مجله‌ی ریاضی شریف
دوره‌ی سوم، شماره‌ی دوم (بهار ۱۴۰۳)

اثباتی برای نامتناهی‌بودن اعداد اول
با استفاده از نظریه‌ی اطلاعات

آلا جواهری

چکیده. در این نوشته اثباتی برای نامتناهی‌بودن اعداد اول ارائه خواهیم داد که از مفهوم انتروپی —که از مفاهیم بنیادین نظریه‌ی اطلاعات است— بهره می‌گیرد. افزون بر این، خواهیم دید که اثبات ما کران پایینی را نیز برای تابع شمارش اعداد اول فراهم می‌کند.

۱. مقدمه

نظریه‌ی اطلاعات بی‌شک از مهم‌ترین دستاوردهای علمی قرن بیستم است. کلاود شانون در ۱۹۴۸ با انتشار [۱] نظریه‌ای را بنیان نهاد که به مدد آن امروزه قادر هستیم گستره‌ای از فناوری‌ها را —از دیسک‌های فشرده[1] تا اینترنت 5G— در اختیار داشته باشیم. یکی از اساسی‌ترین کارهای شانون در [۱] این بود که روشی را برای کمی‌سازی مفهوم «اطلاعات» پیشنهاد داد —و البته باید در نظر داشت که رویکرد شانون تنها راه ممکن برای نیل به چنین مقصودی نیست. او بدین منظور از مفهوم انتروپی بهره گرفت. برای یک متغیر تصادفی گسسته‌ی X که مقادیر $x_۱,\dots,x_n$ را با احتمال $p_۱,\dots,p_n$ اخذ می‌کند، انتروپی X که با $H(X)$ نمایش داده می‌شود، به صورت

$$H(X) \equiv \sum_{i=۱}^{n} p_i \log \frac{۱}{p_i}$$

تعریف می‌شود، که در آن لگاریتم‌ها در پایه‌ی دو هستند. شانون نشان داد که کمیت فوق —که به تعبیری میانگین اطلاعات موجود در یک متغیر تصادفی است— تعیین‌کننده‌ی نرخ نهایی قابل حصول برای برخی از مهم‌ترین وظایف نظریه‌ی مخابرات —فشرده‌سازی داده و انتقال داده روی یک کانال در معرض نویز— است. با این همه، اهمیت کار شانون فقط به کاربردهای آن در نظریه‌ی مخابرات محدود نمی‌شود. ابزارهای نظریه‌ی اطلاعاتی را می‌توان در زمینه‌های متفاوتی از ریاضیات به کار گرفت. در ادامه به عنوان نمونه‌ای از این کاربردها، نامتناهی‌بودن اعداد اول را با به‌کارگیری چند نامساوی انتروپیک ثابت می‌کنیم. این قضیه از کهن‌ترین قضایای نظریه‌ی اعداد است که یونانیان باستان نیز از آن مطلع بودند، و اولین اثباتی که برای آن در دست است، اثباتی است در اصول، شاهکار ماندگار اقلیدس [۲]. با این حال از آن زمان تا به امروز اثبات‌های متعدد دیگری نیز برای این قضیه ارائه شده است. خواننده می‌تواند لیست مفصلی از این اثبات‌ها را در [۴] بیابد. اثبات ما در این نوشته برگرفته از اثباتی است که نخستین‌بار در [۳] بیان شده است.

۲. اثبات نامتناهی‌بودن اعداد اول

برای عدد طبیعی $n \geq ۲$، تعداد اعداد اول کوچکتر از یا مساوی با n را با $\pi(n)$ نمایش می‌دهیم. به این تابع، تابع شمارش اعداد اول[2] گفته می‌شود. فرض کنید $p_۱ < p_۲ < \cdots < p_{\pi(n)}$ اعداد اول کوچکتر از یا مساوی با n باشند.

N را یک متغیر تصادفی یکنواخت روی مجموعه‌ی $\{۱,۲,\dots,n\}$ در نظر بگیرید. از قضیه‌ی اساسی حساب می‌دانیم N تجزیه‌ی یکتایی به عوامل اول دارد. برای $i = ۱,\dots,\pi(n)$، متغیر تصادفی X_{p_i} را توان عدد اول p_i در تجزیه‌ی N تعریف

[1] CD
[2] Prime-counting function

کنید. از یکتایی تجزیه می‌توان نتیجه گرفت که توزیع توأم $(X_{p_1}, X_{p_2}, \dots, X_{p_{\pi(n)}})$ با توزیع N یکسان است. بنابراین

$$H(N) = H(X_{p_1}, X_{p_2}, \dots, X_{p_{\pi(n)}}).$$

از نحوه‌ی تعریف N می‌توان نتیجه گرفت $H(N) = \log n$. حال از یک نامساوی نظریه‌ی اطلاعاتی برای یافتن کران بالایی برای $H(N)$ استفاده می‌کنیم.

لم ۱.۲. برای متغیرهای تصادفی $X_1, X_2, \dots, X_k$،

$$H(X_1, X_2, \dots, X_k) \leq \sum_{i=1}^{k} H(X_i).$$

با استفاده از لم فوق داریم:

$$\log n = H(N) = H(X_{p_1}, X_{p_2}, \dots, X_{p_{\pi(n)}}) \leq \sum_{i=1}^{\pi(n)} H(X_{p_i}).$$

اکنون با استفاده از نامساوی انتروپیک دیگری کران بالایی برای هر $H(X_{p_i})$ ارائه می‌دهیم.

لم ۲.۲. فرض کنید X یک متغیر تصادفی گسسته است. اگر تعداد اعضای $\mathrm{Supp}(X)$ برابر با d باشد،

$$H(X) \leq \log d.$$

برای هر X_{p_i}، تعداد اعضای $\mathrm{Supp}(X_{p_i})$ کوچکتر از یا مساوی با $\log n + 1$ است. بنابراین داریم:

$$\log n = H(N) = H(X_{p_1}, X_{p_2}, \dots, X_{p_{\pi(n)}}) \leq \sum_{i=1}^{\pi(n)} H(X_{p_i}) \leq \pi(n) \log(\log n + 1),$$

که نتیجه می‌دهد

$$\frac{\log n}{\log(\log n + 1)} \leq \pi(n).$$

از آن‌جا که سمت چپ نامساوی فوق وقتی $n \to \infty$، به بی‌نهایت میل می‌کند، تعداد اعداد اول نمی‌تواند متناهی باشد. توجه کنید که در اثبات فوق نه تنها نشان دادیم که تعداد اعداد اول نامتناهی است، بلکه کران پایینی برای تابع شمارش اعداد اول پیدا کردیم.

می‌توان با کلک هوش‌مندانه‌ای کران پایینی را که در این اثبات برای $\pi(n)$ به دست آمده، بهتر کرد. فرض کنید در تجزیه‌ی N به عوامل اول، آن را به صورت $N = M \times {p_1}^{Y_{p_1}} \times \dots \times {p_{\pi(n)}}^{Y_{p_{\pi(n)}}}$ بنویسیم، که در آن M بزرگ‌ترین عدد مربع کاملی است که N را می‌شمارد. در این صورت $\mathrm{Supp}(M)$ حداکثر $\sqrt{n}$ عضوی است، و برای هر i، $\mathrm{Supp}(Y_{p_{\pi(i)}})$ دو عضوی است. لذا

$$\log n = H(N) = H(M, Y_{p_1}, Y_{p_2}, \dots, Y_{p_{\pi(n)}}) \leq H(M) + \sum_{i=1}^{\pi(n)} H(Y_{p_i}) \leq \frac{1}{2} \log n + \pi(n) \log 2,$$

که نتیجه می‌دهد

$$\frac{\log n}{2 \log 2} \leq \pi(n).$$

۳. اثبات لم‌ها

در این بخش برای اثبات لم‌های ۱.۲ و ۲.۲ کمیت انتروپیک جدیدی را معرفی می‌کنیم که به انتروپی نسبی دو متغیر تصادفی مشهور است. برای دو متغیر تصادفی گسسته‌ی X و Y که مقادیر $\omega_1, \dots, \omega_n$ را به ترتیب با احتمال‌های $p_1, \dots, p_n$ و $q_1, \dots, q_n$ اخذ می‌کنند، انتروپی نسبی از X به Y، که آن را با $D(X||Y)$ نمایش می‌دهیم، به صورت

$$D(X||Y) \equiv \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i},$$

تعریف می‌شود[1]. ویژگی مهم این کمیت آن است که نامنفی است. برای اثبات نامنفی‌بودن، از نامساوی $\ln x \leq x - ۱$ استفاده می‌کنیم. داریم:

$$\sum_i p_i \ln \frac{q_i}{p_i} \leq \sum_i p_i \left(\frac{q_i}{p_i} - ۱\right) = \sum_i q_i - ۱ = ۰.$$

بنابراین $\sum_i p_i \log \frac{p_i}{q_i} \geq ۰$.

حال می‌توان دید که اثبات لم ۲.۲ نتیجه‌ی سرراستی از نامنفی‌بودن انتروپی نسبی است. کافی است برای متغیر تصادفی X که مقادیر $x_۱, \ldots, x_d$ را با احتمال ناصفر اخذ می‌کند، متغیر تصادفی Y را به عنوان متغیر تصادفی یکنواخت روی $\mathrm{Supp}(X)$ تعریف کنیم. به این ترتیب داریم:

$$۰ \leq D(X||Y) = \sum_i p_i \log \frac{p_i}{\frac{۱}{d}} = \sum_i p_i \log p_i - \sum_i p_i \log \frac{۱}{d} = \sum_i p_i \log p_i - \log \frac{۱}{d},$$

که نتیجه می‌دهد $\log \frac{۱}{d} \leq \sum_i p_i \log p_i$ یا **معادلاً** $H(X) \leq \log d$.

برای اثبات لم ۱.۲ کافی است حکم را برای حالت $k = ۲$ ثابت کنیم. بدین منظور، متغیرهای تصادفی **مستقل** X' و Y' را به ترتیب **هم‌توزیع** با X و Y تعریف می‌کنیم. به این ترتیب داریم:

$$\begin{aligned}
۰ \leq D(X,Y||X',Y') &= \sum_{i,j} \Pr[X = x_i, Y = y_j] \log \frac{\Pr[X = x_i, Y = y_j]}{\Pr[X' = x_i, Y' = y_j]} \\
&= \sum_{i,j} \Pr[X = x_i, Y = y_j] \log \Pr[X = x_i, Y = y_j] - \sum_{i,j} \Pr[X = x_i, Y = y_j] \log \Pr[X' = x_i, Y' = y_j] \\
&= -H(X,Y) - \sum_{i,j} \Pr[X = x_i, Y = y_j] \log \Pr[X' = x_i] \Pr[Y' = y_j] \\
&= -H(X,Y) - \sum_{i,j} \Pr[X = x_i, Y = y_j] \log \Pr[X' = x_i] - \sum_{i,j} \Pr[X = x_i, Y = y_j] \Pr[Y' = y_j] \\
&= -H(X,Y) - \sum_{i} \Pr[X = x_i] \log \Pr[X' = x_i] - \sum_{j} \Pr[Y = y_j] \Pr[Y' = y_j] \\
&= -H(X,Y) + H(X) + H(Y)
\end{aligned}$$

که نتیجه می‌دهد $H(X,Y) \leq H(X) + H(Y)$.

مراجع

[1] Shannon, C. E. (1948). A mathematical theory of communication. The Bell system technical journal, 27(3), 379-423.

[2] Heath, T. L. (Ed.). (1956). The thirteen books of Euclid’s Elements. Courier Corporation.

[3] Chaitin, G. J. (1977). Toward a Mathematical Definition of Life, 2. IBM Thomas J. Watson Research Division.

[4] Meštrović, R. (2012). Euclid’s theorem on the infinitude of primes: a historical survey of its proofs (300 BC–2022) and another new proof. arXiv preprint arXiv:1202.3670.

[5] Cover, T. M. (1999). Elements of information theory. John Wiley & Sons.

[1] تعریفی که در این‌جا ارائه کرده‌ایم، حالت خاصی از تعریف رایج در ادبیات نظریه‌ی اطلاعات است. برای اطلاع از تعریف رایج به [5] مراجعه کنید.