دانشکده‌ی علوم ریاضی

احتمال و کاربرد آن ۱۴ اردیبهشت ۹۳

جلسه‌ی ۱۹: بررسی چند مسئله

مدرّس: دکتر شهرام خزائی *نگارنده: میلاد پورسلطانی*

۱ هدف

در این جلسه، به کمک ویژگی خطی‌بودن مقدار متوسط احتمال به حل چند مثال می‌پردازیم. همچنین با استفاده از مفهوم جذر متوسط مجذور فاصله[1] مسئله قدم‌زدن تصادفی[2] را بررسی می‌کنیم. در انتها، به کمک یک مثال، به تفصیل توزیع مشترک توابعی از دو متغیر تصادفی را محاسبه خواهیم کرد.

۲ کاربرد خطی بودن مقدار متوسط احتمال

سؤال ۱ *چند بار یک تاس را پرتاب کنیم تا با «احتمال خوبی» همه‌ی وجوه آن مشاهده شود؟*

نکته ۱ *برای دستیابی به احتمال خوب، مقدار متوسط تعداد پرتاب‌های لازم معیار مناسبی است. این مطلب از نامساوی مارکوف که در انتهای درس مطرح خواهد شد نتیجه می‌شود. طبق نامساوی مارکوف برای متغیر تصادفی نامنفی X و مقدار $a > ۰$ رابطه‌ی*

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

برقرار است. بنابراین با انتخاب مناسب ضریب k و جایگزینی مقدار $a = k\mathrm{E}[X]$ در نامساوی مارکوف می‌توان به احتمال به اندازه دلخواه خوب

$$\Pr\{X < k\mathrm{E}[X]\} > ۱ - \frac{۱}{k}$$

دست یافت. مثلاً برای $k = ۳$ داریم $\Pr\{X < ۳\mathrm{E}[X]\} > \frac{۲}{۳}$.

در اینجا حالت کلی‌تر این مسئله را بررسی می‌کنیم.

مثال ۱ *در کیسه‌ای N توپ داریم که از ۱ تا N شماره‌گذاری شده‌اند. هر بار به تصادف توپی از کیسه برمی‌داریم، شماره‌ی آن را یادداشت کرده و دوباره به کیسه باز می‌گردانیم. اگر متغیر تصادفی T را برابر تعداد دفعاتی که باید آزمایش انجام شود تا همه‌ی شماره‌ها مشاهده شود، تعریف کنیم، مطلوب است محاسبه‌ی* $\mathrm{E}[T]$.

[1] Root mean square distance (RMSD)
[2] Random walks

پاسخ برای حل مسئله، متغیر تصادفی X_j را بدین شکل تعریف می‌کنیم: تعداد توپ‌هایی که پس از مشاهده‌ی $j-۱$ امین توپ جدید (با شماره‌ی غیر تکراری) باید از کیسه خارج کنیم تا j امین توپ جدید مشاهده شود.
X_j دارای توزیع هندسی با پارامتر $p_j = \frac{N-(j-۱)}{N}$ است. در نتیجه:

$$\mathrm{E}[X_j] = \frac{۱}{p_j} = \frac{N}{N-(j-۱)}$$

از طرفی، به روشنی $T = \sum_{j=۱}^{N} X_j$. پس بنابر خطی بودن مقدار متوسط، داریم:

$$\begin{aligned}
\mathrm{E}[T] &= \mathrm{E}[\sum_{j=۱}^{N} X_j] \\
&= \sum_{j=۱}^{N} \mathrm{E}[X_j] \\
&= \sum_{j=۱}^{N} \frac{N}{N-(j-۱)} \\
&= N\sum_{i=۱}^{N} \frac{۱}{i}
\end{aligned}$$

از آنجا که $\ln(N+۱) < \sum_{i=۱}^{N} \frac{۱}{i} < \ln N + ۱$ پس برای N های بزرگ با تقریب خوبی می‌توان گفت:

$$\mathrm{E}[T] \approx N \ln N$$

۳ کاربرد جذر متوسط مجذور فاصله

مثال ۲ (قدم‌زدن تصادفی) شخصی مست با شروع از مبدأ، بر یک محور افقی هر بار یک گام (به صورت کاملاً تصادفی و مستقل از گام‌های قبلی) به سمت راست یا چپ برمی‌دارد. این فرد پس از N گام در چه فاصله‌ای از مبدأ قرار خواهد داشت؟
متغیر تصادفی S_N را موقعیت شخص پس از طی N گام تعریف می‌کنیم. هدف ما یافتن مقدار $\mathrm{E}[|S_N|]$ است. (با توجه به تقارن مسئله، واضح است که $\mathrm{E}[S_N] = ۰$ است.)
متغیر تصادفی X_j را اینگونه تعریف می‌کنیم:

$$X_j = \begin{cases} +۱ & \text{اگر شخص در گام } j \text{ام به سمت راست برود} \\ -۱ & \text{اگر شخص در گام } j \text{ام به سمت چپ برود} \end{cases}$$

طبق این تعریف، واضح است:

$$S_N = \sum_{j=۱}^{N} X_j$$

با استفاده از ویژگی خطی بودن مقدار متوسط، می‌توان نوشت:

$$\begin{aligned}\mathrm{E}[S_N^{۲}] &= \mathrm{E}\Big[\Big(\sum_{j=۱}^{N} X_j\Big)^{۲}\Big] \\ &= \mathrm{E}\Big[\sum_{j=۱}^{N} X_j^{۲} + ۲ \sum\sum_{۱\leq i<j\leq N} X_i X_j\Big] \\ &= \sum_{j=۱}^{N} \mathrm{E}[X_j^{۲}] + ۲\sum\sum_{i<j} \mathrm{E}[X_i X_j]\end{aligned}$$

برای هر i، $\mathrm{E}[X_i^{۲}] = ۱$ است، چرا که X_i تنها یکی از دو مقدار ۱+ و یا ۱− را می‌گیرد و مجذور هر دو، برابر ۱ است. می‌توان نتیجه گرفت:

$$\sum_{j=۱}^{N} \mathrm{E}[X_j^{۲}] = N \qquad (۱)$$

و نیز برای هر $i < j$:

$$\mathrm{E}[X_i X_j] = ۰ \qquad (۲)$$

از (۱) و (۲) نتیجه می‌شود:

$$\mathrm{E}[S_N^{۲}] = N \qquad (۳)$$

برای ادامه‌ی حل، به مفهوم تازه‌ای به نام جذر متوسط مجذور فاصله [۳] نیاز داریم که بدین شکل تعریف می‌شود:

$$RMSD = \sqrt{\mathrm{E}[|S_N|^{۲}]} \qquad (۴)$$

این مقدار در بسیاری از موارد به عنوان معیاری (نوعی) از فاصله کفایت می‌کند و تقریب خوبی برای $\mathrm{E}[|S_N|]$ است. پس از (۳) و (۴)، پاسخ تقریبی مسئله برابر $\sqrt{N}$ است.

نکته ۲ می‌دانیم:

$$\mathrm{E}[X^{۲}] - (\mathrm{E}[X])^{۲} = Var(X) \geq ۰ \quad \Rightarrow \quad (\mathrm{E}[X])^{۲} \leq \mathrm{E}[X^{۲}]$$

پس در این مسئله داریم:

$$\mathrm{E}[|S_N|] \leq \sqrt{N}$$

می‌توان نشان داد:

$$\lim_{N\to\infty} \mathrm{E}[|S_N|] = \sqrt{\frac{۲N}{\pi}}$$

یعنی اگر N به اندازه‌ی کافی بزرگ باشد، دقیق‌تر آن است که $\sqrt{\frac{۲N}{\pi}}$ به عنوان جواب اعلام گردد.

مثال ۳ (تعمیم مثال ۲ به ابعاد بالاتر) پاسخ مسئله‌ی قدم زدن تصادفی دو بُعدی (حرکت شخص مست روی صفحه) نیز از مرتبه‌ی $\sqrt{N}$ است و $\mathrm{E}[S_N^{۲}] = \sqrt{N}$.
برای بُعد d داریم:

$$\mathrm{E}[|S_N|] = \sqrt{\frac{۲N}{d}\,\frac{\Gamma(\frac{d+۱}{۲})}{\Gamma(\frac{d}{۲})}}$$

[۳] Root mean square distance (RMSD)

که در آن تابع Γ (گاما) بدین شکل تعریف می‌شود:

$$\Gamma(t) = \int_{۰}^{\infty} x^{t-۱} e^{-x} dx$$

که به نوعی تعمیمی از تابع فاکتوریل است، چنانکه:

$$\Gamma(n) = (n-۱)!$$

۴ محاسبه‌ی توزیع مشترک توابعی از دو متغیر تصادفی

در این بخش ابتدا یک مثال مطرح می‌کنیم و با روش مستقیم آنرا حل می‌کنیم. سپس یک قضیه کاربردی مطرح می‌کنیم که نشان می‌دهد چگونه مسائلی از این دست را می‌توان راحت‌تر حل نمود.

۱.۴ یک مثال

مثال ۴ *فرض کنید X و Y دو متغیر تصادفی با پارامتر $\lambda = ۱$ بوده و توابع چگالی احتمال آنها به قرار زیر باشد:*

$$f_X(x) = \begin{cases} e^{-x} & x \geq ۰ \\ ۰ & x < ۰ \end{cases}$$

$$f_Y(y) = \begin{cases} e^{-y} & y \geq ۰ \\ ۰ & y < ۰ \end{cases}$$

دو متغیر تصادفی $U = X + Y$ و $V = \frac{X}{Y}$ را تعریف می‌کنیم. مطلوب، محاسبه‌ی توزیع مشترک U و V است.

پاسخ *می‌توان نوشت:*

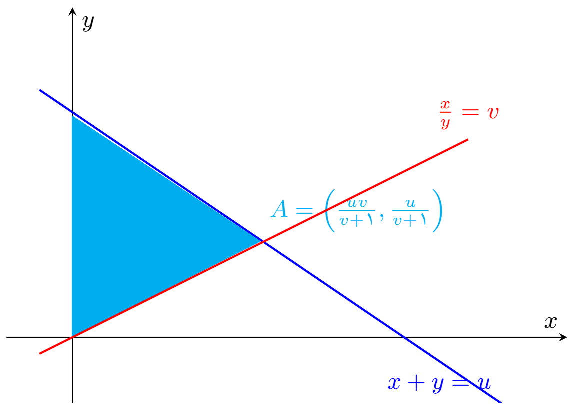
$$\begin{aligned} f_{U,V}(u,v) &= \Pr\{U \leq u,\, V \leq v\} \\ &= \Pr\{X+Y \leq u,\, \frac{X}{Y} \leq v\} \\ &= \iint\limits_{x+y\leq u,\, \frac{x}{y}\leq v} f_{X,Y}(x,y)\,dx\,dy \end{aligned} \tag{۵}$$

که با توجه به استقلال X و Y داریم:

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) = \begin{cases} e^{-(x+y)} & x,y \geq ۰ \\ ۰ & otherwise \end{cases}$$

اکنون دو حالت در نظر می‌گیریم:

- ***الف) حداقل یکی از u یا v مثبت است.** از آنجا که متغیرهای تصادفی X و Y هر دو مقدار نامنفی دارند، پس $f_{U,V}(u,v) = ۰$.*
- ***ب) u و v هر دو نامنفی اند.** در این حالت، محدوده‌ی انتگرال دوگانه رابطه (۵)، مثلث رنگی در شکل ۱ است.*



شکل ۱: محدوده‌ی انتگرال دوگانه رابطه‌ی (۵) برای $u, v > 0$

پس داریم:

$$
\begin{aligned}
f_{U,V}(u,v) &= \iint\limits_{x+y\leq u,\ \frac{x}{y}\leq v} f_{X,Y}(x,y)\,dx\,dy \\
&= \int_{x=0}^{\frac{uv}{v+1}} \int_{y=\frac{x}{v}}^{u-x} e^{-(x+y)}\,dx\,dy \\
&= e^{-u}\frac{u}{(v+1)^2}
\end{aligned}
$$

در نتیجه تابع توزیع مشترک U و V به دست می‌آید:

$$
f_{U,V}(u,v) = \begin{cases} e^{-u}\frac{u}{(v+1)^2} & u, v \geq 0 \\ 0 & oth. \end{cases}
$$

سؤال ۲ *آیا U و V مستقل‌اند؟*

۲.۴ یک قضیه کاربردی

روش ساده‌تری برای حل مسائل مشابه مثال (۴) وجود دارد که آن را به عنوان قضیه‌ای بیان می‌کنیم.

قضیه ۱ *فرض کنید X و Y دو متغیر تصادفی پیوسته با تابع چگالی احتمال توأم $f_{X,Y}(x,y)$ باشند. توابع دو متغیره‌ی حقیقی-مقدار h_1 و h_2 و متغیرهای تصادفی $U = h_1(X,Y)$ و $V = h_2(X,Y)$ را در نظر بگیرید. فرض کنید:*

(الف) *تبدیل $(x,y) \mapsto (u,v) = (h_1(x,y), h_2(x,y))$، تبدیلی یک‌به‌یک از مجموعه‌ی R در صفحه‌ی xy به مجموعه‌ی Q در صفحه‌ی uv باشد. به بیان دیگر، برای هر $(u,v) \in Q$، دستگاه*

$$
\begin{cases} h_1(x,y) = u \\ h_2(x,y) = v \end{cases}
$$

جواب یکتایی به شکل $x = w_1(u,v)$ و $y = w_2(u,v)$ داشته باشد؛ و نیز:

(ب) توابع $w_۱$ و $w_۲$، دارای مشتقات پاره‌ای پیوسته بوده و ماتریس ژاکوبی تبدیل $x = w_۱(u,v)$ و $y = w_۲(u,v)$ در تمام نقاط $(u,v) \in Q$ ناصفر باشد، یعنی

$$J = \det \begin{bmatrix} \frac{\partial w_۱}{\partial u} & \frac{\partial w_۱}{\partial v} \\ \frac{\partial w_۲}{\partial u} & \frac{\partial w_۲}{\partial v} \end{bmatrix} = \frac{\partial w_۱}{\partial u}\frac{\partial w_۲}{\partial v} - \frac{\partial w_۱}{\partial v}\frac{\partial w_۲}{\partial u} \neq ۰$$

آنگاه، متغیرهای تصادفی U و V توأماً پیوسته اند و تابع چگالی احتمال توأم آنها بدین شکل خواهد بود:

$$f_{U,V}(u,v) = \begin{cases} f_{X,Y}\Big(w_۱(u,v),\ w_۲(u,v)\Big)|J| & (u,v) \in Q \\ ۰ & elsewhere \end{cases}$$

۳.۴ حل مثال با استفاده از قضیه کاربردی

حال با استفاده از این قضیه، به حل مثال (۴) می‌پردازیم.
جواب دستگاه، همان مختصات نقطه‌ی A است؛ یعنی

$$x = w_۱(u,v) = \frac{uv}{v+۱}$$
$$y = w_۲(u,v) = \frac{u}{v+۱}$$

پس می‌توان J را محاسبه کرد:

$$J = \det \begin{bmatrix} \frac{v}{۱+v} & \frac{u}{(۱+v)^۲} \\ \frac{۱}{۱+v} & \frac{-u}{(۱+v)^۲} \end{bmatrix} = \frac{-u}{(۱+v)^۲} \Rightarrow |J| = \frac{u}{(۱+v)^۲}$$

دقت کنید که $\frac{uv}{v+۱}, \frac{u}{v+۱} \geq ۰$ اگر و فقط اگر $u, v \geq ۰$. پس می‌توان نوشت:

$$f_{X,Y}\Big(\frac{uv}{v+۱}, \frac{u}{v+۱}\Big) = \begin{cases} e^{-\left(\frac{uv}{v+۱} + \frac{u}{v+۱}\right)} = e^{-u} & u,v \geq ۰ \\ ۰ & oth. \end{cases}$$

پس خواهیم داشت:

$$f_{U,V}(u,v) = \begin{cases} e^{-u}\frac{u}{(v+۱)^۲} & u,v \geq ۰ \\ ۰ & oth. \end{cases}$$