پیوست پ شمارش و احتمال

در این پیوست، تئوری اولیه ترکیبات و احتمال را بررسی می‌کنیم. اگر با این موضوعات آشنایی دارید، می‌توانید به بخش‌های بعدی بپردازید. بعضی از فصل‌ها به احتمالات نیاز ندارند.

بخش پ – ۱، نتایج اولیه تئوری شمارش را مرور می‌کند، از جمله فرمول‌های استاندارد برای شمارش جایگشت‌ها و ترکیبات. اصول احتمال و حقایق اصلی مربوط به توزیع احتمال، در بخش پ – ۲ آمده است. متغیرهای تصادفی در بخش پ – ۳ معرفی شدند، در همین بخش، خواص انتظار (امید) و واریانس معرفی شدند. بخش پ – ۴ توزیع دوجمله‌ای و هندسی را بحث می‌کند که ناشی از مطالعه آزمایش‌های برنولی است. مطالعه توزیع دوجمله‌ای در بخش پ – ۵ ادامه می‌یابد، که بحث پیشرفته‌ای از توزیع را ارائه می‌کند.

پ–۱ شمارش

تئوری شمارش سعی می‌کند به پرسش "چه مقدار" پاسخ دهد، بدون این‌که واقعاً شمارش کند. برای مثال، ممکن‌است بپرسیم "چند عدد n بیتی مختلف وجود دارند؟" ، یا " چند ترتیب از n عنصر متفاوت وجود دارند؟". در این بخش، عناصر تئوری شمارش را مرور می‌کنیم. چون بعضی از مفاهیم به مجموعه‌ها ارتباط دارد، بهتر است بخش ب – ۱ را مطالعه کنید.

قوانین جمع و ضرب

مجموعه‌ای از اقلام که می‌خواهیم شمارش کنیم، گاهی می‌توانند به صورت اجتماع مجموعه‌های جدا از هم یا حاصل‌ضرب دکارتی مجموعه‌ها بیان شوند. **قانون جمع** می‌گوید که تعداد راه‌ها برای انتخاب عنصری از یکی از دو مجموعه جدا ازهم، برابر با مجموع اعداد اصلی[1] دو مجموعه است. یعنی، اگر A و B دو مجموعه متناهی باشند که هیچ عضو مشترکی نداشته باشند، آنگاه $|A\cup B|=|A|+|B|$، که از معادله (ب – ۳) به دست می‌آید. برای مثال، هر مکان در شماره اتومبیل، یک حرف یا رقم است. بنابراین تعداد حالت‌های ممکن برای هر مکان، برابر است با $26 + 10 = 36$ ، زیرا اگر حرف باشد ۲۶ انتخاب و اگر رقم باشد، ۱۰ انتخاب وجود دارد.

قانون ضرب می‌گوید که تعداد راه‌ها برای انتخاب زوج مرتب، برابر با تعداد راه‌ها برای انتخاب اولین عنصر ضرب در تعداد راه‌ها برای انتخاب عنصر دوم است. یعنی، اگر A و B دو مجموعه متناهی باشند، آنگاه

1. cardinality

$|A\times B|=|A|.|B|$، که معادله (ب – ۴) است. برای مثال اگر فروشگاه بستنی، ۲۸ طعم بستنی و ۴ مخلفات روی بستنی ارائه دهد، تعداد بستنی‌های میوه‌ای با یک پیمانه بستنی و یک مخلفات، برابر با $28\times 4=112$ است.

رشته‌ها

رشته‌ای روی مجموعه متناهی S ، دنباله‌ای از عناصر S است. برای مثال، ۸ رشته دودویی به طول ۳ وجود دارد:

$$000, 001, 010, 011, 100, 101, 110, 111$$

گاهی رشته به طول k را **رشته – k** می‌نامیم. **زیررشته** s' از رشته s ، دنباله مرتبی از عناصر متوالی s است. **زیررشته – k** از یک رشته، زیررشته‌ای به طول k است. برای مثال، 010 ، زیررشته – ۳ از 01101001 است (زیررشته – ۳ که در مکان ۴ شروع می‌شود)، اما 111 زیررشته‌ای از 01101001 نیست.

رشته – k روی مجموعه S به عنوان عنصری از ضرب دکارتی S^k از k تایی‌ها محسوب می‌شود؛ بنابراین، تعداد $|S|^k$ رشته به طول k وجود دارد. برای مثال، تعداد رشته – k دودویی، برابر با 2^k است. از نظر شهودی، برای ساخت رشته – k روی مجموعه – n ، n روش برای انتخاب اولین عنصر وجود دارد؛ برای هر یک از این انتخاب‌ها، n روش برای انتخاب عنصر دوم وجود دارد، و در نتیجه k روش انتخاب وجود دارد. این ساختار، منجر به حاصلضرب k تایی $n.n...n=n^k$ به عنوان تعداد رشته – k می‌شود.

جایگشت‌ها

جایگشت مجموعه متناهی S ، دنباله مرتبی از تمام عناصر S است، به طوری که هر عنصر فقط یک بار ظاهر می‌شود. برای مثال، اگر $S=\{a, b, c\}$ ، آنگاه ۶ جایگشت برای S وجود دارد:

$$abc, acb, bac, bca, cab, cba$$

تعداد $n!$ جایگشت از مجموعه n عنصری وجود دارد، زیرا اولین عنصر دنباله می‌تواند به n روش، دومی به $n-1$ روش، و سومی به $n-2$ روش انتخاب شود و غیره.

جایگشت – k مربوط به S ، دنباله مرتبی از k عنصر S است، که هیچ عنصری بیش از یک بار در دنباله ظاهر نمی‌شود (بنابراین، جایگشت معمولی، فقط جایگشت – n از مجموعه – n است). ۱۲ جایگشت – ۲ از مجموعه $\{a, b, c\}$ عبارتنداز:

$$ab, ac, ad, ba, bc, bd, ca, cb, cd, da, db, dc$$

تعداد جایگشت – k از مجموعه – n عبارت‌است‌از:

$$n(n-1)(n-2)\cdots(n-k+1)=\frac{n!}{(n-k)!} \qquad \text{(پ–۱)}$$

زیرا n روش برای انتخاب عنصر اول و $n-1$ روش برای انتخاب عنصر دوم وجود دارد، و در نتیجه تا k عنصر انتخاب می‌شوند، به طوری که آخری از $n-k+1$ عنصر انتخاب می‌گردد.

ترکیبات

ترکیب – k از مجموعه – n به نام S ، زیرمجموعه – k از S است. برای مثال، شش ترکیب – ۲ از مجموعه – ۴ $\{a, b, c, d\}$ وجود دارد:

$$ab, ac, ad, bc, bd, cd$$

(در این‌جا مجموعه – ۲ $\{a, b\}$ را برای اختصار به صورت ab نشان می‌دهیم و غیره). می‌توان ترکیب – k از مجموعه – n را با انتخاب k عنصر مختلف از مجموعه – n ، ایجاد کرد.

تعداد ترکیب – k از مجموعه – n می‌تواند برحسب تعداد جایگشت – k از مجموعه – n بیان شود. برای هر ترکیب – k ، دقیقاً $k!$ جایگشت از عناصرش وجود دارد، که هر کدام یک جایگشت – k مجزا از مجموعه – n است. بنابراین، تعداد جایگشت‌های – k از مجموعه – n ، برابر با تعداد جایگشت‌های – k تقسیم بر $k!$ است، بنا به معادله (پ–۱)، این کمیت برابر است با:

$$\frac{n!}{k!\,(n-k)!} \tag{پ–۲}$$

برای $k = 0$ ، این فرمول به ما می‌گوید که تعداد راه‌های انتخاب صفر عنصر از مجموعه – n ، یک است (نه صفر)، زیرا $0! = 1$.

ضرایب دوجمله‌ای

از نماد $\binom{n}{k}$ (بخوانید n ، k را انتخاب می‌کند) برای نشان‌دادن تعداد ترکیب‌های – k از مجموعه – n استفاده می‌کنیم. بنا به معادله (پ–۲) داریم:

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

این فرمول، برحسب k و $n-k$ متقارن است:

$$\binom{n}{k} = \binom{n}{n-k} \tag{پ–۳}$$

این اعداد را **ضرایب دوجمله‌ای** نیز می‌نامند، زیرا در **بسط دوجمله‌ای** ظاهر می‌شوند:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \tag{پ–۴}$$

حالت خاصی از بسط دوجمله‌ای، وقتی به وجود می‌آید که $x = y = 1$:

$$2^n = \sum_{k=0}^{n} \binom{n}{k}$$

این فرمول متناظر با شمارش 2^n رشته – n دودویی، توسط تعداد یک‌هایی که شامل آن هستند؛ تعداد $\binom{n}{k}$ رشته – n دودویی وجود دارند که دقیقاً شامل k عدد یک هستند، زیرا $\binom{n}{k}$ روش برای انتخاب k مکان از n مکان وجود دارد که می‌توان یک را در آن‌جا قرار داد.

کران‌های دوجمله‌ای

گاهی لازم است اندازه ضریب دوجمله‌ای محدود شود. برای $1 \le k \le n$، کران زیر را داریم:

$$\begin{aligned}\binom{n}{k} &= \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1} \\ &= \left(\frac{n}{k}\right)\left(\frac{n-1}{k-1}\right)\cdots\left(\frac{n-k+1}{1}\right) \\ &\ge \left(\frac{n}{k}\right)^k\end{aligned}$$

با استفاده از امتیاز نامعادله $k! \geq (k/e)^k$ که از تقریب stirling (۳-۱۷) مشتق می‌شود، کران‌های بالای زیر را به دست می‌آوریم:

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1} \leq \frac{n^k}{k!} \leq \left(\frac{en}{k}\right)^k \quad \text{(پ-۵)}$$

برای تمام $0 \leq k \leq n$ ، برای اثبات کران زیر می‌توان از استقرا استفاده کرد (تمرین پ-۱-۱۲ را ببینید):

$$\binom{n}{k} \leq \frac{n^n}{k^k(n-k)^{n-k}} \quad \text{(پ-۶)}$$

که برای سهولت، فرض می‌کنیم $0^0 = 1$. برای $k = \lambda n$، که $0 \leq \lambda \leq 1$، این کران می‌تواند به صورت زیر بازنویسی شود:

$$\binom{n}{\lambda n} \leq \frac{n^n}{(\lambda n)^{\lambda n}((1-\lambda)n)^{(1-\lambda)n}} = \left(\left(\frac{1}{\lambda}\right)^{\lambda}\left(\frac{1}{1-\lambda}\right)^{1-\lambda}\right)^n = 2^{n H(\lambda)}$$

که در آن، تابع زیر، تابع آنتروپی[1] (دودویی) است، و برای سهولت فرض می‌کنیم $0 \lg 0 = 0$ ، به طوری که $H(0) = H(1) = 0$:

$$H(\lambda) = -\lambda \lg \lambda - (1-\lambda)\lg(1-\lambda) \quad \text{(پ-۷)}$$

تمرین‌های بخش پ – ۱

تمرین پ-۱-۱: یک رشته - n ، چند زیررشته k دارد؟ (رشته‌های – k یکسان در مکان‌های مختلف را، متفاوت درنظربگیرید). یک رشته – n در مجموع، چند زیررشته دارد؟

تمرین پ-۱-۲: تابع بولی با n ورودی و n خروجی، تابعی از $\{TRUE, FALSE\}^n$ به $\{TREU, FALSE\}^m$ است. چند تابع با یک ورودی و یک خروجی وجود دارند؟ چند تابع با n ورودی و m خروجی وجود دارند؟

تمرین پ-۱-۳: n پروفسور به چند روش می‌توانند روی یک میز کنفرانس دایره‌ای بنشینند؟ دو صندلی در صورتی یکسان هستند که یکی بتواند بچرخد تا به دومی تبدیل شود.

تمرین پ-۱-۴: برای انتخاب سه عدد مختلف از مجموعه $\{1, 2, \ldots, 100\}$ ، به طوری که مجموع آن‌ها زوج باشد، چند روش وجود دارد؟

تمرین پ-۱-۵: برای $0 < k \leq n$، تساوی زیر را اثبات کنید:

$$\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1} \quad \text{(پ-۸)}$$

1. entropy function

تمرین پ-۱-۶: برای $0 \le k < n$، تساوی زیر را ثابت کنید:

$$\binom{n}{k} = \frac{n}{n-k}\binom{n-1}{k}$$

تمرین پ-۱-۷: برای انتخاب k شیء از n شیء، می‌توانید یکی از اشیا را متمایز کنید و مشخص کنید که آیا این شیء انتخاب می‌شود یا خیر؟ با استفاده از این روش، تساوی زیر را اثبات کنید:

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$$

تمرین پ-۱-۸: با استفاده از نتیجه تمرین پ-۱-۷، جدولی برای $n = 0, 1, \ldots, 6$ و $0 \le k \le n$ از ضرایب دوجمله‌ای $\binom{n}{k}$ با $\binom{0}{0}$ در بالا، $\binom{1}{0}$ و $\binom{1}{1}$ در خط بعدی، و غیره بسازید. چنین جدولی از ضرایب دوجمله‌ای، **مثلث پاسکال** نام دارد.

تمرین پ-۱-۹: ثابت کنید:

$$\sum_{i=1}^{n} i = \binom{n+1}{2}$$

تمرین پ-۱-۱۰: نشان دهید که برای $n \ge 0$ و $0 \le k \le n$، مقدار ماکزیمم $\binom{n}{k}$ وقتی به دست می‌آید که $k = \lfloor n/2 \rfloor$ یا $k = \lceil n/2 \rceil$.

★ **تمرین پ-۱-۱۱:** ثابت کنید که برای هر $n \ge 0$، $j \ge 0$، $k \ge 0$ و $j + k \le n$، داریم:

$$\binom{n}{j+k} \le \binom{n}{j}\binom{n-j}{k} \tag{پ-۹}$$

یک اثبات جبری و یک بحث بر اساس روش انتخاب $j + k$ قلم از n قلم ارائه دهید. مثالی ارائه دهید که در آن، تساوی اتفاق نمی‌افتد.

★ **تمرین پ-۱-۱۲:** با استفاده از استقرا روی $k \le n/2$، نامساوی (پ-۶) را اثبات کنید، و با استفاده از معادله (پ-۳)، آن را برای تمام $k \le n$ بسط دهید.

★ **تمرین پ-۱-۱۳:** با استفاده از تقریب stirling ثابت کنید:

$$\binom{2n}{n} = \frac{2^{2n}}{\sqrt{\pi n}}(1 + O(1/n)) \tag{پ-۱۰}$$

★ **تمرین پ-۱-۱۴:** با دیفرانسیل‌گیری تابع آنتروپی $H(\lambda)$، نشان دهید که مقدار ماکزیمم آن در $\lambda = 1/2$ به دست آید. $H(1/2)$ چیست؟

★ **تمرین پ-۱-۱۵:** برای هر مقدار صحیح $n \ge 0$ نشان دهید که:

$$\sum_{k=0}^{n} \binom{n}{k} k = n2^{n-1} \tag{پ-۱۱}$$

پ-۲ احتمال

احتمال، ابزار اساسی برای طراحی و تحلیل الگوریتم‌های احتمالی و تصادفی است. این بخش، تئوری احتمال پایه را بحث می‌کند.

احتمال را برحسب **فضای نمونه**[1] S تعریف می‌کنیم که مجموعه‌ای است که عناصر آن **پیشامدهای اولیه**[2] (یا رویدادهای اولیه) نامیده می‌شوند. هر پیشامد اولیه را می‌توان نتیجه‌ی ممکن از یک آزمایش دانست. برای آزمایش پرتاب سکه‌های متمایز، فضای نمونه را مجموعه‌ای از رشته‌های – ۲ روی {H, T} درنظرمی‌گیریم:

$$S = \{\text{HH, HT, TH, TT}\}$$

پیشامد، زیرمجموعه‌ای از فضای نمونه S است. برای مثال، در آزمایش پرتاب دو سکه، پیشامد به دست آوردن یک شیر و یک خط، برابر با {HT, TH} است. پیشامد S را **پیشامد مطمئن**[3] و پیشامد $\varnothing$ را **پیشامد تهی** می‌نامیم. اگر $A \cap B = \varnothing$ می‌گوییم پیشامدهای A و B **انحصار متقابل‌اند**[4]. گاهی با پیشامد اولیه $s \in S$ به عنوان پیشامد {s} رفتار می‌کنیم. بنا به تعریف، تمام پیشامدهای اولیه، انحصار متقابل‌اند.

اصول احتمال

توزیع احتمال[5] $\Pr\{\}$ روی فضای نمونه S ، نگاشتی از پیشامدهای S به اعداد حقیقی است، به طوری که **اصل احتمال** زیر برآورده شود:

1. برای هر پیشامد A داریم $\Pr\{A\} \geq 0$.
2. $\Pr\{S\} = 1$.
3. برای هر دو پیشامد انحصار متقابل A و B ، داریم $\Pr\{A \cup B\} = \Pr\{A\} + \Pr\{B\}$. به طور کلی‌تر، برای هر دنباله (متناهی یا نامتناهی شمارش‌پذیر) از پیشامدهای ...,A_1, A_2 که دو به دو انحصار متقابل‌اند، داریم:

$$\Pr\left\{\bigcup_i A_i\right\} = \sum_i \Pr\{A_i\}$$

$\Pr\{A\}$ را **احتمال** پیشامد A می‌گوییم. در این‌جا توجه دارید که اصل ۲، یک نیازمندی نرمال‌سازی[6] است: نکته خاصی راجع‌به انتخاب ۱ به عنوان احتمال پیشامد مطمئن وجود ندارد، مگر این‌که طبیعی و آسان است.

از این اصول و تئوری مجموعه پایه، نتایج زیر به دست می‌آید (بخش ب – ۱ را ببینید). پیشامد تهی دارای احتمال $\Pr\{\varnothing\} = 0$ است. اگر $A \subseteq B$ ، آنگاه $\Pr\{A\} \leq \Pr\{B\}$. با استفاده از $\bar{A}$ به عنوان پیشامد S-A (**مکمل A**)، داریم $\Pr\{\bar{A}\} = 1 - \Pr\{A\}$. برای هر دو پیشامد A و B ، داریم:

$$\Pr\{A \cup B\} = \Pr\{A\} + \Pr\{B\} - \Pr\{A \cap B\} \qquad \text{(پ-۱۲)}$$

$$\leq \Pr\{A\} + \Pr\{B\} \qquad \text{(پ-۱۳)}$$

در مثال پرتاب سکه، فرض کنید هر یک از چهار پیشامد اولیه دارای احتمال ¼ باشند. آنگاه احتمال به دست آوردن حداقل یک شیر به صورت زیر است:

$$\begin{aligned}\Pr\{\text{HH, HT, TH}\} &= \Pr\{\text{HH}\} + \Pr\{\text{HT}\} + \Pr\{\text{TH}\} \\ &= 3/4\end{aligned}$$

از طرف دیگر، چون احتمال به دست آوردن اکیداً کمتر از یک شیر، برابر با $\Pr\{TT\} = 1/4$ است، احتمال به دست آوردن حداقل یک شیر، برابر است با $1 - 1/4 = 3/4$.

1. sample space 2. elementary events 3. certain event 4. mutually exclusive
5. probability distribution 6. normalization

توزیع‌های احتمال گسسته

توزیع احتمال، در صورتی **گسسته** است که روی فضای نمونه متناهی یا نامتناهی شمارش‌پذیر تعریف شده باشد. فرض کنید S فضای نمونه باشد. آنگاه برای هر پیشامد A ، داریم:

$$\Pr\{A\} = \sum_{s \in A} \Pr\{s\}$$

زیرا پیشامدهای ابتدایی، به خصوص آن‌هایی که در A هستند، انحصار متقابل‌اند. اگر S متناهی و هر پیشامد اولیه $s \in S$ دارای احتمال زیر باشد، آنگاه **توزیع احتمال یکنواخت** روی S را خواهیم داشت:

$$\Pr\{s\} = 1/|S|$$

در چنین موردی، آزمایش معمولاً به صورت "انتخاب تصادفی عنصری از S " توصیف می‌شود.

به عنوان مثال، فرآیند پرتاب **سکه سالم** را درنظربگیرید، که برای آن، احتمال به دست آوردن شیر برابر با احتمال به دست آوردن خط، یعنی 1/2 است. اگر سکه را n بار پرتاب کنیم، توزیع احتمال یکنواخت تعریف‌شده روی فضای نمونه $S = \{H, T\}^n$ را به دست می‌آوریم که مجموعه‌ای به اندازه 2^n است. هر پیشامد اولیه در S می‌تواند به صورت رشته‌ای به طول n روی {H, T} نمایش داده شود، و هر کدام با احتمال $1/2^n$ رخ دهد. پیشامد زیر، زیرمجموعه‌ای از S به اندازه $|A| = \binom{n}{k}$ است، زیرا $\binom{n}{k}$ رشته به طول n روی {H, T} وجود دارند که دقیقاً k عدد H دارند:

A = {دقیقاً k شیر و دقیقاً n − k خط رخ می‌دهد}

بنابراین، احتمال پیشامد A برابر است با $\Pr\{A\} = \binom{n}{k}/2^n$.

توزیع احتمال یکنواخت پیوسته

توزیع احتمال یکنواخت پیوسته، مثالی از توزیع احتمال است که در آن، تمام زیرمجموعه‌های فضای حالت به عنوان پیشامد در نظر گرفته نمی‌شود. توزیع احتمال یکنواخت پیوسته، روی فاصله‌ی بسته‌ی [a, b] از مقادیر حقیقی تعریف می‌شود که $a < b$. از نظر شهودی، می‌خواهیم احتمال هر نقطه در فاصله [a, b] یکسان باشد. اما، تعداد شمارش‌ناپذیری از نقاط وجود دارد، که اگر به هر کدام احتمال مثبت یکسانی نسبت دهیم، نمی‌توانیم همزمان اصول ۲ و ۳ را برآورده کنیم. به همین دلیل، فقط می‌خواهیم به بعضی از زیرمجموعه‌های S احتمالی را نسبت بدهیم به طوری که اصول احتمال برای این پیشامدها برآورده شوند.

برای هر فاصله‌ی بسته‌ی [c, d] که $a \le c \le d \le b$، توزیع احتمال یکنواخت پیوسته، احتمال پیشامد [c, d] را به صورت زیر تعریف می‌کند:

$$\Pr\{[c, d]\} = \frac{d - c}{b - a}$$

توجه کنید که برای هر نقطه $x = [x, x]$ ، احتمال x برابر با صفر است. اگر نقاط انتهایی فاصله [c, d] را حذف کنیم، فاصله بازِ (c, d) را به دست می‌آوریم. چون $[c,c] \cup (c,d) \cup [d,d]$ ، از اصل ۳ خواهیم داشت: $\Pr\{[c,d]\} = \Pr\{(c,d)\}$. به طور کلی، مجموعه‌ای از پیشامدها برای توزیع احتمال یکنواخت پیوسته، هر زیرمجموعه‌ای از فضای نمونه [a, b] است که می‌تواند با اجتماع متناهی یا شمارش‌پذیر فاصله‌های باز و بسته به دست آید.

احتمال شرطی و استقلال

گاهی، از قبل، دانش جزیی درباره نتیجه‌ی آزمایشی داریم. برای مثال، فرض کنید دوستی سکه‌های سالم را پرتاب کرد و به شما گفت که حداقل یکی از سکه‌ها شیر را نشان داد. احتمال این‌که هر دو شیر باشند چقدر است؟ اطلاعات داده‌شده، احتمال دو خط را حذف کرده است. سه پیشامد اولیه‌ی باقیمانده، احتمال یکسانی دارند، لذا نتیجه می‌گیریم که هر کدام با احتمال 1/3 رخ می‌دهند. چون فقط یکی از این پیشامدهای اولیه دو شیر را نشان می‌دهند، پاسخ پرسش ما 1/3 است.

احتمال شرطی، فرضیه دانش جزیی قبلی درباره نتیجه‌ی آزمایش را فرمول‌بندی می‌کند. احتمال شرطی این‌که پیشامد A موجب پیشامد B شود، وقتی $\Pr\{B\} \neq 0$ باشد، به صورت زیر است:

$$\Pr\{A \mid B\} = \frac{\Pr\{A \cap B\}}{\Pr\{B\}} \qquad \text{(پ-۱۴)}$$

($\Pr\{A \mid B\}$ را "احتمال این‌که پیشامد A موجب پیشامد B شود، بخوانید"). از نظر شهودی، چون می‌دانیم که پیشامد B رخ می‌دهد، پیشامدی که A نیز رخ دهد برابر است با $A \cap B$، یعنی، $A \cap B$ مجموعه‌ای از نتایج است که در آن، هم A و هم B رخ می‌دهند. چون نتیجه، یکی از پیشامدهای اولیه در B است، احتمالات تمام پیشامدهای اولیه در B را توسط تقسیم آن‌ها بر $\Pr\{B\}$ نرمال‌سازی می‌کنیم، به طوری که مجموع آن‌ها یک شود. بنابراین، احتمال این‌که شرط A موجب B شود، برابر با نسبت احتمال پیشامد $A \cap B$ به احتمال پیشامد B است. در مثال فوق، A پیشامدی است که هر دو سکه شیر هستند، و B پیشامدی است که حداقل یکی از سکه‌ها شیر است. بنابراین، $\Pr\{A \mid B\} = (1/4)/(3/4) = 1/3$.

دو پیشامد در صورتی **مستقل** هستند که:

$$\Pr\{A \cap B\} = \Pr\{A\}\Pr\{B\} \qquad \text{(پ-۱۵)}$$

که اگر $\Pr\{B\} \neq 0$ معادل زیر است:

$$\Pr\{A \mid B\} = \Pr\{A\}$$

برای مثال، فرض کنید دو سکه سالم پرتاب شدند و نتایج مستقل‌اند. آنگاه احتمال دو شیر برابر است با $(1/2)(1/2) = 1/4$. اکنون فرض کنید یک پیشامد این‌است‌که سکه اول شیر می‌آید و پیشامد دیگر این‌است‌که پرتاب سکه‌ها نتایج متفاوتی دارند. هر یک از این دو پیشامد، با احتمال 1/2 رخ می‌دهند، و احتمال این‌که هر دو پیشامد رخ دهند، یک است. بنابراین، بر اساس تعریف استقلال – پیشامدها مستقل‌اند – گرچه ممکن‌است تصور کنید که هر دو پیشامد به سکه اول بستگی دارند. سرانجام، فرض کنید که سکه‌ها به هم مرتبط هستند، به طوری که هر دو شیر یا هر دو خط می‌آیند و احتمال هر دو نیز یکسان است. پس، احتمال این‌که هر سکه شیر بیاید 1/2 است، ولی احتمال این‌که هر دو شیر بیایند برابر با $(1/2)(1/2) \neq 1/2$ است. در نتیجه، پیشامدی که شیر می‌آید و پیشامدی که دیگری شیر بیاید، مستقل نیستند.

کلکسیون $A_1, A_2, \ldots, A_n$ از پیشامدها را **دو به دو مستقل** می‌گویند اگر برای تمام $1 \leq i < j \leq n$ داشته باشیم:

$$\Pr\{A_i \cap A_j\} = \Pr\{A_i\}\Pr\{A_j\}$$

می‌گوییم رویدادهای این کلکسیون (متقابلاً) **مستقل‌اند** اگر هر زیرمجموعه k ، $A_{i_1}, A_{i_2}, ..., A_{i_k}$ از این کلکسیون، که $2 \le k \le n$ و $1 \le i_1 < i_2 < ... < i_k \le n$، در رابطه زیر صدق کند:

$$\Pr\{A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}\} = \Pr\{A_{i_1}\}\Pr\{A_{i_2}\}\cdots\Pr\{A_{i_k}\}$$

برای مثال، فرض کنید سکه‌های سالم را پرتاب می‌کنیم. فرض کنید A_1 پیشامدی باشد که سکه اول شیر می‌آید، و A_2 پیشامدی باشد که سکه دوم شیر می‌آید، و A_3 پیشامدی باشد که دو سکه متفاوت‌اند. داریم:

$$\begin{aligned}
\Pr\{A_1\} &= 1/2\,,\\
\Pr\{A_2\} &= 1/2\,,\\
\Pr\{A_3\} &= 1/2\,,\\
\Pr\{A_1 \cap A_2\} &= 1/4\,,\\
\Pr\{A_1 \cap A_3\} &= 1/4\,,\\
\Pr\{A_2 \cap A_3\} &= 1/4\,,\\
\Pr\{A_1 \cap A_2 \cap A_3\} &= 0
\end{aligned}$$

چون برای $1 \le i < j \le 3$ داریم $\Pr\{A_i \cap A_j\} = \Pr\{A_i\}\Pr\{A_j\} = 1/4$ ، پیشامدهای A_1 ، A_2 و A_3 دو به دو مستقل‌اند. پیشامدها متقابلاً مستقل نیستند، زیرا $\Pr\{A_1 \cap A_2 \cap A_3\} = 0$ و $\Pr\{A_1\}\Pr\{A_2\}\Pr\{A_3\} = 1/8 \ne 0$.

قضیه بیز[1]

از تعریف احتمال شرطی (پ-۱۴) و قانون جابه‌جایی $A \cap B = B \cap A$ ، نتیجه می‌شود که برای دو پیشامد A و B ، هر کدام با احتمال غیرصفر، داریم:

$$\begin{aligned}
\Pr\{A \cap B\} &= \Pr\{B\}\Pr\{A \mid B\}\\
&= \Pr\{A\}\Pr\{B \mid A\}
\end{aligned} \tag{پ-۱۶}$$

با حل برای $\Pr\{A \mid B\}$ ، داریم:

$$\Pr\{A \mid B\} = \frac{\Pr\{A\}\Pr\{B \mid A\}}{\Pr\{B\}} \tag{پ-۱۷}$$

که **قضیه بیز** نام دارد. مخرج $\Pr\{B\}$ ثابت نرمال‌کننده‌ای است که می‌تواند به صورت زیر بیان شود. چون $B = (B \cap A) \cup (B \cap \overline{A})$ و $B \cap A$ و $B \cap \overline{A}$ پیشامدهایی با انحصار متقابل‌اند، داریم:

$$\begin{aligned}
\Pr\{B\} &= \Pr\{B \cap A\} + \Pr\{B \cap \overline{A}\}\\
&= \Pr\{A\}\Pr\{B \mid A\} + \Pr\{\overline{A}\}\Pr\{B \mid \overline{A}\}
\end{aligned}$$

با جایگزینی در معادله (پ-۱۷)، شکل معادلی از قضیه بیز را به دست می‌آوریم:

$$\Pr\{A \mid B\} = \frac{\Pr\{A\}\Pr\{B \mid A\}}{\Pr\{A\}\Pr\{B \mid A\} + \Pr\{\overline{A}\}\Pr\{B \mid \overline{A}\}}$$

قضیه بیز می‌تواند محاسبه احتمالات شرطی را ساده نماید. برای مثال، فرض کنید یک سکه سالم و یک سکه ناقص داریم که همیشه شیر می‌آیند. آزمایشی انجام می‌دهیم که شامل سه پیشامد مستقل باشد: یکی از دو

1. Bayes's theorem

سکه به طور تصادفی انتخاب می‌شود، سکه یک بار پرتاب می‌شود، و سپس دوباره پرتاب می‌گردد. فرض کنید سکه‌ی انتخاب‌شده در هر دو بار شیر می‌آید. احتمال این‌که آن سکه ناقص باشد چند است؟

این مسئله را با استفاده از قضیه بیز حل می‌کنیم. فرض کنید A پیشامد انتخاب سکه ناقص باشد، و B پیشامدی باشد که سکه در هر دو بار شیر می‌آید. می‌خواهیم $\Pr\{A \mid B\}$ را تعیین کنیم. داریم $\Pr\{A\}=1/2$، $\Pr\{B \mid A\}=1$، $\Pr\{\overline{A}\}=1/2$ و $\Pr\{B \mid \overline{A}\}=1/4$. در نتیجه داریم:

$$\begin{aligned}\Pr\{A \mid B\} &= \frac{(1/2)\cdot 1}{(1/2)\cdot 1 + (1/2)\cdot(1/4)} \\ &= 4/5\end{aligned}$$

تمرین‌های بخش پ – ۲

تمرین پ-۲-۱: نامعادله بول را ثابت کنید: برای هر دنباله متناهی یا نامتناهی شمارش‌پذیر از پیشامدهای $A_1, A_2, \ldots$ داریم:

$$\Pr\{A_1 \cup A_2 \cup \cdots\} \le \Pr\{A_1\} + \Pr\{A_2\} + \cdots \qquad \text{(پ-۱۸)}$$

تمرین پ-۲-۲: پروفسور Rosencrants یک بار سکه سالمی را می‌اندازد. پروفسور Guildenstern سکه سالمی را دو بار می‌اندازد. احتمال این‌که پروفسور Rosencrants نسبت به پروفسور Guildenstern شیرهای بیشتری به دست آورد چقدر است؟

تمرین پ-۲-۳: دسته‌ای از ۱۰ کارت که از ۱ تا ۱۰ شماره‌گذاری شدند، مخلوط هستند. سه کارت از این دسته خارج می‌شوند، به طوری که هر بار یکی از این کارت‌ها خارج خواهند شد. احتمال این‌که سه کارت انتخاب‌شده به ترتیب مرتب باشند چیست؟

★**تمرین پ-۲-۴:** رویه‌ای را شرح دهید که دو مقدار صحیح a و b را به عنوان ورودی می‌گیرد به طوری که $0 < a < b$ و با استفاده از پرتاب‌های سکه سالم، شیرها را با احتمال a/b و خطها را با احتمال $(b-a)/b$ به عنوان خروجی تولید می‌کند. روی تعداد مورد انتظار پرتاب سکه‌ها، که $O(1)$ است، کرانی را تعیین کنید (*راهنمایی:* a/b را به صورت دودویی نشان دهید).

تمرین پ-۲-۵: ثابت کنید:

$$\Pr\{A \mid B\} + \Pr\{\overline{A} \mid B\} = 1$$

تمرین پ-۲-۶: ثابت کنید برای هر کلکسیون از رویدادهای $A_1, A_2, \ldots, A_n$ داریم:

$$\begin{aligned}\Pr\{A_1 \cap A_2 \cap \cdots \cap A_n\} = {}& \Pr\{A_1\} \cdot \Pr\{A_2 \mid A_1\} \cdot \Pr\{A_3 \mid A_1 \cap A_2\} \cdots \\ & \Pr\{A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1}\}\end{aligned}$$

★**تمرین پ-۲-۷:** نشان دهید که چگونه می‌توان مجموعه‌ای از n رویداد را ساخت که دو به دو مستقل باشند، اما هیچ زیرمجموعه $k > 2$ از آن‌ها، متقابلاً مستقل نباشند.

★**تمرین پ-۲-۸:** دو پیشامد A و B **مستقل شرطی‌اند،** اگر با توجه به C ، داشته باشیم:

$$\Pr\{A \cap B \mid C\} = \Pr\{A \mid C\} \cdot \Pr\{B \mid C\}$$

مثال ساده ولی غیربدیهی از دو پیشامد ارائه دهید که مستقل نیستند ولی با توجه به پیشامد سوم، مستقل شرطی‌اند.

★**تمرین پ-۲-۹:** فرض کنید در یک بازی که جایزه پشت یکی از سه پرده پنهان است، شما یکی از رقبا هستید. در صورتی جایزه را می‌برید که پرده درست را انتخاب کنید. پس از این‌که را انتخاب کردید ولی قبل از کناررفتن پرده، رقیب شما، پرده دیگری را کنار می‌زند. می‌دانید که یک صفحه‌ی خالی ظاهر می‌شود، و از شما خواسته می‌شود که می‌توانید پرده دیگری را انتخاب کنید. اگر پرده را عوض کنید، شانس شما چگونه تغییر می‌کند.

★ **تمرین پ-۲-۱۰:** رییس زندان، یکی از زندانی‌ها را به طور تصادفی انتخاب کرد تا آزاد کند. دو زندانی دیگر باقی می‌مانند. نگهبان می‌داند کدام‌یک آزاد می‌شود، اما نباید اطلاعات مربوط به زندانی را در اختیارش قرار دهد. زندانی‌ها را X ، Y و Z می‌نامیم. زندانی X به طور مخفی از نگهبان سوال می‌کند که کدام‌یک از Y یا Z باقی می‌مانند، با توجه به این‌که او می‌داند یکی از زندانی‌ها آزاد خواهد شد، نگهبان هیچ اطلاعاتی در اختیار او قرار نمی‌دهد. نگهبان به X و Y می‌گوید که Z آزاد خواهد شد، که به معنای این‌است‌که احتمال آزادی او اکنون 1/2 است. آیا او درست می‌گوید، یا هنوز شانس او 1/3 است؟ توضیح دهید.

پ-۳ متغیرهای تصادفی گسسته

متغیر تصادفی (گسسته) X ، تابعی از فضای نمونه نامتناهی شمارش‌پذیر یا متناهی S به اعداد حقیقی است. این تابع، یک عدد حقیقی را به هر نتیجه‌ی آزمایش نسبت می‌دهد، که به ما اجازه می‌دهد با توزیع احتمال ایجادشده روی مجموعه‌ای از اعداد حاصل کار کنیم. متغیرهای تصادفی می‌توانند برای فضاهای نمونه نامتناهی شمارش‌ناپذیر نیز تعریف شوند، اما مشکلات تکنیکی را به وجود می‌آورند که لازم نیست برای اهداف ما، برطرف شوند. بنابراین، فرض خواهیم کرد که متغیرهای تصادفی گسسته‌اند.

برای متغیر تصادفی X و عدد حقیقی x ، رویداد $X = x$ را $\{s \in S : X(s) = x\}$ تعریف می‌کنیم. بنابراین، داریم:

$$\Pr\{X = x\} = \sum_{s \in S : X(s) = x} \Pr\{s\}$$

تابع زیر، یک **تابع چگال احتمال**[1] از متغیر تصادفی X است:

$$f(x) = \Pr\{X = x\}$$

با توجه به اصول احتمال، $\Pr\{X = x\} \geq 0$ و $\sum_x \Pr\{X = x\} = 1$.

به عنوان مثال، آزمایش چرخاندن یک جفت تاس ۶ وجهی را درنظربگیرید. ۳۶ پیشامد اولیه ممکن در فضای نمونه وجود دارد، فرض می‌کنیم توزیع احتمال، یکنواخت است، به طوری که هر پیشامد اولیه $s \in S$ احتمال یکسانی دارند: $\Pr\{s\} = 1/36$. متغیر تصادفی X را *ماکزیمم* دو مقدارِ نشان‌داده‌شده در تاس درنظربگیرید. داریم $\Pr\{X = 3\} = 5/36$ ، زیرا X مقدار ۳ تا ۵ را به ۳۶ پیشامد اولیه‌ی ممکن، یعنی (1, 3) ، (2, 3) ، (3, 3) ، (3, 2) و (3, 1) نسبت می‌دهد.

متداول است که چندین متغیر تصادفی روی یک فضای نمونه تعریف شوند. اگر X و Y متغیرهای تصادفی باشند، تابع زیر، **تابع چگالی احتمال توأم**[2] X و Y است:

$$f(x, y) = \Pr\{X = x \text{ and } Y = y\}$$

برای مقدار ثابت y داریم:

$$\Pr\{Y = y\} = \sum_x \Pr\{X = x \text{ and } Y = y\}$$

و به طور مشابه، برای مقدار ثابت x داریم:

$$\Pr\{X = x\} = \sum_y \Pr\{X = x \text{ and } Y = y\}$$

1. probability density function 2. joint probability density function

با استفاده از تعریف (پ-۱۴) از احتمال شرطی، داریم:

$$\Pr\{X = x \mid Y = y\} = \frac{\Pr\{X = x \text{ and } Y = y\}}{\Pr\{Y = y\}}$$

دو متغیر تصادفی X و Y را **مستقل** تعریف می‌کنیم اگر برای تمام x و y ، پیشامدهای X = x و Y = y مستقل باشند، یا معادل آن، اگر برای تمام x و y ، داشته باشیم $\Pr\{X = x \text{ and } Y = y\} = \Pr\{X = x\}\Pr\{Y = y\}$.

با توجه به مجموعه‌ای از متغیرهای تعریف‌شده روی یک فضای نمونه، می‌توان متغیرهای تصادفی را به عنوان مجموع، حاصلضرب، یا توابع دیگری از متغیرهای اصلی تعریف کرد.

مقدار مورد انتظار و متغیر تصادفی

ساده‌ترین و مفیدترین خلاصه از توزیع متغیر تصادفی، "میانگین" مقادیری است که روی آن عمل می‌شود. **مقدار مورد انتظار**[1] (یا به طور متقارن، **انتظار** یا **میانگین**) متغیر تصادفی گسسته به صورت زیر است:

$$\mathrm{E}[X] = \sum_x x \Pr\{X = x\} \qquad \text{(پ-۱۹)}$$

که اگر مجموع، متناهی باشد یا به طور مطلق همگرا باشد، خوش‌تعریف است. گاهی انتظار X به صورت μ_X نوشته می‌شود، و وقتی که متغیر تصادفی حذف می‌شود، به صورت μ نمایش داده خواهد شد.

یک بازی را درنظربگیرید که در آن، دو سکه سالم را پرتاب می‌کنید. برای هر شیر ۳ دلار می‌گیرید ولی برای هر خط ۲ دلار از دست می‌دهید. مقدار مورد انتظار متغیر تصادفی X که دستاورد شما را نشان می‌دهد، به صورت زیر است:

$$\begin{aligned} \mathrm{E}[X] &= 6 \cdot \Pr\{2\ \text{H's}\} + 1 \cdot \Pr\{1\ \text{H}, 1\ \text{T}\} - 4 \cdot \Pr\{2\ \text{T's}\} \\ &= 6(1/4) + 1(1/2) - 4(1/4) \\ &= 1 \end{aligned}$$

انتظار مجموع دو متغیر تصادفی، برابر با مجموع انتظارات آن‌ها است، یعنی هر وقت E[X] و E[Y] تعریف می‌شوند، داریم:

$$\mathrm{E}[X + Y] = \mathrm{E}[X] + \mathrm{E}[Y] \qquad \text{(پ-۲۰)}$$

این خاصیت را **خطی‌بودن انتظار** می‌نامیم، که حتی اگر X و Y مستقل نباشند، برقرار است. علاوه‌براین، به حاصل‌جمع‌های متناهی و همگرای مطلق انتظارات بسط داده می‌شود. خطی‌بودن انتظار، خاصیت مهمی است که ما را قادر می‌سازد تحلیل احتمالی را با استفاده از متغیرهای تصادفی شاخص انجام دهیم (بخش ۲-۱۵).

اگر X متغیر تصادفی باشد، هر تابع g(x) متغیر تصادفی جدید g(X) را تعریف می‌کند. اگر انتظار g(X) تعریف شده باشد، آنگاه داریم:

$$\mathrm{E}[g(X)] = \sum_x g(x) \Pr\{X = x\}$$

با قراردادن g(x) = ax ، برای هر ثابت a داریم:

$$\mathrm{E}[aX] = a\mathrm{E}[X] \qquad \text{(پ-۲۱)}$$

1. expected value

در نتیجه، انتظارات خطی‌اند: برای هر دو متغیر تصادفی X و Y و ثابت a ، داریم:

$$\mathrm{E}[aX+Y] = a\mathrm{E}[X] + \mathrm{E}[Y] \tag{پ-۲۲}$$

وقتی دو متغیر تصادفی X و Y مستقل باشند و هر کدام انتظار تعریف‌شده‌ای داشته باشند، داریم:

$$\begin{aligned} \mathrm{E}[XY] &= \sum_x \sum_y xy \Pr\{X = x \text{ and } Y = y\} \\ &= \sum_x \sum_y xy \Pr\{X = x\} \Pr\{Y = y\} \\ &= \left(\sum_x x \Pr\{X = x\}\right) \left(\sum_y y \Pr\{Y = y\}\right) \\ &= \mathrm{E}[X]\mathrm{E}[Y] \end{aligned}$$

به طور کلی، وقتی n متغیر تصادفی $X_1, X_2, \ldots, X_n$ متقابلاً مستقل باشند، داریم:

$$\mathrm{E}[X_1 X_2 \cdots X_n] = \mathrm{E}[X_1]\mathrm{E}[X_2] \cdots \mathrm{E}[X_n] \tag{پ-۲۳}$$

وقتی متغیر تصادفی X مقادیر را از مجموعه‌ای از اعداد طبیعی $\mathbf{N} = \{0,1,2,\ldots\}$ می‌گیرد، فرمول خوبی برای انتظار آن وجود دارد:

$$\begin{aligned} \mathrm{E}[X] &= \sum_{i=0}^{\infty} i \Pr\{X = i\} \\ &= \sum_{i=0}^{\infty} i(\Pr\{X \geq i\} - \Pr\{X \geq i+1\}) \\ &= \sum_{i=1}^{\infty} \Pr\{X \geq i\} \end{aligned} \tag{پ-۲۴}$$

زیرا هر جمله $\Pr\{X \geq i\}$ به تعداد i بار اضافه و $i-1$ بار کم می‌شود (انتظار $\Pr\{X \geq 0\}$، که صفر بار اضافه می‌شود ولی کم نمی‌شود).

وقتی تابع محدب $f(x)$ را به متغیر تصادفی X اِعمال می‌کنیم، با توجه به نامعادله جانسون خواهیم داشت:

$$\mathrm{E}[f(X)] \geq f(\mathrm{E}[X]) \tag{پ-۲۵}$$

به شرطی که انتظارات وجود داشته باشند و متناهی باشند (تابع f(x) **مقعر** است اگر برای هر x و y و برای تمام $0 \leq \lambda \leq 1$، داشته باشیم $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$).

واریانس و انحراف معیار

مقدار مورد انتظار متغیر تصادفی، به ما نمی‌گوید که مقادیر متغیر چگونه توزیع شدند. برای مثال اگر متغیرهای تصادفی X و Y را داشته باشیم که برای آن‌ها:

$$\Pr\{X = 1/4\} = \Pr\{X = 3/4\} = 1/2 \quad \text{و} \quad \Pr\{Y = 0\} = \Pr\{Y = 1\} = 1/2$$

آنگاه هم E[X] و هم E[Y] برابر با 1/2 هستند، و هنوز مقادیر Y نسبت به مقادیر واقعی X ، نسبت به میانگین دورتر هستند.

مفهوم واریانس، از نظر ریاضی بیان می‌کند که مقدار یک متغیر تصادفی، چقدر از میانگین فاصله دارد. **واریانس** متغیر تصادفی X با میانگین E[X] برابر است با:

$$\begin{aligned}\mathrm{Var}[X] &= \mathrm{E}[(X-\mathrm{E}[X])^2] \\ &= \mathrm{E}[X^2 - 2X\mathrm{E}[X] + \mathrm{E}^2[X]] \\ &= \mathrm{E}[X^2] - 2\mathrm{E}[X\mathrm{E}[X]] + \mathrm{E}^2[X] \\ &= \mathrm{E}[X^2] - 2\mathrm{E}^2[X] + \mathrm{E}^2[X] \\ &= \mathrm{E}[X^2] - \mathrm{E}^2[X]\end{aligned} \qquad \text{(پ-۲۶)}$$

توجیه معادلات $\mathrm{E}[\mathrm{E}^2[X]] = \mathrm{E}^2[X]$ و $\mathrm{E}[X\mathrm{E}[X]] = \mathrm{E}^2[X]$ این‌است‌که E[X] متغیر تصادفی نیست ولی عدد حقیقی است، که معنایش این‌است‌که معادله (پ-۲۱) اعمال می‌شود (به طوری که a = E[X]). معادله (پ-۲۶) می‌تواند بازنویسی شود تا عبارتی برای انتظار مربع متغیر تصادفی به دست آید:

$$\mathrm{E}[X^2] = \mathrm{Var}[X] + \mathrm{E}^2[X] \qquad \text{(پ-۲۷)}$$

واریانس متغیر تصادفی X و واریانس aX مرتبط‌اند (تمرین پ-۳-۱۰ را ببینید):

$$\mathrm{Var}[aX] = a^2\mathrm{Var}[X]$$

وقتی X و Y متغیرهای تصادفی مستقل‌اند، داریم:

$$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$$

به طور کلی، اگر n متغیر تصادفی $X_1, X_2, \ldots, X_n$ دو به دو مستقل باشند، آنگاه داریم:

$$\mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{Var}[X_i] \qquad \text{(پ-۲۸)}$$

انحراف معیار متغیر تصادفی X ، ریشه دوم نامنفی واریانس X است. انحراف معیار متغیر تصادفی X ، گاهی با σ_X یا σ نمایش داده می‌شود. با این نمادگذاری، واریانس X با σ^2 نمایش داده می‌شود.

تمرین‌های بخش پ – ۳

تمرین پ-۳-۱: دو تاس ۶ وجهی پرتاب می‌شوند. انتظار مجموع دومتغیری که نشان داده می‌شوند چیست؟ انتظار ماکزیمم این دو مقدار چیست؟

تمرین پ-۳-۲: آرایه A[1 .. n] شامل n عضو متفاوت است که ترتیب تصادفی دارند، به طوری که احتمال هر جایگشت از n عدد یکسان است. انتظار اندیس عنصر ماکزیمم در آرایه چیست؟ انتظار اندیس عنصر می‌نیمم در آرایه چیست؟

تمرین پ-۳-۳: یک بازی کارناوال شامل سه تاس در یک قفس است. بازیکن می‌تواند دلاری را در هر یک از اعداد ۱ تا ۶ قرار دهد. قفسه شکسته می‌شود و پاداش به صورت زیر است. اگر شماره بازیکن در هیچ‌کدام از تاس‌ها ظاهر نشود، دلار خود را از دست می‌دهد. وگرنه، اگر شماره او دقیقاً روی k تاس از سه تاس ظاهر شود، برای k = 1, 2, 3 ، دلار خود را حفظ می‌کند و k دلار دیگر برنده می‌شود. انتظار حاصل از یک بار بازی کارناوال چیست؟

تمرین پ-۳-۴: ثابت کنید اگر X و Y متغیرهای تصادفی نامنفی باشند، آنگاه داریم:

$$\mathrm{E}[\max(X,Y)] \le \mathrm{E}[X] + \mathrm{E}[Y]$$

★ **تمرین پ-۳-۵:** فرض کنید X و Y متغیرهای تصادفی مستقل باشند. ثابت کنید f(X) و g(Y) برای هر انتخاب توابع f و g مستقل‌اند.

⋆ **تمرین پ-۳-۶:** فرض کنید X متغیر تصادفی نامنفی، و $E[X]$ خوش‌تعریف باشد. ثابت کنید نامعادله مارکوف برای $t > 0$ برقرار است:

$$\Pr\{X \geq t\} \leq E[X]/t \tag{پ-۲۹}$$

⋆ **تمرین پ-۳-۷:** فرض کنید S فضای نمونه باشد، و X و X' متغیرهای تصادفی باشند که برای $s \in S$ داریم $X(s) \geq X'(s)$. ثابت کنید برای هر ثابت حقیقی t داریم:

$$\Pr\{X \geq t\} \geq \Pr\{X' \geq t\}$$

تمرین پ-۳-۸: کدام بزرگ‌تر است: انتظار مربع متغیر تصادفی، یا مربع انتظار آن؟

تمرین پ-۳-۹: نشان دهید که برای هر متغیر تصادفی X که فقط مقادیر 0 و 1 را می‌پذیرد، داریم:

$$\mathrm{Var}[X] = \mathrm{E}[X]\,\mathrm{E}[1 - X]$$

تمرین پ-۳-۱۰: از تعریف (پ-۲۶) درباره واریانس، ثابت کنید $\mathrm{Var}[aX] = a^2\mathrm{Var}[X]$.

پ-۴ توزیع‌های هندسی و دوجمله‌ای

پرتاب سکه، نمونه‌ای از **آزمایش برنولی** است، که به عنوان آزمایشی فقط با دو نتیجه‌ی ممکن تعریف شد: **موفق**، که با احتمال p رخ می‌دهد، و **شکست**، که با احتمال $q = 1 - p$ رخ می‌دهد. وقتی راجع‌به آزمایش‌های برنولی صحبت می‌کنیم، معنایش این‌است‌که آزمایش‌ها متقابلاً مستقل‌اند، مگر این‌که غیر از این بیان کنیم، که هر کدام، برای موفقیت احتمال p دارند، دو توزیع از آزمایش‌های برنولی به دست می‌آید. توزیع هندسی و توزیع دوجمله‌ای.

توزیع هندسی

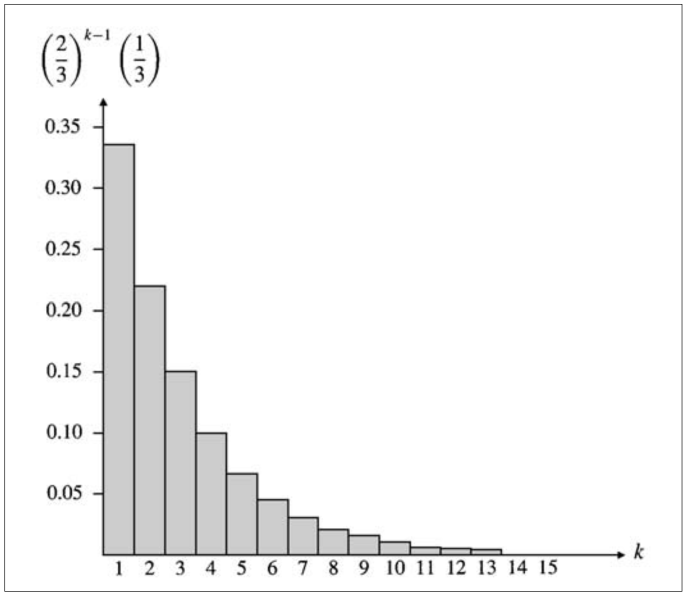
فرض کنید دنباله‌ای از آزمایش‌های برنولی را داریم، که احتمال موفقیت هر کدام p و احتمال شکست هر کدام $q = 1 - p$ است. قبل از موفقیت، چند آزمایش انجام می‌شود؟ فرض کنید متغیر تصادفی X برابر با تعداد آزمایش‌های موردنیاز برای کسب موفقیت است. سپس X دارای مقادیری در بازه‌ی $\{1, 2, \ldots\}$ است، و برای $k \geq 1$ داریم:

$$\Pr\{X = k\} = q^{k-1}p \tag{پ-۳۰}$$

زیرا قبل از یک موفقیت، $k - 1$ شکست داریم. توزیع احتمالی که در معادله (پ-۳۰) صدق می‌کند، **توزیع هندسی** نام دارد. شکل پ – ۱ این توزیع را نشان می‌دهد.

با فرض این‌که $q < 1$، انتظار توزیع هندسی می‌تواند با همانی (الف – ۸) محاسبه شود:

$$\begin{aligned} \mathrm{E}[X] &= \sum_{k=1}^{\infty} kq^{k-1}p \\ &= \frac{p}{q}\sum_{k=0}^{\infty} kq^k \\ &= \frac{p}{q}\cdot\frac{q}{(1-q)^2} \\ &= 1/p \end{aligned} \tag{پ-۳۱}$$



شکل پ – ۱ توزیع احتمال با احتمال $p = 1/3$ موفقیت و $q = 1 - p$ شکست. انتظار توزیع $1/p = 3$ است.

بنابراین، به طور میانگین، قبل از کسب موفقیت، $1/p$ آزمایش انجام می‌گیرد، که نتیجه‌ی شهودی است. واریانس، که می‌تواند به طور مشابه محاسبه شود (ولی با استفاده از تمرین الف–۱–۳)، به صورت زیر است:

$$\mathrm{Var}[X] = q/p^2 \qquad \text{(پ-۳۲)}$$

به عنوان مثال، فرض کنید مکرراً دو تاس را می‌اندازیم تا ۷ یا ۱۱ را به دست آوریم. از ۳۶ نتیجه‌ی ممکن، ۶ منجر به ۷ و ۲ منجر به ۱۱ می‌شود. بنابراین، احتمال موفقیت $p = 8/36 = 2/9$ است و به طور میانگین باید $1/p = 9/2 = 4.5$ بار پرتاب کنیم تا ۷ یا ۱۱ را به دست آوریم.

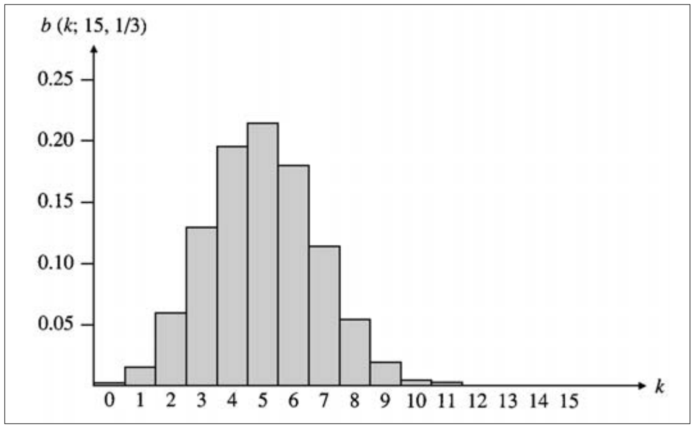
توزیع دوجمله‌ای

چند موفقیت در حین آزمایش‌های برنولی رخ می‌دهد، که در آن، یک موفقیت با احتمال p و شکست با احتمال $q = 1 - p$ رخ خواهد داد؟ متغیر تصادفی X را برابر با تعداد موفقیت‌ها در n آزمایش درنظربگیرید. سپس، X دارای مقادیری در بازه‌ی $\{0, 1, \dots, n\}$ است، و برای $k = 0, \dots, n$ داریم:

$$\Pr\{X = k\} = \binom{n}{k} p^k q^{n-k} \qquad \text{(پ-۳۳)}$$

زیرا $\binom{n}{k}$ روش برای انتخاب k موفقیت از n آزمایش وجود دارد، و احتمال این‌که هر کدام رخ دهند، $p^k q^{n-k}$ است. توزیع احتمالی که در معادله (پ–۳۳) صدق می‌کند، **توزیع دوجمله‌ای** نامیده می‌شود. برای سهولت، خانواده توزیع دوجمله‌ای را با استفاده از نمادگذاری زیر تعریف می‌کنیم:

$$b(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k} \qquad \text{(پ-۳۴)}$$



شکل پ–۲ توزیع دوجمله‌ای $b(k; 15, 1/3)$ ناشی از $n = 15$. آزمایش برنولی، هر یک با احتمال موفقیت $p = 1/3$. امید ریاضی توزیع $np = 5$ است.

شکل پ–۲ توزیع دوجمله‌ای را شرح می‌دهد. نام "دوجمله‌ای" از این حقیقت ناشی می‌شود که (پ–۳۳)، جمله k اُم عبارت $(p+q)^n$ است. در نتیجه، چون $p + q = 1$ ، داریم:

$$\sum_{k=0}^{n} b(k; n, p) = 1 \tag{پ–۳۵}$$

که توسط اصل ۲ از اصول احتمال مورد نیاز است.

می‌توان انتظار متغیر تصادفیِ دارای توزیع دوجمله‌ای را از معادلات (پ–۸) و (پ–۳۵) محاسبه کرد. فرض کنید X متغیر تصادفی باشد که از توزیع دوجمله‌ای $b(k; n, p)$ پیروی می‌کند، و قرار دهید $q = p - 1$. بنا به تعریف انتظار، داریم:

$$\begin{aligned}
\mathrm{E}[X] &= \sum_{k=0}^{n} k \Pr\{X = k\} \\
&= \sum_{k=0}^{n} k\, b(k; n, p) \\
&= \sum_{k=1}^{n} k \binom{n}{k} p^k q^{n-k} \\
&= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} q^{n-k} \\
&= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k q^{(n-1)-k} \\
&= np \sum_{k=0}^{n-1} b(k; n-1, p) \\
&= np
\end{aligned} \tag{پ–۳۶}$$

با استفاده از خاصیت خطی‌بودن انتظار، می‌توان نتیجه مشابهی را با محاسبات جبری کمتری به دست آورد. فرض کنید X_i متغیر تصادفی باشد که تعداد موفقیت‌ها در i اُمین آزمایش را توصیف می‌نماید. سپس

$E[X_i] = p\,.\,1 + q\,.\,0 = p$ ، و بنا به خاصیت خطی‌بودن انتظار (معادله (پ–۲۰))، تعداد موفقیت‌های مورد انتظار برای n آزمایش برابر است با:

$$\begin{aligned} E[X] &= E\left[\sum_{i=1}^{n} X_i\right] \\ &= \sum_{i=1}^{n} E[X_i] \\ &= \sum_{i=1}^{n} p \\ &= np \end{aligned} \qquad \text{(پ–۳۷)}$$

همین روش می‌تواند برای محاسبه واریانس توزیع به کار رود. با استفاده از معادله (پ–۲۶)، داریم $\mathrm{Var}[X_i] = E[X_i^2] - E^2[X_i]$. چون X_i فقط مقادیر 0 و 1 را می‌پذیرد، داریم $E[X_i^2] = E[X_i] = p$ و در نتیجه، داریم:

$$\mathrm{Var}[X_i] = p - p^2 = pq \qquad \text{(پ–۳۸)}$$

برای محاسبه واریانس X ، از امتیاز استقلال n آزمایش استفاده می‌کنیم. بنابراین، بنا به معادله (پ–۲۸) داریم:

$$\begin{aligned} \mathrm{Var}[X] &= \mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] \\ &= \sum_{i=1}^{n} \mathrm{Var}[X_i] \\ &= \sum_{i=1}^{n} pq \\ &= npq \end{aligned} \qquad \text{(پ–۳۹)}$$

همان‌طور که می‌توان در شکل پ–۲ مشاهده کرد، توزیع دوجمله‌ای $b(k; n, p)$ با افزایش k از 0 تا n ، افزایش می‌یابد تا به میانگین np برسد، و سپس کاهش می‌یابد. با مشاهده نسبت جملات موفقیت‌آمیز، می‌توان اثبات کرد که این توزیع همیشه به این روش عمل می‌کند:

$$\begin{aligned} \frac{b(k; n, p)}{b(k-1; n, p)} &= \frac{\binom{n}{k} p^k q^{n-k}}{\binom{n}{k-1} p^{k-1} q^{n-k+1}} \\ &= \frac{n!(k-1)!(n-k+1)!p}{k!(n-k)!n!q} \\ &= \frac{(n-k+1)p}{kq} \\ &= 1 + \frac{(n+1)p - k}{kq} \end{aligned} \qquad \text{(پ–۴۰)}$$

وقتی $(n + 1)(p - k)$ مثبت باشد، این نسبت بزرگ‌تر از یک است. در نتیجه، برای $k < (n + 1)p$ داریم $b(k;n,p) > b(k-1;n,p)$ (توزیع افزایش می‌یابد)، و برای $k > (n + 1)p$ داریم $b(k;n,p) < b(k-1;n,p)$ (توزیع کاهش می‌یابد). اگر $k = (n + 1)p$ صحیح باشد، آنگاه $b(k;n,p) = b(k-1;n,p)$ ، لذا این توزیع دارای

دو ماکزیمم است: در $k = (n+1)p$ و در $k-1 = (n+1)p-1 = np-q$. وگرنه، در مقدار صحیح یکتای k به ماکزیمم می‌رسد که k در بازه‌ی $np-q < k < (n+1)p$ است. لِم زیر کران بالایی را روی توزیع دوجمله‌ای ایجاد می‌کند.

لِم پ-۱

فرض کنید $n \geq 0$ ، $0 < p < 1$ ، $q = 1-p$ و $0 \leq k \leq n$. آنگاه داریم:

$$b(k; n, p) \leq \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k}$$

اثبات: با استفاده از معادله (پ-۶) داریم:

$$\begin{aligned} b(k; n, p) &= \binom{n}{k} p^k q^{n-k} \\ &\leq \left(\frac{n}{k}\right)^k \left(\frac{n}{n-k}\right)^{n-k} p^k q^{n-k} \\ &= \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k} \end{aligned}$$ ■

تمرین‌های بخش پ – ۴

تمرین پ-۴-۱: اصل ۲ از اصول احتمال را برای توزیع هندسی وارسی کنید.

تمرین پ-۴-۲: به طور میانگین چند بار باید ۶ سکه سالم را قبل از به دست آوردن ۳ شیر و ۳ خط پرتاب کنیم؟

تمرین پ-۴-۳: نشان دهید $b(k; n, p) = b(n-k; n, q)$ که $q = 1-p$.

تمرین پ-۴-۴: نشان دهید که مقدار ماکزیمم توزیع دوجمله‌ای $b(k; n, p)$ تقریباً $1/\sqrt{2\pi npq}$ است، که $q = 1-p$.

تمرین پ-۴-۵: نشان دهید که احتمال عدم موفقیت در آزمایش برنولی، هر کدام با احتمال $p = 1/n$ ، تقریباً $1/e$ است. نشان دهید که احتمال دقیقاً یک موفقیت، تقریباً $1/e$ است.

★ **تمرین پ-۴-۶:** پروفسور Rosencrantz یک سکه سالم را n بار پرتاب می‌کند، و پروفسور Guildenstern نیز همین کار را می‌کند. نشان دهید که احتمال این‌که تعداد شیرهای یکسانی را به دست آورند، برابر با $\binom{2n}{n}/4^n$ است (*راهنمایی:* برای پرفسور Rosencrantz ، شیر را موفقیت درنظربگیرید؛ برای Guildenstern ، خط را موفقیت درنظربگیرید). با استفاده از بحث خود، همانی زیر را اثبات کنید:

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}$$

★ **تمرین پ-۴-۷:** نشان دهید که برای $0 \leq k \leq n$ داریم:

$$b(k; n, 1/2) \leq 2^{n H(k/n) - n}$$

که $H(n)$ تابع آنتروپی است (پ-۷).

★ **تمرین پ-۴-۸:** n آزمایش برنولی را درنظربگیرید که برای $i = 1, 2, \ldots, n$ ، آزمایش i اُم دارای احتمال موفقیت p_i است و فرض کنید X متغیر تصادفی باشد که تعداد کل موفقیت‌ها را نشان می‌دهد. فرض کنید برای $i = 1, 2, \ldots, n$ داریم $p \geq p_i$. ثابت کنید برای $1 \leq k \leq n$ داریم:

$$\Pr\{X < k\} \geq \sum_{i=0}^{k-1} b(i; n, p)$$

★ **تمرین پ-۴-۹:** فرض کنید X متغیر تصادفی برای تعداد کل موفقیت‌ها در مجموعه A از n آزمایش برنولی باشد، که آزمایش i اُم دارای احتمال موفقیت p_i است، و فرض کنید X' متغیر تصادفی برای تعداد کل موفقیت‌ها در مجموعه دوم A' از n آزمایش برنولی باشد، که آزمایش i اُم دارای احتمال موفقیت $p'_i \geq p_i$ است. ثابت کنید برای $0 \leq k \leq n$ داریم:

$$\Pr\{X' \geq k\} \geq \Pr\{X \geq k\}$$

(*راهنمایی:* نشان دهید که چگونه می‌توان آزمایش‌های برنولی را در A' به دست آورد (از نتیجه تمرین پ-۳-۷ استفاده کنید).

★پ-۵ دنباله‌های توزیع دوجمله‌ای[1]

احتمالِ داشتن حداقل، یا حداکثر k موفقیت در آزمایش برنولی، هر یک با احتمال موفقیت p ، غالباً جالب‌تر از احتمالَ دقیقاً k موفقیت است. در این بخش، **خط‌های** توزیع دوجمله‌ای را بررسی می‌کنیم: دو ناحیه از توزیع $b(k; n, p)$ که از میانگین np دور هستند. چندین کران مهم را روی (مجموع تمام جملات در) یک دنباله اثبات خواهیم کرد.

ابتدا کرانی را روی دنباله راست توزیع $b(k; n, p)$ فراهم می‌سازیم. کران‌های دنباله چپ می‌تواند با معکوس‌کردن نقش‌های موفقیت و شکست‌ها تعیین شود.

قضیه پ - ۲

دنباله‌ای از آزمایش‌های برنولی را درنظربگیرید، که موفقیت با احتمال p به دست می‌آید. فرض کنید X متغیر تصادفی باشد که تعداد کل موفقیت‌ها را نشان می‌دهد. آنگاه برای $0 \leq k \leq n$، احتمال حداقل k موفقیت برابر است با:

$$\begin{aligned}\Pr\{X \geq k\} &= \sum_{i=k}^{n} b(i; n, p) \\ &\leq \binom{n}{k} p^k\end{aligned}$$

اثبات: برای $S \subseteq \{1,2,\ldots,n\}$ فرض می‌کنیم پیشامد موفق‌بودن آزمایش i اُم برای هر $i \in S$ باشد. بدیهی است که اگر $|S| = k$ ، داریم $\Pr\{A_S\} = p^k$. داریم:

$$\begin{aligned}\Pr\{X \geq k\} &= \Pr\{\text{there exists } S \subseteq \{1, 2, \ldots, n\} : |S| = k \text{ and } A_S\} \\ &= \Pr\left\{\bigcup_{S\subseteq\{1,2,\ldots,n\}:|S|=k} A_S\right\} \\ &\leq \sum_{S\subseteq\{1,2,\ldots,n\}:|S|=k} \Pr\{A_S\} \\ &= \binom{n}{k} p^k ,\end{aligned}$$

که این نامعادله از نامعادله بول (پ-۱۸) به دست می‌آید. ■

قضیه فرعی زیر، این قضیه را برای دنباله‌ی سمت چپ توزیع دوجمله‌ای بیان می‌کند. به طور کلی، اثبات آن را به عهده خواننده واگذار می‌کنیم.

1. tails of the binomial distribution

قضیه فرعی پ-۳

دنباله‌ای از آزمایش‌های برنولی را درنظربگیرید، که موفقیت با احتمال p رخ می‌دهد. اگر X متغیر تصادفی باشد که تعداد کل موفقیت‌ها را نشان می‌دهد، آنگاه برای $0 \le k \le n$، احتمال حداکثر k موفقیت به صورت زیر است:

$$\begin{aligned}\Pr\{X \le k\} &= \sum_{i=0}^{k} b(i; n, p) \\ &\le \binom{n}{n-k}(1-p)^{n-k} \\ &= \binom{n}{k}(1-p)^{n-k} \qquad \blacksquare\end{aligned}$$

کران بعدی که تعیین می‌کنیم به دنباله‌ی چپ توزیع دوجمله‌ای مربوط می‌شود. قضیه فرعی آن نشان می‌دهد که دور از میانگین، دنباله چپ به طور نمایی کاهش می‌یابد.

قضیه پ-۴

دنباله‌ای از n آزمایش برنولی را درنظربگیرید که احتمال موفقیت p و احتمال شکست $q = 1 - p$ است. فرض کنید X متغیر تصادفی باشد که تعداد کل موفقیت‌ها را نشان می‌دهد. آنگاه برای $0 < k < np$ ، احتمال کمتر از k موفقیت برابر است با:

$$\begin{aligned}\Pr\{X < k\} &= \sum_{i=0}^{k-1} b(i; n, p) \\ &< \frac{kq}{np-k}\, b(k; n, p)\end{aligned}$$

اثبات: سری $\sum_{i=0}^{k-1} b(i;n,p)$ را به وسیله سری هندسی با استفاده از تکنیک بخش (الف – ۲) محدود می‌کنیم. برای $i = 1, 2, \ldots, k$ ، از معادله (پ-۴۰) داریم:

$$\begin{aligned}\frac{b(i-1; n, p)}{b(i; n, p)} &= \frac{iq}{(n-i+1)p} \\ &< \frac{iq}{(n-i)p} \\ &\le \frac{kq}{(n-k)p}\end{aligned}$$

اگر قرار دهیم:

$$\begin{aligned}x &= \frac{kq}{(n-k)p} \\ &< \frac{kq}{(n-np)p} \\ &= \frac{kq}{nqp} \\ &= \frac{k}{np} \\ &< 1\end{aligned}$$

نتیجه می‌شود که برای $0 < i \le k$ داریم:

$$b(i-1; n, p) < x\, b(i; n, p)$$

با اعمال مکرر این نامعادله به تعداد $k-i$ بار، به ازای $0 < i \le k$ ، خواهیم داشت:

$$b(i; n, p) < x^{k-i}\, b(k; n, p)$$

و در نتیجه داریم:

$$\begin{aligned}\sum_{i=0}^{k-1} b(i; n, p) &< \sum_{i=0}^{k-1} x^{k-i} b(k; n, p)\\ &< b(k; n, p)\sum_{i=1}^{\infty} x^i\\ &= \frac{x}{1-x}\, b(k; n, p)\\ &= \frac{kq}{np-k}\, b(k; n, p)\end{aligned}$$

■

قضیه فرعی پ-۵

دنباله‌ای از n آزمایش برنولی را درنظربگیرید، که موفقیت با احتمال p و شکست با احتمال $q = 1 - p$ رخ می‌دهد. آنگاه برای $0 < k \le np/2$ ، احتمال کمتر از k موفقیت، کمتر از نیمی از احتمال کمتر از $k+1$ موفقیت است.

اثبات: چون $k \le np/2$، داریم:

$$\begin{aligned}\frac{kq}{np-k} &\le \frac{(np/2)q}{np-(np/2)}\\ &= \frac{(np/2)q}{np/2}\\ &\le 1\end{aligned}$$

زیرا $q \le 1$. فرض کنید X متغیر تصادفی باشد که تعداد موفقیت‌ها را نشان می‌دهد. قضیه پ - ۴ دلالت می‌کند که احتمالِ کمتر از k موفقیت برابر است با:

$$\Pr\{X < k\} = \sum_{i=0}^{k-1} b(i; n, p) < b(k; n, p)$$

بنابراین، داریم:

$$\begin{aligned}\frac{\Pr\{X < k\}}{\Pr\{X < k+1\}} &= \frac{\sum_{i=0}^{k-1} b(i; n, p)}{\sum_{i=0}^{k} b(i; n, p)}\\ &= \frac{\sum_{i=0}^{k-1} b(i; n, p)}{\sum_{i=0}^{k-1} b(i; n, p) + b(k; n, p)}\\ &< 1/2\end{aligned}$$

چون $\sum_{i=0}^{k-1} b(i;n,p) < b(k;n,p)$. ■

تعیین کران‌ها روی دنباله‌ی راست، می‌تواند به طور مشابه تعیین شود. اثبات آن‌ها به عنوان تمرین پ-۵-۲ درنظر گرفته شده است.

قضیه فرعی پ-۶

دنباله‌ای از n آزمایش برنولی را درنظربگیرید، که موفقیت با احتمال p رخ می‌دهد. فرض کنید X متغیر تصادفی باشد که تعداد کل موفقیت‌ها را نشان می‌دهد. آنگاه برای $np < k < n$ ، احتمالِ بیشتر از k موفقیت برابر است با:

$$\begin{aligned}\Pr\{X > k\} &= \sum_{i=k+1}^{n} b(i; n, p) \\ &< \frac{(n-k)p}{k-np}\, b(k; n, p)\end{aligned} \qquad \blacksquare$$

قضیه فرعی پ-۷

دنباله‌ای از n آزمایش برنولی را درنظربگیرید، که موفقیت با احتمال p و شکست با احتمال $q = 1 - p$ رخ می‌دهد. آنگاه برای $(np + n) / 2 < k < n$ ، احتمال بیشتر از k موفقیت، کمتر از نیمی از احتمال بیشتر از $k - 1$ موفقیت است. ■

قضیه بعدی، n آزمایش برنولی را درنظرمی‌گیرد، که برای $i = 1, 2, \ldots , n$ ، احتمال موفقیت هر کدام p_i است. همان‌طور که قضیه فرعی بعدی نشان می‌دهد، می‌توان با استفاده این از قضیه، کرانی را روی دنباله‌ی راست توزیع دوجمله‌ای تعیین کرد. برای این کار، برای هر آزمایش قرار دهید $p_i = p$.

قضیه فرعی پ-۸

دنباله‌ای از n آزمایش برنولی را درنظربگیرید که در آزمایش i اُم، برای $i = 1, 2, \ldots , n$ ، موفقیت با احتمال p_i و شکست با احتمال $q_i = 1 - p_i$ رخ می‌دهد. فرض کنید X متغیر تصادفی باشد که تعداد کل موفقیت‌ها را توصیف می‌نماید، و قرار دهید $\mu = E[X]$. آنگاه، برای $r > \mu$ داریم:

$$\Pr\{X - \mu \geq r\} \leq \left(\frac{\mu e}{r}\right)^r$$

اثبات: چون برای هر $\alpha > 0$، تابع $e^{\alpha x}$ در x اکیداً صعودی است، داریم:

$$\Pr\{X - \mu \geq r\} = \Pr\{e^{\alpha(X-\mu)} \geq e^{\alpha r}\} \qquad \text{(پ-۴۱)}$$

که α بعداً تعیین می‌شود. با استفاده از نامعادله مارکوف (پ-۲۹)، خواهیم داشت:

$$\Pr\{e^{\alpha(X-\mu)} \geq e^{\alpha r}\} \leq \mathrm{E}[e^{\alpha(X-\mu)}]\, e^{-\alpha r} \qquad \text{(پ-۴۲)}$$

قسمت عمده‌ی اثبات، مربوط به محدودکردن $E[e^{\alpha(X-\mu)}]$ و جایگزینی مقدار مناسبی برای α در نامعادله (پ-۴۲) است. ابتدا، $E[e^{\alpha(X-\mu)}]$ را ارزیابی می‌کنیم. با استفاده از نمادگذاری بخش ۲-۵، فرض کنید برای $i = 1, 2, \ldots , n$ داریم $X_i = I\{$آزمایش برنولی i اُم موفق است$\}$ ، یعنی X_i متغیر تصادفی است که وقتی یک است که آزمایش برنولی i اُم موفق باشد و وقتی صفر است که شکست بخورد. بنابراین، داریم:

$$X = \sum_{i=1}^{n} X_i$$

و بنا به خاصیت خطی‌بودن انتظار، داریم:

$$\mu = \mathrm{E}[X] = \mathrm{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{E}[X_i] = \sum_{i=1}^{n} p_i$$

که دلالت می‌کند:

$$X - \mu = \sum_{i=1}^{n} (X_i - p_i)$$

برای ارزیابی $E[e^{\alpha(X-\mu)}]$، برای $X - \mu$ جایگزینی انجام می‌دهیم، و به دست می‌آوریم:

$$\begin{aligned} \mathrm{E}[e^{\alpha(X-\mu)}] &= \mathrm{E}\left[e^{\alpha\sum_{i=1}^{n}(X_i-p_i)}\right] \\ &= \mathrm{E}\left[\prod_{i=1}^{n} e^{\alpha(X_i-p_i)}\right] \\ &= \prod_{i=1}^{n}\mathrm{E}[e^{\alpha(X_i-p_i)}] \end{aligned}$$

که از معادله (پ–۲۳) به دست می‌آید، زیرا استقلال متقابل متغیرهای تصادفی X_i، برای استقلال متقابل متغیرهای تصادفی $e^{\alpha(X_i-p_i)}$ دلالت می‌کند (تمرین پ–۳–۵ را ببینید). بنا به تعریف انتظار، داریم:

$$\begin{aligned} \mathrm{E}[e^{\alpha(X_i-p_i)}] &= e^{\alpha(1-p_i)}p_i + e^{\alpha(0-p_i)}q_i \\ &= p_i e^{\alpha q_i} + q_i e^{-\alpha p_i} \\ &\le p_i e^{\alpha} + 1 \\ &\le \exp(p_i e^{\alpha}) \end{aligned} \tag{پ–۴۳}$$

که $\exp(x)$ تابع نمایی است: $\exp(x) = e^x$ (نامعادله (پ–۴۳) از نامعادلات $\alpha > 0$، $q_i \le 1$، $e^{\alpha q_i} \le e^{\alpha}$ و $e^{-\alpha p_i} \le 1$ به دست می‌آید، و خط آخر از نامعادله (۳–۱۱) به دست می‌آید. در نتیجه داریم:

$$\begin{aligned} \mathrm{E}[e^{\alpha(X-\mu)}] &= \prod_{i=1}^{n}\mathrm{E}[e^{\alpha(X_i-p_i)}] \\ &\le \prod_{i=1}^{n}\exp(p_i e^{\alpha}) \\ &= \exp\left(\sum_{i=1}^{n} p_i e^{\alpha}\right) \\ &= \exp(\mu e^{\alpha}) \end{aligned} \tag{پ–۴۴}$$

زیرا $\mu = \sum_{i=1}^{n} p_i$. بنابراین، از معادله (پ–۴۱) و نامعادلات (پ–۴۲) و (پ–۴۴) نتیجه می‌شود که:

$$\Pr\{X-\mu \ge r\} \le \exp(\mu e^{\alpha} - \alpha r) \tag{پ–۴۵}$$

با انتخاب $\alpha = \ln(r/\mu)$ (تمرین پ–۵–۷) به دست می‌آوریم:

$$\begin{aligned} \Pr\{X-\mu \ge r\} &\le \exp(\mu e^{\ln(r/\mu)} - r\ln(r/\mu)) \\ &= \exp(r - r\ln(r/\mu)) \\ &= \frac{e^r}{(r/\mu)^r} \\ &= \left(\frac{\mu e}{r}\right)^r \end{aligned}$$

■

وقتی به آزمایش‌های برنولی اعمال می‌شود که در آن، هر آزمایش احتمال موفقیت یکسانی دارد، قضیه پ–۸ منجر به قضیه فرعی زیر می‌شود که دنباله سمت راست توزیع دوجمله‌ای را محدود می‌کند.

قضیه فرعی پ–۹

دنباله‌ای از n آزمایش برنولی را درنظربگیرید، که احتمال موفقیت p و احتمال شکست $q = 1 - p$ است. آنگاه، برای $r > np$ داریم:

$$\begin{aligned}\Pr\{X - np \geq r\} &= \sum_{k=\lceil np+r\rceil}^{n} b(k; n, p) \\ &\leq \left(\frac{npe}{r}\right)^r\end{aligned}$$

اثبات: بنا به معادله (پ-۳۶)، داریم $\mu = E[X] = np$. ■

تمرین‌های بخش پ – ۵

★**تمرین پ-۵-۱:** احتمال کدام کمتر است: به وجود نیامدن هیچ شیر وقتی که سکه سالمی را n بار پرتاب می‌کنید، یا به دست آوردن کمتر از n شیر وقتی سکه را $4n$ بار پرتاب می‌کنید؟

★**تمرین پ-۵-۲:** قضیه‌های فرعی پ-۶ و پ-۷ را اثبات کنید.

★**تمرین پ-۵-۳:** برای تمام $a > 0$ و k به طوری که $0 < k < n$ نشان دهید:

$$\sum_{i=0}^{k-1} \binom{n}{i} a^i < (a+1)^n \frac{k}{na - k(a+1)} b(k; n, a/(a+1))$$

★**تمرین پ-۵-۴:** ثابت کنید اگر $0 < k < np$ ، که $0 < p < 1$ و $q = 1 - p$ ، آنگاه:

$$\sum_{i=0}^{k-1} p^i q^{n-i} < \frac{kq}{np-k} \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k}$$

★**تمرین پ-۵-۵:** نشان دهید که شرایط قضیه پ-۸ دلالت می‌کند که:

$$\Pr\{\mu - X \geq r\} \leq \left(\frac{(n-\mu)e}{r}\right)^r$$

به طور مشابه، نشان دهید که شرایط قضیه فرعی پ-۹ دلالت می‌کند که:

$$\Pr\{np - X \geq r\} \leq \left(\frac{nqe}{r}\right)^r$$

★**تمرین پ-۵-۶:** دنباله‌ای از n آزمایش برنولی را درنظربگیرید، که در i آزمایش، برای $i = 1, 2, \ldots, n$ ، موفقیت با احتمال p_i و شکست با احتمال $q_i = 1 - p_i$ رخ می‌دهد. فرض کنید X متغیر تصادفی باشد که تعداد کل موفقیت‌ها را توصیف می‌کند، و $\mu = E[X]$. نشان دهید که برای $r \geq 0$ داریم:

$$\Pr\{X - \mu \geq r\} \leq e^{-r^2/2n}$$

(*راهنمایی:* ثابت کنید که $p_i e^{\alpha q_i} + q_i e^{-\alpha p_i} \leq e^{\alpha^2/2}$. سپس از طرح اثبات قضیه پ-۸ پیروی کنید، به طوری که به جای نامعادله (پ-۳۴) از این نامعادله استفاده کنید).

★**تمرین پ-۵-۷:** نشان دهید که سمت راست نامعادله (پ-۴۵) با انتخاب $\alpha = \ln(r/\mu)$ می‌نیمم شود.

پ–۶ مسئله‌ها

مسئله پ-۱: توپ‌ها و جعبه‌ها.

در این مسئله، اثر فرض‌های مختلف روی تعداد روش‌های قراردادن n توپ در b جعبه مختلف را بررسی می‌کنیم.

الف. فرض کنید n توپ متفاوت‌اند و ترتیب آن‌ها در یک جعبه، مهم نیست. ثابت کنید تعداد روش‌های قراردادن توپ‌ها در جعبه‌ها، b^n است.

ب. فرض کنید توپ‌ها متفاوت‌اند و توپ‌ها در هر جعبه، مرتب‌اند. ثابت کنید دقیقاً $(b+n-1)!/(b-1)!$ روش برای قراردادن توپ‌ها در جعبه‌ها وجود دارد (*راهنمایی:* تعداد روش‌های چیدمان n توپ متمایز و $b-1$ عدد جعبه را در یک سطر درنظربگیرید).

پ. فرض کنید توپ‌ها یکسان هستند، و در نتیجه ترتیب آن‌ها در داخل جعبه مهم نیست. نشان دهید که تعداد روش‌های قراردادن توپ‌ها در جعبه‌ها برابر با $\binom{b+n-1}{n}$ است (*راهنمایی:* از چیدمان‌های قسمت (ب)، نشان دهید که اگر توپ‌ها یکسان باشند، چند بار تکرار می‌شوند).

ت. فرض کنید توپ‌ها یکسان هستند و هیچ جعبه‌ای نمی‌تواند بیش از یک توپ داشته باشد. نشان دهید که تعداد روش‌های قراردادن توپ‌ها، $\binom{b}{n}$ است.

ث. فرض کنید توپ‌ها یکسان هستند و هیچ جعبه‌ای نباید خالی باشد. نشان دهید که تعداد روش‌های قراردادن توپ‌ها برابر با $\binom{n-1}{b-1}$ است.