دانشکده‌ی علوم ریاضی

احتمال و کاربرد آن

جلسه‌ی ۲۱: نامساوی‌های احتمالاتی

مدرّس: دکتر شهرام خزائی
نگارنده: آرمیتا ثابتی‌اشرف

۱ مقدمه

تا اینجا دیدیم که برای محاسبه‌ی مقدار یک احتمال، به تابع توزیع احتمال، تابع جرم احتمال و یا تابع چگالی احتمال نیاز داریم. اغلب پیش می‌آید که برای برخی متغیرهای تصادفی، قادر به تعیین هیچ‌یک از این توابع نیستیم، اما می‌توانیم امید ریاضی و یا واریانس آن‌ها را حساب کنیم. در این موارد با اینکه نمی‌توانیم مقدار دقیق احتمال را محاسبه کنیم، امّا با استفاده از نامساوی مارکوف[۱] و نامساوی چبیشف[۲] می‌توان کرانی برای مقدار احتمال به‌دست آورد.
بررسی این دو نامساوی و نیز همگرایی در احتمال[۳] موضوع بحث این جلسه است.

۲ نامساوی مارکوف

برای متغیرهای تصادفی نامنفی، بدون دانستن توزیع آن و فقط با استفاده از مقدار متوسط، به کمک نامساوی مارکوف می‌توان کران بالایی برای این متغیر به دست آورد.

قضیه ۱ *اگر X متغیر تصادفی نامنفی باشد، برای هر $a > ۰$ داریم:*

$$\Pr\{X \geq a\} \leq \frac{\mathrm{E}[X]}{a}$$

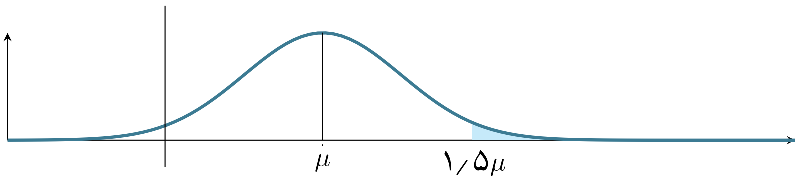
برهان. اثبات درستی این نامساوی برای متغیرهای تصادفی گسسته X، با تابع جرم احتمال $p(x)$ به صورت زیر است:

$$\begin{aligned}\mathrm{E}[X] = \sum_{x>۰} xp(x) &\geq \sum_{x\geq a} xp(x) \\ &\geq \sum_{x\geq a} ap(x) \\ &= a\sum_{x\geq a} p(x) \\ &= a\Pr\{x \geq a\}\end{aligned}$$

■

[۱] Markov's inequality
[۲] Chebyshev's inequality
[۳] Convergence in probability

مثال ۱ *در نمودار شبه گاوسی زیر، احتمال این که متغیر تصادفی بیشتر از $\text{۱٫۵}\mu$ از متوسط فاصله داشته باشد، کمتر از $\frac{۲}{۳}$ است. واضح است که این کران خوبی نیست.*



۳ نامساوی چبیشف

قضیه ۲ *اگر X متغیری تصادفی با مقدار متوسط متناهی μ و واریانس $\sigma^۲$ باشد، آنگاه برای هر $k > ۰$ داریم:*

$$\Pr\{|X-\mu| \geq k\} \leq \frac{\sigma^۲}{k^۲}$$

برهان. از آنجا که $(X-\mu)^۲$ یک متغیر تصادفی نامنفی است، با قرار دادن $a = k^۲$ در نامساوی مارکوف خواهیم داشت:

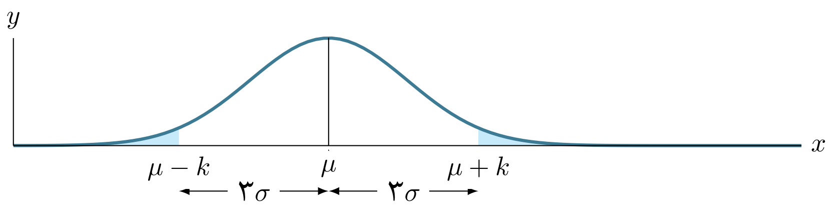
$$\Pr\{(X-\mu)^۲ \geq k^۲\} \leq \frac{\mathrm{E}[(X-\mu)^۲]}{k^۲}$$

چون $(X-\mu)^۲ \geq k^۲$ فقط در صورتی که $|X-\mu| \geq k$ باشد، برقرار است، خواهیم داشت:

$$\Pr\{|X-\mu| \geq k\} \leq \frac{\mathrm{E}[(x-\mu)^۲]}{k^۲} = \frac{\sigma^۲}{k^۲}$$

■

مثال ۲ *در نمودار توزیع گاوسی زیر، اگر k را برابر ۳σ در نظر بگیریم، احتمال این که متغیر تصادفی با مقدار متوسط کمتر از سه برابر انحراف معیار از میانگین فاصله داشته باشد، بیشتر از $\frac{۸}{۹}$ است. مقدار دقیق این احتمال $۲\phi(-۳) \approx ۰٫۰۰۲۶$ است.*



۴ همگرایی در احتمال

قضیه ۳ (قانون ضعیف اعداد بزرگ[۴]) *اگر $X_1, X_2, X_3, \ldots$ دنباله‌ای از متغیرهای تصادفی مستقل با توزیع یکسان[5] باشند و برای $i = 1, 2, \ldots$ داشته باشیم $\mu = \mathrm{E}[X_i]$ و $\sigma^2 = Var(X_i) < \infty$، آنگاه برای هر $\epsilon > 0$ خواهیم داشت:*

$$\lim_{n\to\infty} P\Big(\Big|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\Big| > \epsilon\Big) = 0.$$

تعریف ۱ (همگرایی در احتمال[۶]) *اگر $X_1, X_2, \ldots$ دنباله‌ای از متغیرهای تصادفی پیوسته باشد، گوییم X_n در احتمال به یک متغیر تصادفی X همگرا می‌شود، اگر برای هر $\epsilon > 0$ داشته باشیم:*

$$\lim_{n\to\infty} \Pr\{|X_n - X| > \epsilon\} = 0.$$

از قانون ضعیف اعداد بزرگ می‌توانیم برای تخمین پارامترها استفاده نماییم.

مثال ۳ *کسر f از افراد جامعه نوشابه را به دلستر ترجیح می‌دهند. مقدار f را چگونه می‌توان تخمین زد؟*
برای تخمین f از n نفر به طور تصادفی نمونه‌گیری کرده، متغیر تصادفی X_i را به این صورت تعریف می‌کنیم:

$$X_i = \begin{cases} 1 & \text{اگر شخص } i\text{ام بگوید بله} \\ 0 & oth. \end{cases}$$

چون X_iها توزیع برنولی دارند، برای متغیر تصادفی $M_n = \dfrac{X_1 + \cdots + X_n}{n}$ داریم:

$$\mathrm{E}[M_n] = f\ .$$

بنابراین یک نمونه از متغیر تصادفی M_n را می‌توان به عنوانی تخمینی از f در نظر گرفت. دقت کنید که انجام دفعات مختلف آزمایش، یعنی پرسش از n شخص تصادفی و محاسبه کسری از افراد که جواب مثبت داده‌اند، به تخمین‌های مختلفی برای f می‌انجامد. مطلوب هر مسئله‌ی تخمین، نزدیک بودن جواب به‌دست آمده به جواب واقعی است. برای تعیین کیفیت مقدار تخمین، احتمال زیر را در نظر بگیرید:

$$\Pr\{|M_n - f| \geq \epsilon\} \leq 1 - \alpha$$

که ϵ مقدار خطا و α سطح اطمینان است.
اگر بخواهیم تخمین نهایی، با خطای کمتر از ۰/۰۱ $(\epsilon = 0/01)$ و سطح اطمینان ۰/۹۵ $(\alpha = 0/95)$ به‌دست بیاید، باید داشته باشیم:

$$\Pr\{|M_n - f| \geq 0/01\} \leq 1 - 0/95$$

با استفاده از نامساوی چبیشف داریم:

$$\begin{aligned}\Pr\{|M_n - f| \geq \epsilon\} &\leq \frac{Var(M_n)}{\epsilon^2} = \frac{Var(X_i)}{n\epsilon^2} = \frac{f(1-f)}{n\epsilon^2} \\ &\leq \frac{1}{4n\epsilon^2} \leq (1-\alpha) \Rightarrow \\ n &\geq \frac{1}{4\epsilon^2(1-\alpha)} \Rightarrow \\ n &\geq 50000\end{aligned}$$

[5] Independent and identically distributed random variables (i.i.d)

برای این که جامعه آماری را کوچک‌تر کنیم، می‌توانیم ϵ را کمی زیاد کنیم تا دقت کمی کمتر شود. مثلا اگر $\epsilon = ۰/۰۳$ باشد، n تقریباً ۵۰۰۰ به دست می‌آید. همواره باید سعی کنیم میان ϵ و α تعادل برقرار شود. به ازای هر n و ϵ دل‌خواه، می‌توان α (سطح اطمینان) را محاسبه کرد. یک راه دیگر برای کمتر کردن تعداد نمونه‌های مورد نیاز بدون تغییر دادن کیفیت تخمین، پیدا کردن کران مناسب‌تری برای احتمال فوق است. این مسأله در جلسه آینده بررسی می‌شود.