نابرابریهای احتمالات

امید اعتصامی
پژوهشگاه دانشهای بنیادی، پژوهشکده ریاضیات

۱ انگیزه

خیلی وقتها ما مقدار دقیق احتمال یک پیشامد را نمی‌دانیم ولی می‌خواهیم نشان دهیم وقوع آن پیشامد نادر است (احتمال آن کم است) یا وقوع آن پیشامد نادر نیست (احتمال آن زیاد است).

روش‌های زیادی وجود دارد که بدون بدست آوردن احتمال یک پیشامد، کران پایین یا بالا برای احتمال آن پیشامد به ما می‌دهد.

۲ نابرابری مارکوف

قضیه ۱ فرض کنید f تابع نامنفی است که هرگاه $x \in S$ داریم $f(x) \geq s > 0$. آنگاه

$$\Pr(X \in S) \leq \mathbb{E}\, f(x)/s.$$

اثبات:

$$\mathbb{E}\, f(x) \geq \Pr(x \in S) \cdot s + \Pr(x \notin S) \cdot 0.$$

□

مثلا

$$\Pr(|X| \geq s) \leq \mathbb{E}\, |X|/s.$$

خود مارکوف درباره‌ی حالت خاص قضیه‌ی بالا وقتی $f(x) = |x|^a$ بحث می‌کند. حالت $f(x) = e^{ax}$ هم بسیار مفید است.

۳ نابرابری چبیشف

اگر در نابرابری مارکوف قرار دهیم $f(x) = (X - \mathbb{E}X)^2$ نابرابری چبیشف بدست می‌آید.

قضیه ۲

$$\Pr(|X - \mathbb{E}X| \geq r) \leq \mathrm{var}(X)/r^2.$$

۴ نامساوی ینسن [۱]

تابع حقیقی f روی بازه‌ی I از اعداد حقیق محدب خوانده می‌شود هرگاه به ازای هر $x, y \in I$ و $p, q \geq 0$ به طوری که $p + q = 1$ داشته باشیم

$$f(px + qy) \leq pf(x) + qf(y).$$

قضیه ۳ اگر f تابعی محدب روی بازه‌ی I باشد و متغیر تصادفی X همواره مقداری عضو I داشته باشد، آنگاه

$$f(\mathbb{E}X) \leq \mathbb{E}f(X).$$

اگر f مشتق دوم داشته باشد، محدب بودن f معادل نامنفی بودن مشتق دوم است. به عنوان مثال e^{ax} به ازای هر ثابت a و x^{2n} به ازای هر عدد صحیح نامنفی محدبند.

به ازای $x > 0$، تابع x^n به ازای هر عدد صحیح n محدب است. علی الخصوص $1/x$ محدب است.

به ازای $x > 0$، توابع $\ln(1/x)$ و $x \ln x$ محدبند.

۵ اصل گشتاور اول

قضیه ۴ اگر X متغیر تصادفی با مقادیر صحیح نامنفی باشد، آنگاه

$$\Pr(X > 0) \leq \mathbb{E}X.$$

اثبات:

$$\Pr(X > 0) = \sum_{n \geq 1} \Pr(X = n) \leq \sum_{n \geq 1} n \Pr(X = n) = \mathbb{E}X.$$

$\square$

گرافی با مجموعه رئوس $\{1, \ldots, n\}$ در نظر بگیرید که هر یال در آن به احتمال p آمده است. احتمال این که گراف حداقل یک مثلث داشته باشد چیست؟ با کمک قضیه‌ی بالا، کران بالایی برای این احتمال بدست می‌آوریم.

به ازای $1 \leq u < v < w \leq n$، متغیر تصادفی X_{uvw} را متغیر تصادفی نشانگری تعریف کنید که برابر ۱ است اگر هر سه یال (u, v) و (v, w) و (w, u) حضور داشته باشند و در غیر این صورت $X_{uvw} = 0$. احتمال مورد نظر برابر است با

$$\Pr(\sum_{u,v,w} X_{uvw} > 0) \leq \mathbb{E} \sum_{u,v,w} X_{uvw} = \sum_{u,v,w} \mathbb{E}X_{uvw} = \binom{n}{3} p^3.$$

۶ اصل گشتاور دوم

اصل گشتاور دوم، بر خلاف اصل گشتاور اول که کران بالا می‌داد، برای مثبت بودن متغیر تصادفی کران پایین می‌دهد.

قضیه ۵ اگر X متغیر تصادفی نامنفی باشد، آنگاه

$$\Pr(X > 0) \geq (\mathbb{E}X)^2 / (\mathbb{E}X^2).$$

اثبات:

$$\mathbb{E}X^2 = \Pr(X > 0)\mathbb{E}(X^2 | X > 0) \geq \Pr(X > 0)(\mathbb{E}(X | X > 0))^2 = \Pr(X > 0)((\mathbb{E}X) / \Pr(X > 0))^2.$$

$\square$

مثال گراف با مجموعه رئوس $\{1, \ldots, n\}$ را به یاد بیاورید که هر یال در آن به احتمال p آمده است. این بار می‌خواهیم برای احتمال وجود حداقل یک مثلث کران پایین بدست بیاوریم. تعریف کنید

$$X = \sum_{1 \leq u < v < w \leq n} X_{uvw}$$

که در آن X_{uvw} همچون بالا تعریف شده است. گشتاور دوم X را این گونه محاسبه می‌کنیم:

$$\begin{aligned}
\mathbb{E}X^2 &= \mathbb{E}(\sum_{u,v,w} X_{uvw})(\sum_{u',v',w'} X_{u'v'w'}) \\
&= \sum_{u,v,w,u',v',w'} \mathbb{E}(X_{uvw} X_{u'v'w'}) \\
&= \sum_{k=0}^{3} \sum_{\substack{u,v,w,u',v',w' \\ |\{u,v,w\} \cap \{u',v',w'\}| = k}} \mathbb{E}(X_{uvw} X_{u'v'w'}) \\
&= 20\binom{n}{6} p^6 + 30\binom{n}{5} p^6 + 12\binom{n}{4} p^5 + \binom{n}{3} p^3.
\end{aligned}$$

حال داریم

$$\Pr(X > 0) \geq (\mathbb{E}X)^2 / (\mathbb{E}X^2) = \frac{\binom{n}{3}^2 p^6}{20\binom{n}{6} p^6 + 30\binom{n}{5} p^6 + 12\binom{n}{4} p^5 + \binom{n}{3} p^3}.$$

۷ نابرابری امید ریاضی شرطی

نابرابری امید ریاضی شرطی راس [۱] [۲] در حالتهایی خاص‌تر از اصل گشتاور دوم کار می‌کند، اما در این حالت‌ها همواره کران بهتر یا مساوی اصل گشتاور دوم به ما می‌دهد.

قضیه ۶ اگر $X = X_1 + \ldots + X_n$ که در آن X_iها متغیرهای تصادفی دودویی هستند، آنگاه

$$\Pr(X > 0) \geq \sum_{j=1}^{n} \frac{\mathbb{E}X_j}{\mathbb{E}(X|X_j = 1)}.$$

اثبات: چون تابع $f(x) = 1/x$ به ازای $x > 0$ محدب است، طبق نامساوی ینسن داریم:

$$\begin{aligned}
\Pr(X > 0) &= \sum_{j=1}^{n} \Pr(X_j > 0)\mathbb{E}(X_j/X|X_j > 0) \\
&= \sum_{j=1}^{n} (\mathbb{E}X_j)\mathbb{E}(1/X|X_j = 1) \\
&\leq \sum_{j=1}^{n} (\mathbb{E}X_j)\frac{1}{\mathbb{E}(X|X_j = 1)}.
\end{aligned}$$

□

۸ کران پایین برای چندجمله‌ای قابلیت اطمینان

فرض کنید $f(y_1, \ldots, y_n)$ تابعی با ورودی‌های دودویی و خروجی دودویی باشد، مثلا

$$f(y_1, \ldots, y_5) = y_1y_2y_3 \vee y_2y_3y_4 \vee y_4y_5.$$

اگر ورودی‌های f متغیرهای تصادفی دودویی مستقل Y_1، ...، Y_n باشند به طوری که $p_i = \Pr(Y_i = 1) = \mathbb{E}Y_i$، آنگاه

$$\Pr(f(Y_1, \ldots, Y_n) = 1)$$

چندجمله‌ای با متغیرهای p_1، ...، p_n خواهد بود که به آن چندجمله‌ای قابلیت اطمینان [۲] می‌گویند. (علت این نامگذاری این است که اگر یک سیستم دارای n مولفه باشد که y_i نشان‌دهنده‌ی کار کردن یا نکردن مولفه‌ی iام باشد، و $f(y_1, \ldots, y_n)$ نشان‌دهنده‌ی کار کردن کل سیستم باشد، و مولفه iام با احتمال p_i کار کند، آنگاه چندجمله‌ای قابلیت اطمینان میزان اطمینان از کل سیستم (احتمال کار کردن کل سیستم) را نشان می‌دهد.)

در مثال تابع پنج متغیره‌ی بالا، تعریف کنید

$$X_1 = Y_1Y_2Y_3, \;\; X_2 = Y_2Y_3Y_4, \;\; X_3 = Y_4Y_5, \;\; X = X_1 + X_2 + X_3.$$

آنگاه طبق نابرابری امید ریاضی شرطی، برای چندجمله‌ای قابلیت اطمینان که همان $\Pr(X > 0)$ است داریم:

$$\Pr(X > 0) \geq \frac{p_1p_2p_3}{1 + p_4 + p_4p_5} + \frac{p_2p_3p_4}{p_1 + 1 + p_5} + \frac{p_4p_5}{p_1p_2p_3 + p_2p_3 + 1}.$$

[۱] Ross conditinoal expectation inequality
[۲] reliability polynomial

۹ نابرابری FKG

دو خاصیت از گرافها را در نظر بگیرید که نسبت به یالها یکنوای صعودی [۳] باشند (یعنی وقتی گرافی آن خاصیت را داشته باشد، با افزودن به یالها همچنان آن خاصیت را داشته باشد.) به عنوان مثال خاصیت اول را هم‌بند بودن گراف در نظر بگیرید و خاصیت دوم را چهار رنگ ناپذیر بودن. آیا می‌توان گفت که این دو خاصیت با هم هم‌بستگی مثبت [۴] دارند؟ (یعنی اگر با نماد ایورسون

$$f = [\text{graph is connected}], \quad g = [\text{graph is not 4-colorable}]$$

آنگاه

$$\mathbb{E}(fg) \geq (\mathbb{E}f)(\mathbb{E}g)$$

را داریم؟)

نابرابری FKG منسوب به C. M. Fortuin و P. W. Kasteleyn و J. Ginibre به ما می‌گوید که زمانی که یالها به طور مستقل از هم انتخاب شده باشند آنگاه این هم‌بستگی مثبت وجود دارد، البته این نابرابری کلی‌تر است.

قضیه ۷ فرض کنید یک زیر مجموعه S از مجموعه‌ی متناهی U انتخاب می‌شود. احتمال انتخاب S را با $\mu(S)$ نمایش می‌دهیم. فرض کنید به ازای هر S و T زیرمجموعه‌ی U داشته باشیم

$$\mu(S)\mu(T) \leq \mu(S \cap T)\mu(S \cup T).$$

در نهایت فرض کنید دو تابع f و g از زیرمجموعه‌های U به اعداد حقیقی یکنوای صعودی‌اند (یعنی اگر $T \subseteq S$ آنگاه $f(S) \leq f(T)$ و $g(S) \leq g(T)$). آنگاه

$$\mathbb{E}(fg) \geq (\mathbb{E}f)(\mathbb{E}g).$$

مراجع

[1] Jensen, J. L. W. V. (1906). Sur les fonctions convexes et les inégalités entre les valeurs moyennes. Acta Mathematica 30 (1): 175–193.

[2] Ross, S. M. (1994). Probability, Statistics, and Optimization (New York, Wiley), pages 185-190.

[۳] monotonic
[۴] positive correlation