و نشان داده‌ایم که چگونه می‌توان بدون متوسل شدن به نظریهٔ اندازه، هستهٔ کلاسیک نظریهٔ احتمال را مستقیماً بر پایهٔ مارتینگلهای مبتنی بر نظریهٔ بازیها بنا کرد.

دوباره احتمال مفهومی ثانویه می‌شود، اما حالا برحسب مارتینگلها تعریف شده است؛ ما نشان داده‌ایم که این شیوه می‌تواند در خدمت بسیاری از اهداف باشد. این رویکرد نه تنها برای قضایای حدی کلاسیک، بلکه برای بسیاری از کاربردهای احتمال در مسائل مالی نیز کاملاً مناسب است. در بخش بعد، این رویکرد را با یک مثال شرح می‌دهیم.

۶. ارزش اکتشافیِ تصادفی بودنِ الگوریتمی

نظریهٔ تصادفی بودنِ الگوریتمیِ کولموگوروف همچنان پژوهشهای جدید بسیار جالبی را برمی‌انگیزد (مثلاً ر.ک. [۹، ۳۳]). با این حال، شاید بیشترین فایدهٔ نظریهٔ الگوریتمی، استفاده از آن به‌عنوان ابزاری برای اکتشاف باشد تا زبانی برای توصیف ریاضی با کاربردهای عملی. این نظریه، چون از دیدگاه شهودی بسیار دقیق است، بیشترین کاربرد را در اکتشاف دارد. اما جزئیات ابزارهای ریاضی که این نظریه را چنین دقیق و صریح می‌گرداند ــ ثابتهای جمعی، محاسبه‌پذیری در حد، و غیره ــ هنگامی که می‌خواهیم به توصیف و کاربرد بپردازیم، دست‌وپاگیر می‌شوند. لذا به محض اینکه نتیجه‌ای ثابت شد، با حذف ایده‌های الگوریتمی از آن، می‌توان به دستاوردهای زیادی نائل شد. بدین طریق، تصادفی بودن الگوریتمی زایل می‌شود و نقش هدایتگر آن از دید استفاده‌کنندگان آتی نتیجه، پنهان می‌ماند.

چون معمولاً نتایج الگوریتمی حتی منتشر نمی‌شوند، ارائهٔ مثالهایی از این تبدیل در چارچوب یک مقالهٔ توصیفی ساده نیست. برای مثال، می‌توانیم به نتایج متعددی در کتاب خود[۲۹] به‌عنوان مثالهایی از نتایج الگوریتمی اشاره کنیم که در شکلی مبتنی بر نظریهٔ بازیها به تحریر در آمده‌اند، اما به علت فقدان آثار چاپ‌شده یا حتی توصیف صیقل‌یافته‌ای از نتایج الگوریتمی اصلی، این کار فقط برای عدهٔ کمی از خوانندگان روشنگر خواهد بود. با این حال یک مثال در اینجا می‌آوریم. برای سادگی و کنار گذاشتن ترجیحات فلسفی کولموگوروف، این مثال را از نوع نامتناهی برگزیده‌ایم.

نظریهٔ احتمال الگوریتمی نامتناهیانه، آن‌گونه که مارتین‌ـلوف آن را عرضه کرد، تمایز آشکاری بین دنباله‌های نامتناهی تصادفی و غیرتصادفی را مجاز می‌دارد که به قضیه‌های حدی قوی نقطه به نقطه منجر می‌شوند. برای مثال می‌توانیم قانون قوی اعداد بزرگِ بورل را با گفتن اینکه همهٔ دنباله‌های دودویی نامتناهی تصادفی $x_1 x_2 \ldots$ در

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{2} \qquad (۲)$$

صدق می‌کنند، دوباره بیان کنیم.

اشنور[۲۵، ۲۷] خیلی زود متوجه شد که (۲) برای بسیاری از دنباله‌های غیرتصادفی نیز برقرار است؛ در واقع برای آنکه (۲) برقرار باشد کافی است که برای کاستیِ تصادفی بودنِ پاره‌های آغازین $x_1 x_2 \ldots$ شرط آهنگ رشد زیرنمایی را قائل شویم. این نتیجه در [۳۵] به دو قضیهٔ حدی قوی دیگر تعمیم داده شده است: قانون لگاریتم مکرر و خاصیت بازگشتی. شرط اعمال‌شده بر آهنگ رشد کاستیِ تصادفی بودن برای این دو قانون ظریفتر، بسیار قوی‌تر است: برای مثال، یکی از نتایج [۳۵] این بود که اگر آهنگ رشد $o(\ln\ln n)$ باشد آنگاه

$$\limsup_{n\to\infty} \frac{1}{\sqrt{n\ln\ln n}}\left|\sum_{i=1}^{n} x_i - n/2\right| = \frac{1}{\sqrt{2}}.$$

از طرف دیگر، برای هر تابع $f(n) \to \infty$، وقتی $n \to \infty$، دنباله‌ای با آهنگ رشد $f(n)\ln\ln n$ موجود است به‌طوری که

$$\limsup_{n\to\infty} \frac{1}{\sqrt{n\ln\ln n}}\left|\sum_{i=1}^{n} x_i - n/2\right| = \infty.$$

همان‌طور که اشنور[۲۶] شرح می‌دهد، سریع‌ترین آهنگ رشد قابل قبول برای کاستیِ تصادفی بودن، اهمیت قضیهٔ حدی مورد نظر را نشان می‌دهد؛ این حقیقت که برای قانون قوی اعداد بزرگ این آهنگ می‌تواند تقریباً به سرعت نمایی باشد، حاکی از آن است که این قانون یکی از قضیه‌های حدی بسیار اساسی احتمال است.

حال توضیح می‌دهیم که چگونه ایدهٔ الگوریتمی رده‌بندی قضیه‌های حدی قوی را می‌توان در چارچوب نظریهٔ بازیها[۲۹] بیان کرد.[۱] به‌جای دنبالهٔ دودویی $x_1 x_2 \ldots$، دنباله‌ای از اعداد حقیقی را در نظر می‌گیریم که از نظر قدرمطلق کراندار باشد (بدون کاستن از کلیت، با کران ۱)، و به یک پیشگو نیز اجازه می‌دهیم که به‌ازای هر x_n، مقدار پیشگویی‌شدهٔ m_n را اعلام کند. قرارداد زیر را در نظر می‌گیریم.

بازی پیش‌بینی کراندار

بازیکنان: پیشگو، شکاک، واقعیت

قرارداد:

$\mathcal{K}_0 := 1$.

به‌ازای $n = 1, 2, \ldots$:

پیشگو $m_n \in [-1, 1]$ را اعلام می‌کند.

شکاک $M_n \in \mathbb{R}$ را اعلام می‌کند.

واقعیت $x_n \in [-1, 1]$ را اعلام می‌کند.

$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(x_n - m_n)$.

از لحاظ شهودی، m_n مقدار مورد انتظار پیشگو از حرکتهای واقعیت، x_n، است. این بدان معنی است که او آماده است هر تعداد حقیقی بلیت (مثبت، صفر یا منفی)، هرکدام به قیمت m_n، به شکاک بفروشد که حاضر به پرداخت x_n است. تعداد بلیتهایی که شکاک برای خرید انتخاب می‌کند M_n است و سرمایه‌اش در پایان دورِ nام بازی، $\mathcal{K}_n$ است (با فرض اینکه سرمایهٔ اولیه‌اش ۱ بوده است).

یک استراتژی را برای شکاک احتیاط‌آمیز می‌نامیم در صورتی که سرمایهٔ شکاک، $\mathcal{K}_n$، صرف‌نظر از اینکه پیشگو و واقعیت چه حرکتی می‌کنند، هرگز منفی نشود.

---

۱. قدم اساسی در ایجادِ خودِ این چارچوب، جرح و تعدیل تعریف مارتینگلی اشنور از کاستی تصادفی بودن برای در نظرگرفتن اصل prequential داوید[۶، ۷] بود. به محض اینکه تصویر صحنه روشن شد، کاستی تصادفی بودن رفع شد و ارجاع به نوشتگان احتمال الگوریتمی موضوعیت خود را از دست داد و حذف شد.

قضیهٔ ۱. شکاک در بازی پیشگویی کراندار استراتژی احتیاط‌آمیزی دارد به‌طوری که «به‌صورت نمایی ثروتمند» می‌شود؛ به این معنی که

$$\limsup_{n\to\infty} \frac{1}{n}\mathcal{K}_n > 0 \qquad (۳)$$

در مسیرهایی که

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}(x_i - m_i) = 0 \qquad (۴)$$

نقض شود.

برهان. بدون کاستن از کلیت، فرض می‌کنیم $m_n \equiv 0$. شکاک صرفاً می‌تواند از استراتژی $\mathcal{P}$ی زیر (توصیف شده در[۲۹]، ص۶۹؛ تمام این برهان، شکل وارونهٔ برهان قضیهٔ ۲.۳ در [۲۹] است) استفاده کند. به‌ازای هر عدد حقیقی ϵ، فرض کنید $\mathcal{P}_\epsilon$ این استراتژی باشد که همواره $\epsilon\alpha$ بلیت خریده می‌شود، که در آن α سرمایهٔ فعلی است. استراتژی $\mathcal{P}$ عبارت است از تقسیم سرمایهٔ اولیهٔ ۱ به دو دنباله از حسابها، A_k^+ و A_k^-، $k = 1, 2, \ldots$، با سرمایهٔ اولیهٔ 2^{-k-1} در حسابهای A_k^+ و A_k^-، سپس به‌کار بردن $\mathcal{P}_\epsilon$ با $\epsilon := 2^{-k}$ برای هر حساب A_k^+، و به‌کار بردن $\mathcal{P}_\epsilon$ با $\epsilon := -2^{-k}$ برای هر حساب A_k^-.

فرض کنید (۴) نقض شده است، مثلاً به‌ازای δای مثبت و بینهایت مقدار n

$$\frac{1}{n}\sum_{i=1}^{n} x_i > \delta .$$

اگر k را به‌قدری بزرگ اختیار کنیم که $\epsilon := 2^{-k}$ در $\epsilon < \delta/2$ صدق کند، می‌توانیم ببینیم که، برای بینهایت مقدار n

$$\frac{1}{n}\sum_{i=1}^{n} x_i > 2\epsilon .$$

تلفیق این نابرابری با $x_i^2 \leq 1$ ایجاب می‌کند که

$$\epsilon\sum_{i=1}^{n} x_i - \epsilon^2\sum_{i=1}^{n} x_i^2 > \epsilon^2 n .$$

علاوه بر این، با استفاده از نابرابری $t - t^2 \leq \ln(1+t)$ (که به‌ازای $t \geq -1/2$ درست است) به‌دست می‌آوریم

$$\sum_{i=1}^{n}\ln(1+\epsilon x_i) > \epsilon^2 n .$$

از آنجا که سرمایهٔ کسب شدهٔ $\mathcal{K}_n^{\mathcal{P}_\epsilon}$ با استراتژی $\mathcal{P}_\epsilon$ برابر با

$$2^{-k-1}\prod_{i=1}^{n}(1+\epsilon x_i)$$

است، می‌بینیم

$$\limsup_{n\to\infty}\frac{1}{n}\ln\mathcal{K}_n^{\mathcal{P}_\epsilon} > 0$$

که البته مستلزم (۳) است.

حالت دیگری که در آن (۴) نقض می‌شود، یعنی وقتی به‌ازای δای مثبت و بینهایت مقدار n

$$\frac{1}{n}\sum_{i=1}^{n} x_i < -\delta$$

به‌صورت مشابه بررسی می‌شود: k را آن‌قدر بزرگ اختیار می‌کنیم که $\epsilon := -2^{-k}$ در $|\epsilon| < \delta/2$ صدق کند. ■

قضیهٔ ۱ متناظر با صورتی از قانون قوی اعداد بزرگ است که اشنور در[۲۵] آورده و در بالا به آن اشاره شد. حکمهای مشابهی را هم می‌توان برای قانون لگاریتم مکرر و خاصیت بازگشتی ثابت کرد؛ وقتی که این قانونها نقض شده باشند، سرعت تضمین‌شدهٔ رشد سرمایهٔ شکاک البته بسیار کند خواهد بود.

حتی اگر حق با ما باشد که تصادفی بودن الگوریتمی عمدتاً به‌عنوان ابزاری برای کشف ارزشمند است، همین امر آموزش آن به محققان آینده را، در سطحی بسیار گسترده‌تر از امروز توجیه می‌کند.

مراجع

1. Eugene A. Asarin. Some properties of Kolmogorov δ-random finite sequences. *Theory of Probability and its Applications*, 32:507-508, 1987.
2. Eugene A. Asarin. On some properties of finite objects random in the algorithmic sense. *Soviet Mathematics Doklady*, 36:109-112, 1988.
3. Eugene A. Asarin. *On some properties of finite objects random in the algorithmic sense*. PhD thesis, Moscow State University, 1988. Academic advisor: Andrei N. Kolmogorov.
4. Jacob Bernoulli. *Ars Conjectandi*. Thurnisius, Basel, 1713.
5. Alonzo Church. On the concept of a random sequence. *Bulletin of American Mathematical Society*, 46(2):130-135, 1940.
6. A. Philip Dawid. Statistical theory: the prequential approach. *Journal of the Royal Statistical Society* A. 147-278-292, 1984.
7. A. Philip Dawid. Calibration based empirical probability (with discussion). *Annals of Statistics*, 13:1251-1285, 1985.
8. J. L. Doob. *Stochastic Processes*. Wiley, New York, 1953.
9. Peter Gács, J. Tromp. and Paul Vitányi. Algorithmic statistics. *IEEE Transactions on Information Theory*, 47 (6):2443-2463, 2001.
10. Andrei N. Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer, Berlin, 1933. English translation (1950): *Foundations of the theory of probability*. Chelsea. New York.
11. Andrei N. Kolmogorov. Теория вероятностей и ее применения. In Математика и естествознание в СССР, pages 51-61. GONTI, Moscow and Leningrad, 1938.

12. Andrei N. Kolmogorov. Вероятность. In Большая Советская Энник лопедия, volume 7, pages 508-510. Bol'shaya Sovetskaya Ehntsiklopediya, Moscow, 2nd edition, 1951.

13. Andrei N. Kolmogorov. Теория вероятностей. In Математика, ее содержание, методы и значение, volume 2, pages 252-284. Izdatel'stvo AN SSSR, Moscow, 1956.

14. Andrei N. Kolmogorov. Теория вероятностей. In Математика в СССР за сорок лет, volume 1, pages 781-795. Fizmatgiz, Moscow, 1959.

15. Andrei N. Kolmogorov. On tables of random numbers. *Sankhya. Indian Journal of Statistics A*, **25(4):369-376,** 1963.

16. Andrei N. Kolmogorov. Three approaches to the quantitative definition of information. *Problems of Information Transmission*, 1:1-7, 1965.

17. Andrei N. Kolmogorov. Logical basis for information theory and probability theory. *IEEE Transactions of Information Theory*, IT-14:662-664, 1968.

18. Andrei N. Kolmogorov. Combinatorial foundations of information theory and the calculus of probabilities. *Russian Mathematical Surveys*, 38(4):29-40, 1983.

19. Andrei N. Kolmogorov. On logical foundations of probability theory. In Yu. V. Prokhorov and K. Itô, editors, *Probability Theory and Mathematical Statistics*, volume 1021 of *Lecture Notes in Mathematics*, pages 1-5. Springer, 1983.

20. Andrei N. Kolmogorov and Boris V. Gnedenko. Теория вероятностей. In Математика в СССР за тридцать лет, pages 701-727. Gostekhizdat, Moscow and Leningrad, 1948.

21. Leonid A. Levin. On the notion of a random sequence. *Soviet Mathematics Doklady*, 14:1413, 1973.

22. Ming Li and Paul Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications.* Springer, New York, 2nd edition, 1997.

23. Per Martin-Löf. The definition of random sequences. *information and Control*, 9:602-619, 1966.

24. Richard von Mises. *Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik.* F. Deuticke, Leipzig and Vienna, 1931.

25. C. P. Schnorr. Klassifikation der Zufallsgesetze nach Komplexität und Ordnung. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 16:1-21, 1970.

26. C. P. Schnorr. A unified approach to the definition of random sequences. *Math Systems Theory*, 5:246-258, 1971.

27. C. P. Schnorr. *Zufälligkeit und Wahrscheinlichkeit.* Springer, Berlin, 1971.

28. C. P. Schnorr. Process complexity and effective random tests. *Journal of Computer and System Sciences*, 7(4):376-388. 1973.

29. Glenn Shafer and Vladimir Vovk. *Probability and Finance: It's Only a Game!* Wiley, New York, 2001.

30. Alexander Shen'. The concept of Kolmogorov (α, β)-stochasticity and its properties. *Soviet Mathematics Doklady*, 28:295-299, 1983.

31. Ray J. Solomonoff. A preliminary report on a general theory of inductive inference. Technical Report ZTB-138, Zator Company, Cambridge, MA, November 1960.

32. Ray J. Solomonoff. A formal theory of inductive inference. Parts I and II. *Information and Control*, 7:1-22 and 224-254. 1964.

33. Nikolai Vereshchagin and Paul Vitányi. Kolmogorov's structure functions with an application to the foundations of model selection. In *Proceedings of the 43rd Annual Symposium on Foundations of Computer Science.* IEEE Computer Society, 2002. To appear.

34. Jean Ville. *Etude critique de la notion de collectif.* Gauthier-Villars, Paris, 1939.

35. Vladimir Vovk. The law of the iterated logarithm for random Kolmogorov, or chaotic, sequences. *Theory of Probability and its Applications*, 32:413-425, 1987.

36. Vladimir Vovk. Kolmogorov's complexity conception of probability. In Vincent F. Hendricks, Stig Andur Pedersen. and Klaus Frovin Jϕ rgensen, editors, *Probability Theory: Philosophy. Recent History and Relations to Science*, Pages 51-69. Kluwer, Dordrecht. 2001.

37. Vladimir V. V'yugin. On nonstochastic objects. *Problems of Information Transmission*, 21:3-9, 1985.

38. Vladimir V. V'yugin. On the defect of randomness of a finite object with respect to measures with given complexity bounds. *Theory of Probability and its Applications*, 32:508-512, 1987.

39. Abraham Wald. Die Widerspruchfreiheit des Kollectivbegriffes der Wahrscheinlichkeitsrechnung. *Ergebnisse eines Mathematischen Kolloquiums*, 8:38-72, 1937.

* * * * * *

- Vladimir Vovk and Glenn Shafer, "Kolmogorov's contributions to the foundations of probability", *Problems of Information Transmission*, **39** (2003) 21-31.

* ولادیمیر ووک، دانشگاه لندن، بریتانیا
گلن شیفر، مدرسهٔ بازرگانی راتگرز، آمریکا