پیوست پ
شمارش و احتمال

در این پیوست، ترکیب‌های مقدماتی و نظریه‌ی احتمال را بررسی می‌کنیم. اگر با این موضوعات آشنایی دارید، می‌توانید به بخش‌های بعدی بپردازید. بعضی از فصل‌های این کتاب، به احتمالات نیاز ندارند.

بخش پ – ۱، نتایج اولیه‌ی نظریه‌ی شمارش را مرور می‌کند، از جمله فرمول‌های استاندارد برای شمارش جایگشت‌ها و ترکیبات. اصول احتمال و قوانین اصلی مربوط به توزیع احتمال، در بخش پ – ۲ آمده است. متغیرهای تصادفی در بخش پ – ۳ معرفی شدند، در همین بخش، خواص امید ریاضی و واریانس نیز معرفی شدند. بخش پ – ۴ توزیع‌های دوجمله‌ای و هندسی را بحث می‌کند که در مطالعه‌ی آزمایش‌های برنولی ظاهر می‌شوند. مطالعه‌ی توزیع دوجمله‌ای در بخش پ – ۵ ادامه می‌یابد که بحث پیشرفته‌ای از توزیع را ارائه می‌کند.

پ–۱ شمارش

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

قوانین جمع و ضرب

گاهی می‌توانیم مجموعه‌ای از اقلام را که می‌خواهیم شمارش کنیم، به صورت اجتماع مجموعه‌های جدا از هم یا حاصل‌ضرب دکارتی مجموعه‌ها بیان کنیم.

قانون جمع می‌گوید که تعداد راه‌ها برای انتخاب عنصری از یکی از دو مجموعه‌ی جدا ازهم، برابر با مجموع اعداد اصلی[1] (اندازه‌ی) دو مجموعه است. یعنی، اگر A و B دو مجموعه‌ی متناهی باشند که هیچ عضو مشترکی نداشته باشند، آن‌گاه $|A \cup B| = |A| + |B|$ که از معادله‌ی (ب – ۳) به دست می‌آید. برای مثال، هر مکان در پلاک 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 یک 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 مجموعه‌ی متناهی باشند،

1. Cardinality

آن‌گاه $|A\times B|=|A|.|B|$ که همان معادله‌ی (ب–۴) است. برای مثال اگر فروشگاه بستنی، ۲۸ نوع بستنی میوه‌ای و ۴ افزودنی مختلف ارائه دهد، تعداد بستنی‌های میوه‌ای همراه با یک ماده‌ی افزودنی برابر با $28\cdot 4=112$ است.

رشته‌ها

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

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

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

رشته‌ی k عنصری بر روی مجموعه‌ی 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 رشته‌ی k عنصری دودویی، برابر با 2^k است. از نظر شهودی، برای ساخت رشته‌ی k عنصری بر روی مجموعه‌ی n عنصری، 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 k بار ضرب $\underbrace{n.n...n}_{\text{n بار}}=n^k$ به عنوان تعداد رشته‌ی k عنصری می‌شود.

جایگشت‌ها

یک جایگشت از مجموعه‌ی متناهی S ، دنباله‌ی 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 مجموعه‌ی 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 مرتبه‌ی k از S ، دنباله‌ی مرتبی از k عنصر S است که هیچ عنصری بیش از یک بار در دنباله ظاهر نمی‌شود (بنابراین، جایگشت معمولی، فقط جایگشت مرتبه‌ی n از مجموعه‌ی n عضوی است). ۱۲ جایگشت ۲ عضوی از مجموعه‌ی $\{a, b, c, d\}$ عبارتنداز:

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

تعداد جایگشت k عضوی از مجموعه‌ی n عضوی 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 می‌شود. برای مثال مذکور، با $n=4$ و $k=2$ فرمول (پ–۱) به $4!/2!=12$ ارزیابی می‌شود که با جایگشت–۲ ارائه شده مطابقت دارد.

ترکیب‌ها

ترکیب k عضوی از یک مجموعه‌ی n عضوی به نام S ، زیرمجموعه‌ی k عضوی از S است. برای مثال، شش ترکیب ۲ عضوی از مجموعه‌ی ۴ عضوی $\{a, b, c, d\}$ وجود دارد:

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

(در این‌جا برای اختصار، آکولادهای اطراف مجموعه را حذف می‌کنیم). می‌توان ترکیب k عضوی از مجموعه‌ی n عضوی را با انتخاب k عنصر مختلف از مجموعه‌ی n عضوی، ایجاد کرد. ترتیب انتخاب عناصر اهمیتی ندارد.

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

$$\frac{n!}{k!\,(n-k)!}. \qquad \text{(پ-۲)}$$

برای $k=0$ ، این فرمول به ما می‌گوید که تعداد راه‌های انتخاب صفر عنصر از مجموعه‌ی n عضوی، 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 عضوی از مجموعه‌ی n عضوی استفاده می‌کنیم. بنا به معادله‌ی (پ-۲) داریم:

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

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

$$\binom{n}{k}=\binom{n}{n-k}. \qquad \text{(پ-۳)}$$

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

$$(x+y)^n=\sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}\ , \qquad \text{(پ-۴)}$$

که $n\in\mathbb{N}$ و $x,y\in\mathbb{R}$. سمت راست معادله‌ی (پ-۴) بسط دوجمله‌ای سمت چپ است. 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 رشته‌ی دودویی به طول n ، توسط تعداد یک‌هایی است که دارند: تعداد $\binom{n}{k}$ رشته‌ی n عنصری دودویی وجود دارند که دقیقاً شامل k عدد ۱ است، زیرا $\binom{n}{k}$ روش برای انتخاب k مکان از n مکان وجود دارد که می‌توان یک را در آن‌جا قرار داد.

بسیاری از اتحادها شامل ضرایب دوجمله‌ای هستند. در تمرین‌ها بعضی از آن‌ها را اثبات خواهید کرد.

کران‌های دوجمله‌ای

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

$$\begin{aligned}\binom{n}{k} &= \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1}\\ &= \left(\frac{n}{k}\right)\left(\frac{n-1}{k-1}\right)\cdots\left(\frac{n-k+1}{1}\right)\ \ge\ \left(\frac{n}{k}\right)^k. \qquad \text{(پ-۵)}\end{aligned}$$

با استفاده از نامعادله‌ی $k! \geq (k/e)^k$ که از تقریب استرلینگ (۳-۲۵) مشتق می‌شود، کران‌های بالای زیر را به دست می‌آوریم:

$$\begin{aligned}\binom{n}{k} &= \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1} \\ &\leq \frac{n^k}{k!} \\ &\leq \left(\frac{en}{k}\right)^k\end{aligned} \tag{پ-۶}$$

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

$$\binom{n}{k} \leq \frac{n^n}{k^k (n-k)^{n-k}}, \tag{پ-۷}$$

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:

$$\begin{aligned}\binom{n}{\lambda n} &\leq \frac{n^n}{(\lambda n)^{\lambda n}((1-\lambda)n)^{(1-\lambda)n}} \\ &= \left(\left(\frac{1}{\lambda}\right)^{\lambda}\left(\frac{1}{1-\lambda}\right)^{1-\lambda}\right)^n \\ &= 2^{nH(\lambda)},\end{aligned}$$

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

$$H(\lambda) = -\lambda \lg \lambda - (1-\lambda)\lg(1-\lambda) \tag{پ-۸}$$

تمرین‌ها

تمرین پ-۱-۱: یک رشته‌ی n عنصری، چند زیررشته‌ی k عنصری دارد؟ (رشته‌های k عنصری یکسان در مکان‌های مختلف را، متفاوت در نظر بگیرید). یک رشته‌ی n عنصری در مجموع، چند زیررشته دارد؟

تمرین پ-۱-۲: **تابع بولی** با n ورودی و m خروجی، تابعی از $\{0,1\}^n$ به $\{0,1\}^m$ است. چند تابع با n ورودی و یک خروجی وجود دارند؟ چند تابع با n ورودی و m خروجی وجود دارند؟

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

تمرین پ-۱-۴: برای انتخاب سه عدد مختلف از مجموعه‌ی $\{1, 2, \ldots, 99\}$ ، 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:

$$\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1} \tag{پ-۹}$$

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

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

1. entropy function

تمرین پ-۱-۷: برای انتخاب k شیء از n شیء، می‌توانید یکی از اشیا را متمایز از اشیای دیگر در نظر بگیرید 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 نتیجه‌ی 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، $\binom{2}{0}$، $\binom{2}{1}$ و $\binom{2}{2}$ 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 بیشینه $\binom{n}{k}$ 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:

$$\binom{n}{j+k} \le \binom{n}{j}\binom{n-j}{k}. \qquad \text{(پ-۱۰)}$$

هم اثبات جبری و هم یک برهان بر اساس روش انتخاب $j + k$ قلم از n قلم ارائه دهید. مثالی ارائه دهید که در آن، تساوی برقرار نیست.

★ **تمرین پ-۱-۱۲:** با استفاده از استقرا روی $0 \le k \le n/2$، نامساوی (پ-۷) را اثبات کنید و با استفاده از معادله‌ی (پ-۳)، آن را برای تمام $0 \le k \le n$ بسط دهید.

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

$$\binom{2n}{n} = \frac{2^{2n}}{\sqrt{\pi n}}(1 + O(1/n)). \qquad \text{(پ-۱۱)}$$

★ **تمرین پ-۱-۱۴:** با مشتق‌گیری از تابع آنتروپی $H(\lambda)$، نشان دهید که مقدار بیشینه‌ی 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:

$$\sum_{k=0}^{n} \binom{n}{k} k = n2^{n-1}. \qquad \text{(پ-۱۲)}$$

★ **تمرین پ-۱-۱۶:** نامعادله‌ی (پ-۵) کران پایینی روی ضریب دو جمله‌ای $\binom{n}{k}$ ارائه می‌دهد. برای مقادیر کوچک k، یک کران قوی برقرار است. اثبات کنید که برای $k \le \sqrt{n}$ داریم:

$$\binom{n}{k} \ge \frac{n^k}{4k!}. \qquad \text{(پ-۱۳)}$$

پ-۲ احتمال

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

احتمال را برحسب **فضای نمونه‌ی**[1] S تعریف می‌کنیم که مجموعه‌ای است که عناصر آن **پیشامدهای مقدماتی**[2] یا **پیشامدهای ساده** نامیده می‌شوند. هر پیشامد ساده را می‌توان نتیجه‌ی ممکن از یک آزمایش دانست. برای آزمایش مربوط به پرتاب دو سکه‌ی متمایز که نتیجه‌ی هر پرتاب شیر (H) یا خط (T) است، می‌توان فضای نمونه را مجموعه‌ای از رشته‌های ۲ عنصری بر روی {H, T} دانست:

$$S = \{\mathrm{HH}, \mathrm{HT}, \mathrm{TH}, \mathrm{TT}\}.$$

پیشامد، زیرمجموعه‌ای از فضای نمونه‌ی 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 حتمی**[3] و پیشامد $\varnothing$ را **پیشامد تهی** یا **پوچ** می‌نامیم. اگر $A \cap B = \varnothing$ می‌گوییم پیشامدهای A و B **دو به دو ناسازگار هستند**[4]. گاهی با پیشامد ساده‌ی s به عنوان پیشامد {s} رفتار می‌کنیم. بنا به تعریف، تمام پیشامدهای ساده، دو به دو ناسازگار هستند.

اصول موضوعی احتمال (axioms of probability)

توزیع احتمال[5] Pr{} روی فضای نمونه‌ی S ، نگاشتی از پیشامدهای S به اعداد حقیقی است، به طوری که در **اصول موضوعی احتمال** صدق می‌کند:

۱. برای هر پیشامد A داریم $\Pr\{A\} \geq 0$.

۲. $\Pr\{S\} = 1$.

۳. برای هر دو پیشامد دو به دو ناسازگار A و B ، داریم $\Pr\{A \cup B\} = \Pr\{A\} + \Pr\{B\}$. به طور کلی‌تر، برای هر دنباله‌ی (متناهی یا نامتناهی شمارش‌پذیر) از پیشامدهای $A_1, A_2, \ldots$ که دو به دو ناسازگار هستند، 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 موضوعی ۲، یک الزام **نرمال‌سازی**[6] است: نکته‌ی خاصی راجع‌به انتخاب ۱ به عنوان احتمال پیشامد **حتمی** وجود ندارد، مگر این‌که طبیعی و مناسب است.

از این اصول و نظریه‌ی مقدماتی مجموعه‌ای، 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:

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

$$\leq \Pr\{A\} + \Pr\{B\}. \qquad \text{(پ-۱۵)}$$

در مثال پرتاب سکه، فرض کنید هر یک از چهار پیشامد ساده دارای احتمال 1/4 باشند. آن‌گاه احتمال به دست آوردن حداقل یک شیر به صورت زیر است:

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

از طرف دیگر، چون احتمال به دست آوردن اکیداً کمتر از یک شیر، برابر با $\Pr\{\mathrm{TT}\} = 1/4$ 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 بر روی فضای نمونه‌ی متناهی یا نامتناهی شمارش‌پذیر تعریف شده باشد. فرض کنید S یک فضای نمونه باشد. آن‌گاه برای هر پیشامد A ، داریم:

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

زیرا پیشامدهای ساده، به خصوص آن‌هایی که در A هستند، دو به دو ناسازگار هستند. اگر S متناهی و هر پیشامد ساده‌ی $s \in S$ دارای احتمال زیر باشد، آن‌گاه **توزیع احتمال یکنواخت** بر 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 **سکه‌ی سالم** را در نظر بگیرید 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 بر روی فضای نمونه‌ی $S = \{H, T\}^n$ را به دست می‌آوریم که مجموعه‌ای به اندازه‌ی 2^n است. هر پیشامد ساده در S می‌تواند به صورت رشته‌ای به طول n بر روی {H, T} نمایش داده شود و هر کدام با احتمال $1/2^n$ رخ دهد. پیشامد زیر، زیرمجموعه‌ای از S با اندازه‌ی $|A| = \binom{n}{k}$ است، زیرا $\binom{n}{k}$ رشته به طول n بر روی {H, T} وجود دارند که دقیقاً k عدد H دارند:

$$A = \{\text{دقیقاً } k \text{ شیر و دقیقاً } n-k \text{ خط رخ می‌دهد}\}$$

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

توزیع احتمال یکنواخت پیوسته

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

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

با قرار دادن $c = d$ احتمال یک نقطه برابر با صفر است. اگر نقاط انتهایی [c, c] و [d, d] دو سرِ فاصله‌ی [c, d] را حذف کنیم، فاصله‌ی بازِ (c, d) را به دست می‌آوریم. چون $[c, d] = [c,c] \cup (c,d) \cup [d,d]$ ، 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 نمونه‌ی [a, b] است که می‌تواند با اجتماع متناهی یا شمارش‌پذیر فاصله‌های باز و بسته و چند مجموعه‌ی پیچیده‌تر به دست آید.

احتمال شرطی و استقلال

گاهی، از قبل، اطلاعاتی جزیی درباره‌ی نتیجه‌ی یک آزمایش داریم. برای مثال، فرض کنید دوستی دو سکه‌ی 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 آمدن دو خط را حذف کرده است. سه پیشامد ساده‌ی باقیمانده، احتمال یکسانی دارند، لذا نتیجه می‌گیریم که هر کدام با احتمال 1/3 رخ می‌دهند. چون فقط یکی از این پیشامدهای ساده دو شیر را نشان می‌دهد، پاسخ پرسش ما 1/3 است.

احتمال شرطی، مفهوم اطلاعات جزیی قبلی درباره‌ی نتیجه‌ی یک آزمایش را فرمول‌بندی می‌کند. احتمال شرطیِ این‌که پیشامد A موجب پیشامد B شود، وقتی $\Pr\{B\} \neq 0$ باشد، به صورت زیر است:

$$\Pr\{A \mid B\} = \frac{\Pr\{A \cap B\}}{\Pr\{B\}} \qquad \text{(پ-۱۶)}$$

(عبارت $\Pr\{A \mid B\}$ را به صورت **احتمال پیشامد A به شرط وقوع پیشامد B** بخوانید). از نظر شهودی، بنا به معادله (پ-۱۶) چون وقوع پیشامد B داده شده است، پیشامدی که A نیز رخ دهد برابر با $A \cap B$ است، یعنی، $A \cap B$ مجموعه‌ای از نتایج است که در آن، هم A و هم B رخ می‌دهند. چون این نتیجه، یکی از پیشامدهای ساده در B است، احتمالات تمام پیشامدهای ساده در B را با تقسیم آن‌ها بر $\Pr\{B\}$ نرمال‌سازی می‌کنیم، به طوری که مجموع آن‌ها برابر با یک شود. بنابراین، احتمال شرطی A با معلوم بودن B ، برابر با نسبت احتمال پیشامد $A \cap B$ به احتمال پیشامد B است. در مثال فوق، A پیشامدی است که در آن هر دو سکه شیر آمده باشد و B پیشامدی است که حداقل یکی از سکه‌ها شیر آمده باشد. بنابراین $\Pr\{A \mid B\} = (1/4)/(3/4) = 1/3$.

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

$$\Pr\{A \cap B\} = \Pr\{A\}\Pr\{B\}, \qquad \text{(پ-۱۷)}$$

که اگر $\Pr\{B\} \neq 0$ باشد، با شرط زیر هم‌ارز است:

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

برای مثال، فرض کنید دو سکه‌ی سالم پرتاب شدند و نتایج مستقل هستند. آن‌گاه احتمال آمدن دو شیر برابر است با $(1/2)(1/2) = 1/4$. اکنون فرض کنید یک پیشامد این‌است‌که **سکه‌ی 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، 1/4 است. بنابراین، بر اساس تعریف استقلال، این پیشامدها مستقل هستند، گرچه ممکن‌است تصور کنید که هر دو پیشامد به سکه‌ی اول بستگی دارند. سرانجام، فرض کنید که سکه‌ها با هم جور می‌شوند، 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 $1/2 \neq (1/2)(1/2)$ است. در نتیجه، پیشامد این‌که یکی شیر بیاید و پیشامدی که دیگری شیر بیاید، مستقل نیستند.

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

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

1. pairwise independent

می‌گوییم رویدادهای این کلکسیون (**متقابلاً مستقل**[1] **هستند** اگر هر زیرمجموعه‌ی k عضوی، $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$ از این کلکسیون که $2 \le k \le n$ و $1 \le i_1 < i_2 < \ldots < i_k \le n$، در رابطه‌ی 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 دو سکه‌ی سالم را پرتاب می‌کنیم. فرض کنید A_1 پیشامدی باشد که **سکه‌ی اول شیر می‌آید** و A_2 پیشامدی باشد که **سکه‌ی دوم شیر می‌آید** و A_3 پیشامدی باشد که **دو سکه** حالت‌های متفاوت داشته باشند. 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 مستقل هستند. پیشامدها متقابلاً مستقل نیستند، زیرا $\Pr\{A_1 \cap A_2 \cap A_3\} = 0$ و $\Pr\{A_1\}\Pr\{A_2\}\Pr\{A_3\} = 1/8 \neq 0$.

قضیه‌ی بیز[2]

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:

$$\begin{aligned}
\Pr\{A \cap B\} &= \Pr\{B\}\Pr\{A \mid B\}\\
&= \Pr\{A\}\Pr\{B \mid A\}.
\end{aligned} \qquad \text{(پ-۱۸)}$$

با حل آن نسبت به $\Pr\{A \mid B\}$، داریم:

$$\Pr\{A \mid B\} = \frac{\Pr\{A\}\Pr\{B \mid A\}}{\Pr\{B\}}, \qquad \text{(پ-۱۹)}$$

که **قضیه‌ی بیز** نام دارد. مخرج $\Pr\{B\}$ ثابت نرمال‌کننده‌ای است که می‌تواند به صورت زیر بیان شود. چون $B = (B \cap A) \cup (B \cap \bar{A})$ و $B \cap A$ و $B \cap \bar{A}$ پیشامدهایی دو به دو ناسازگار هستند، داریم:

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

با جایگزینی در معادله‌ی (پ-۱۹)، شکل هم‌ارز قضیه‌ی بیز را به دست می‌آوریم:

$$\Pr\{A \mid B\} = \frac{\Pr\{A\}\Pr\{B \mid A\}}{\Pr\{A\}\Pr\{B \mid A\} + \Pr\{\bar{A}\}\Pr\{B \mid \bar{A}\}}. \qquad \text{(پ-۲۰)}$$

قضیه‌ی بیز می‌تواند محاسبه‌ی احتمالات شرطی را ساده نماید. برای مثال، فرض کنید یک سکه‌ی سالم و یک سکه‌ی ناقص داریم که همیشه شیر می‌آید. آزمایشی انجام می‌دهیم که شامل سه پیشامد مستقل باشد:

1. (mutually) Independent 2. Bayes's theorem

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

این مسأله را با استفاده از قضیه‌ی بیز حل می‌کنیم. فرض کنید A پیشامد انتخاب سکه‌ی 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:

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

تمرین‌ها

تمرین پ-۲-۱: پروفسور روزنکرانتس (Rosencrantz) یک بار سکه‌ی سالمی را می‌اندازد. پروفسور گیلدنسترن (Guildenstern) سکه‌ی سالمی را دو بار می‌اندازد. احتمال این‌که پروفسور روزنکرانتس نسبت به پروفسور گیلدنسترن شیرهای بیشتری به دست آورد چقدر است؟

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

$$\Pr\{A_1 \cup A_2 \cup \cdots\} \le \Pr\{A_1\} + \Pr\{A_2\} + \cdots. \qquad \text{(پ-۲۱)}$$

تمرین پ-۲-۳: دسته‌ای از ۱۰ کارت را که از ۱ تا ۱۰ شماره‌گذاری شدند، بُر می‌زنیم. 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: ثابت کنید:

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

تمرین پ-۲-۵: ثابت کنید برای هر کلکسیون از رویدادهای $A_1, A_2, \ldots, A_n$ داریم:

$$\begin{aligned}\Pr\{A_1 \cap A_2 \cap \cdots \cap A_n\} = \Pr\{A_1\}\cdot\Pr\{A_2 \mid A_1\}\cdot\Pr\{A_3 \mid A_1 \cap A_2\}\cdots\\ \Pr\{A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1}\}.\end{aligned} \qquad \text{(پ-۲۲)}$$

★ **تمرین پ-۲-۶:** نشان دهید که چگونه می‌توان مجموعه‌ای از n رویداد را ساخت که دو به دو مستقل باشند، اما هیچ زیرمجموعه‌ی $k > 2$ از آن‌ها، **متقابلاً مستقل** نباشند.

★ **تمرین پ-۲-۷:** دو پیشامد A و B **مستقل شرطی هستند**، 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 شرطی هستند.

★ **تمرین پ-۲-۸:** پروفسور گور یک کلاس موسیقی درباره‌ی ریتم تدریس می‌کند که در آن سه دانش‌آموزِ جف، تیم و کارمین در معرض مردودی قرار دارند. پروفسور گور به این سه دانش‌آموز می‌گوید یکی از آن‌ها قبول و دو نفر دیگر مردود می‌شوند. کارمین به طور خصوصی از پروفسور می‌پرسد کدام یک از جف و تیم مردود می‌شوند، با این استدلال که می‌داند حداقل یکی از آن‌ها مردود می‌شود. پروفسور هیچ اطلاعاتی درباره‌ی نتیجه‌ی کارمین فاش نمی‌کند. در نقض حریم خصوصی، پروفسور گور به کارمین می‌گوید که جف مردود می‌شود. کارمین اکنون تا حدودی احساس آرامش می‌کند و فکر می‌کند که او یا تیم قبول می‌شوند، به طوری که احتمال قبولی او 1/2 است. آیا کارمین درست فکر می‌کند یا هنوز شانس قبولی او 1/3 است.

پ–۳ متغیرهای تصادفی گسسته

متغیر تصادفی (گسسته‌ی) X ، تابعی از یک فضای نمونه‌ی نامتناهی یا متناهی شمارش‌پذیر 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 بر روی مجموعه‌ای از اعداد حاصل کار کنیم. متغیرهای تصادفی را می‌توان برای فضاهای نمونه‌ی نامتناهی شمارش‌ناپذیر نیز تعریف کرد، اما مشکلات تکنیکی را به وجود می‌آورند که لازم نیست برای اهداف ما، برطرف شوند. بنابراین، فرض خواهیم کرد که متغیرهای تصادفی گسسته هستند.

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

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

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

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

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

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

معمولاً چندین متغیر تصادفی را بر روی یک فضای نمونه تعریف می‌کنیم. اگر X و Y متغیرهای تصادفی باشند، تابع زیر، **تابع چگالی احتمال مشترک**[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:

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

با استفاده از تعریف (پ–۱۶) از احتمال شرطی، داریم:

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

بنا به تعریف، دو متغیر تصادفی X و Y **مستقل** هستند 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 بر روی یک فضای نمونه، می‌توان متغیرهای تصادفی را به عنوان مجموع، حاصل‌ضرب، یا توابع دیگری از متغیرهای اصلی تعریف کرد.

1. Probability density function 2. Joint probability density function

امید ریاضی یک متغیر تصادفی

ساده‌ترین و مفیدترین بیان از توزیع متغیر تصادفی، **میانگین** مقادیری است که آن متغیر اختیار می‌کند. **مقدار امید ریاضی**[1] (یا مترادف آن، **امید ریاضی** یا **میانگین**) متغیر تصادفی گسسته به صورت زیر است:

$$\mathrm{E}[X] = \sum_x x \cdot \Pr\{X = x\}, \qquad (\text{پ-۲۳})$$

که اگر مجموع آن متناهی باشد یا مطلقاً همگرا باشد، **خوش‌تعریف** است. گاهی امید ریاضی X به صورت μ_X نوشته می‌شود و وقتی که متغیر تصادفی حذف می‌شود، به صورت μ نمایش داده خواهد شد.

یک بازی را در نظر بگیرید که در آن، دو سکه‌ی سالم را پرتاب می‌کنید. برای هر شیر ۳ دلار می‌گیرید ولی برای هر خط ۲ دلار از دست می‌دهید. مقدار امید ریاضی متغیر تصادفی X که دریافتی شما را نشان می‌دهد، به صورت زیر است:

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

ویژگی خطی بودن امید ریاضی بیان می‌کند که مجموع دو متغیر تصادفی، برابر با مجموع امیدهای ریاضی آن‌ها است، یعنی هر وقت E[X] و E[Y] تعریف می‌شوند، داریم:

$$\mathrm{E}[X + Y] = \mathrm{E}[X] + \mathrm{E}[Y], \qquad (\text{پ-۲۴})$$

ویژگی **خطی‌بودن امید ریاضی**، حتی اگر X و Y مستقل نباشند، برقرار است. علاوه‌براین، به مجموع‌های متناهی و مطلقاً همگرایی از امیدهای ریاضی بسط داده می‌شود. خطی‌بودن امید ریاضی، ویژگی مهمی است که ما را قادر می‌سازد تحلیل احتمالی را با استفاده از متغیرهای تصادفی شاخص انجام دهیم (بخش ۵-۲).

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

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

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

$$\mathrm{E}[aX] = a\mathrm{E}[X]. \qquad (\text{پ-۲۵})$$

در نتیجه، امیدهای ریاضی خطی هستند: برای هر دو متغیر تصادفی X و Y و ثابت a ، داریم:

$$\mathrm{E}[aX + Y] = a\mathrm{E}[X] + \mathrm{E}[Y]. \qquad (\text{پ-۲۶})$$

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

$$\begin{aligned}\mathrm{E}[XY] &= \sum_x \sum_y xy \cdot \Pr\{X = x \text{ and } Y = y\} \\ &= \sum_x \sum_y xy \cdot \Pr\{X = x\} \Pr\{Y = y\} && \text{(بر اساس استقلال X و Y)} \\ &= \left(\sum_x x \cdot \Pr\{X = x\}\right)\left(\sum_y y \cdot \Pr\{Y = y\}\right) \\ &= \mathrm{E}[X]\mathrm{E}[Y] . && \text{(بر اساس معادله‌ی (پ-۲۳))}\end{aligned}$$

1. Expected value

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

$$\mathrm{E}[X_1 X_2 \cdots X_n] = \mathrm{E}[X_1]\,\mathrm{E}[X_2] \cdots \mathrm{E}[X_n]. \qquad \text{(پ-۲۷)}$$

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

$$\begin{aligned}
\mathrm{E}[X] &= \sum_{i=0}^{\infty} i \cdot \Pr\{X = i\} \\
&= \sum_{i=0}^{\infty} i\,(\Pr\{X \geq i\} - \Pr\{X \geq i+1\}) \\
&= \sum_{i=1}^{\infty} \Pr\{X \geq i\}\ , \qquad \text{(پ-۲۸)}
\end{aligned}$$

زیرا هر جمله‌ی $\Pr\{X \geq i\}$ به تعداد i بار جمع و $i-1$ بار کم می‌شود (به جز $\Pr\{X \geq 0\}$ که صفر بار جمع می‌شود ولی اصلاً کم نمی‌شود).

تابع f(x) محدب است اگر برای تمام x و y و تمام $0 \leq \lambda \leq 1$ داشته باشیم:

$$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) \ + (1-\lambda) f(y) \qquad \text{(پ-۲۹)}$$

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

$$\mathrm{E}[f(X)] \geq f(\mathrm{E}[X]), \qquad \text{(پ-۳۰)}$$

به شرطی که امیدهای ریاضی وجود داشته باشند و متناهی باشند.

واریانس و انحراف معیار

امید ریاضی یک متغیر تصادفی، به ما نمی‌گوید که مقادیر متغیر چگونه توزیع (پخش) شدند. برای مثال اگر متغیرهای تصادفی X و Y را داشته باشیم که برای آن‌ها:

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

آن‌گاه هم E[X] و هم E[Y] برابر با 1/2 هستند، اما مقادیر واقعی که Y اختیار می‌کند متفاوت از میانگین مقادیر واقعی است که X اختیار می‌کند.

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

$$\begin{aligned}
\mathrm{Var}[X] &= \mathrm{E}[(X - \mathrm{E}[X])^2] \\
&= \mathrm{E}[X^2 - 2X\mathrm{E}[X] + \mathrm{E}^2[X]] \\
&= \mathrm{E}[X^2] - 2\mathrm{E}[X\mathrm{E}[X]] + \mathrm{E}^2[X] \\
&= \mathrm{E}[X^2] - 2\mathrm{E}^2[X] + \mathrm{E}^2[X] \\
&= \mathrm{E}[X^2] - \mathrm{E}^2[X]. \qquad \text{(پ-۳۱)}
\end{aligned}$$

برای توجیه معادله‌ی $\mathrm{E}[\mathrm{E}^2[X]] = \mathrm{E}^2[X]$ توجه کنید که چون E[X] یک عدد حقیقی است، نه یک متغیر تصادفی، $\mathrm{E}^2[X]$ نیز چنین است. معادله‌ی $\mathrm{E}[X\mathrm{E}[X]] = \mathrm{E}^2[X]$ با فرض $a = \mathrm{E}[X]$ از معادله‌ی (پ-۲۵) نتیجه می‌شود. معادله‌ی (پ-۳۱) می‌تواند بازنویسی شود تا عبارتی برای امید ریاضی مربع متغیر تصادفی به دست آید:

$$\mathrm{E}[X^2] = \mathrm{Var}[X] + \mathrm{E}^2[X]. \qquad \text{(پ-۳۲)}$$

واریانس متغیر تصادفی X و واریانس aX به هم وابسته هستند (تمرین پ-۳-۱۰ را ببینید):

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

وقتی X و Y متغیرهای تصادفی مستقل باشند، 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:

$$\mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{Var}[X_i]. \qquad (\text{پ-۳۳})$$

انحراف معیار متغیر تصادفی X ، ریشه‌ی دوم نامنفی واریانس X است. انحراف معیار متغیر تصادفی X ، گاهی با σ_X یا در صورتی که متغیر X از متن فهمیده شود، با σ نمایش داده می‌شود. با این نمادگذاری، واریانس X با σ^2 نمایش داده می‌شود.

تمرین‌ها

تمرین پ-۳-۱: دو تاس ۶ وجهی پرتاب می‌شوند. امید ریاضی مجموع دو مقداری که بر روی تاس‌ها نشان داده می‌شوند چیست؟ امید ریاضی بیشینه‌ی این دو مقدار چیست؟

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

تمرین پ-۳-۳: یک بازی کارناوال شامل سه تاس در یک قفسه است. بازیکن می‌تواند یک دلار را برای هر یک از شماره‌های ۱ تا ۶ شرطبندی کند. قفسه تکان داده می‌شود و تصفیه به صورت زیر است. اگر شماره‌ی بازیکن در هیچ‌کدام از تاس‌ها ظاهر نشود، دلار خود را از دست می‌دهد. وگرنه، اگر شماره‌ی 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. امید ریاضی حاصل از یک بار بُرد 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 مستقل هستند.

★ **تمرین پ-۳-۶:** فرض کنید X متغیر تصادفی نامنفی و E[X] خوش‌تعریف باشد. ثابت کنید **نامعادله‌ی مارکوف**[1] برای $t > 0$ برقرار است:

$$\Pr\{X \ge t\} \le \mathrm{E}[X]/t \qquad (\text{پ-۳۴})$$

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

$$\Pr\{X \ge t\} \ge \Pr\{X' \ge t\}.$$

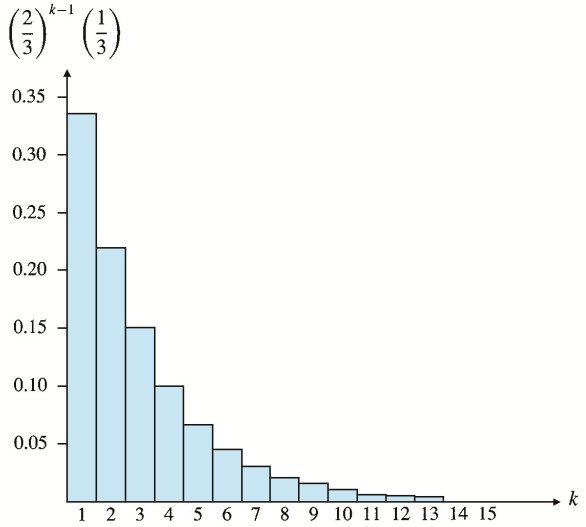
تمرین پ-۳-۸: کدام بزرگ‌تر است: امید ریاضی مربع متغیر تصادفی، یا مربع امید ریاضی آن؟

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-۳۱) درباره‌ی واریانس، ثابت کنید $\mathrm{Var}[aX] = a^2\mathrm{Var}[X]$.

1. Markov's inequality



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

پ–۴ توزیع‌های هندسی و دوجمله‌ای

پرتاب سکه را می‌توانیم نمونه‌ای از **آزمایش برنولی** در نظر بگیریم که آزمایشی فقط با دو نتیجه‌ی ممکن است: **موفقیت** که با احتمال p رخ می‌دهد و **شکست** که با احتمال $q = 1 - p$ رخ می‌دهد. وقتی راجع به آزمایش‌های برنولی صحبت می‌کنیم، معنایش این است که آزمایش‌ها **متقابلاً مستقل هستند**، مگر این‌که غیر از این بیان کنیم که هر کدام، برای موفقیت، احتمال یکسان p دارند، دو توزیع مهم 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 دست یافتن به یک موفقیت، چند آزمایش انجام می‌شود؟ فرض کنید متغیر تصادفی X برابر با تعداد آزمایش‌های موردنیاز برای کسب موفقیت باشد. سپس X دارای مقادیری در بازه‌ی $\{1, 2, \ldots\}$ است و برای $k \geq 1$ داریم:

$$\Pr\{X = k\} = q^{k-1} p\,, \qquad (پ–۳۵)$$

زیرا قبل از یک موفقیت، $k - 1$ شکست داریم. توزیع احتمالی که در معادله‌ی (پ–۳۵) صدق می‌کند، **توزیع هندسی** نام دارد. شکل پ – ۱ این توزیع را نشان می‌دهد.

با فرض این‌که $q < 1$، امید ریاضی توزیع هندسی می‌تواند به صورت زیر محاسبه شود:

$$\begin{aligned}
\mathrm{E}[X] &= \sum_{k=1}^{\infty} k q^{k-1} p \\
&= \frac{p}{q} \sum_{k=0}^{\infty} k q^{k} \\
&= \frac{p}{q} \cdot \frac{q}{(1-q)^2} \quad \text{(بنا به معادله‌ی (الف–۱۱))} \\
&= \frac{p}{q} \cdot \frac{q}{p^2} \\
&= 1/p\,. \qquad (پ–۳۶)
\end{aligned}$$

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 پ-۴-۳)، به صورت زیر است:

$$\mathrm{Var}[X] = q/p^2. \qquad (\text{پ-۳۷})$$

به عنوان مثال، فرض کنید مکرراً دو تاس را می‌اندازیم تا یک ۷ یا یک ۱۱ را به دست آوریم. از ۳۶ نتیجه‌ی ممکن، ۶ بار منجر به ۷ و ۲ بار 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 یک ۷ یا یک ۱۱ را به دست آوریم.

توزیع دوجمله‌ای

در حین n آزمایش برنولی چند موفقیت رخ می‌دهد که در آن، 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 در نظر بگیرید. سپس، X دارای مقادیری در بازه‌ی $\{0, 1, \ldots, n\}$ است و برای $k = 0, 1, \ldots, n$ داریم:

$$\Pr\{X = k\} = \binom{n}{k} p^k q^{n-k}, \qquad (\text{پ-۳۸})$$

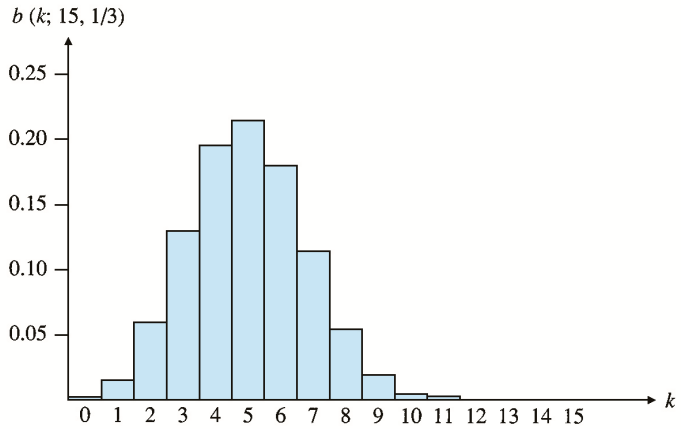
زیرا $\binom{n}{k}$ روش برای انتخاب k موفقیت از n آزمایش وجود دارد و احتمال این‌که هر کدام رخ دهند، $p^k q^{n-k}$ است. توزیع احتمالی که در معادله‌ی (پ-۳۸) صدق می‌کند، **توزیع دوجمله‌ای** نامیده می‌شود. برای سهولت، خانواده‌ی توزیع دوجمله‌ای را با استفاده از نمادگذاری زیر تعریف می‌کنیم:

$$b(k; n, p) = \binom{n}{k} p^k (1 - p)^{n-k}. \qquad (\text{پ-۳۹})$$

شکل پ-۲ توزیع دوجمله‌ای را شرح می‌دهد. نام **دوجمله‌ای** از این حقیقت ناشی می‌شود که (پ-۳۸)، جمله‌ی k اُم بسط $(p+q)^n$ است. در نتیجه، چون $p + q = 1$، بنا به معادله‌ی (پ-۴) داریم:

$$\sum_{k=0}^{n} b(k; n, p) = 1, \qquad (\text{پ-۴۰})$$

که اصل ۲ از اصول موضوعی احتمال به آن نیاز دارد.



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.

می‌توان امید ریاضی متغیر تصادفیِ دارای توزیع دوجمله‌ای را از معادلات (پ-۹) و (پ-۴۰) 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 امید ریاضی، داریم:

$$\begin{aligned}
\mathrm{E}[X] &= \sum_{k=0}^{n} k \cdot \Pr\{X = k\} \\
&= \sum_{k=0}^{n} k \cdot b(k; n, p) \\
&= \sum_{k=1}^{n} k \binom{n}{k} p^k q^{n-k} \\
&= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} q^{n-k} \quad \text{(بنا به معادله‌ی (پ-۹))} \\
&= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k q^{(n-1)-k} \\
&= np \sum_{k=0}^{n-1} b(k; n-1, p) \\
&= np \quad \text{(بنا به معادله‌ی (پ-۴۰))}
\end{aligned} \tag{پ-۴۱}$$

با استفاده از خاصیت خطی‌بودن امید ریاضی، می‌توان نتیجه‌ی 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 می‌کند. سپس $E[X_i] = p \,.\, 1 + q \,.\, 0 = p$ و بنا به خاصیت خطی بودن امید ریاضی (معادله‌ی (پ-۲۱))، امید ریاضی تعداد موفقیت‌ها برای n آزمایش برابر است با:

$$\begin{aligned}
\mathrm{E}[X] &= \mathrm{E}\left[\sum_{i=1}^{n} X_i\right] \\
&= \sum_{i=1}^{n} \mathrm{E}[X_i] \quad \text{(بنا به معادله‌ی (پ-۲۴))} \\
&= \sum_{i=1}^{n} p \\
&= np\,.
\end{aligned} \tag{پ-۴۲}$$

همین روش را می‌توانیم برای محاسبه‌ی واریانس توزیع به کار ببریم. با استفاده از معادله‌ی (پ-۳۱)، داریم $\mathrm{Var}[X_i] = \mathrm{E}[X_i^2] - \mathrm{E}^2[X_i]$. چون X_i فقط مقادیر 0 و 1 را می‌پذیرد، داریم $X_i^2 = X_i$ که نتیجه می‌شود $\mathrm{E}[X_i^2] = \mathrm{E}[X_i] = p$ و در نتیجه، داریم:

$$\mathrm{Var}[X_i] = p - p^2 = p(1-p) = pq\,. \tag{پ-۴۳}$$

برای محاسبه‌ی واریانس X، از امتیاز استقلال n آزمایش استفاده می‌کنیم. بنابراین، بنا به معادله‌ی (پ-۳۳) داریم:

$$\begin{aligned}
\mathrm{Var}[X] &= \mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] \\
&= \sum_{i=1}^{n} \mathrm{Var}[X_i] \\
&= \sum_{i=1}^{n} pq \\
&= npq\,.
\end{aligned} \tag{پ-۴۴}$$

همان‌طور که شکل پ-۲ نشان می‌دهد، توزیع دوجمله‌ای $b(k; n, p)$ با k افزایش می‌یابد تا به میانگین np برسد و سپس کاهش می‌یابد. با مشاهده‌ی نسبت جملات موفقیت‌آمیز، می‌توان اثبات کرد که این توزیع همیشه به این روش رفتار می‌کند:

$$\begin{aligned}\frac{b(k;n,p)}{b(k-1;n,p)} &= \frac{\binom{n}{k}p^k q^{n-k}}{\binom{n}{k-1}p^{k-1}q^{n-k+1}}\\ &= \frac{n!\,(k-1)!\,(n-k+1)!\,p}{k!\,(n-k)!\,n!\,q}\\ &= \frac{(n-k+1)p}{kq} \qquad \text{(پ-۴۵)}\\ &= 1+\frac{(n-k+1)p-kq}{kq}\\ &= 1+\frac{(n-k+1)p-k(1-p)}{kq}\\ &= 1+\frac{(n+1)p-k}{kq}\,.\end{aligned}$$

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 صعودی) و برای $k>(n+1)p$ داریم $b(k;n,p)<b(k-1;n,p)$ (توزیع نزولی). اگر $k=(n+1)p$ صحیح باشد، آن‌گاه $b(k;n,p)/b(k-1;n,p)$، لذا این توزیع دارای دو مقدار بیشینه است: در $k=(n+1)p$ و در $k-1=(n+1)p-1=np-q$. وگرنه، در مقدار صحیح یکتای k به بیشینه می‌رسد که k در بازه‌ی $np-q<k<(n+1)p$ است.

لِم زیر کران بالایی را بر روی توزیع دوجمله‌ای ایجاد می‌کند.

لِم پ-۱

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

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

اثبات: داریم:

$$\begin{aligned}b(k;n,p) &= \binom{n}{k}p^k q^{n-k}\\ &\le \left(\frac{n}{k}\right)^k\left(\frac{n}{n-k}\right)^{n-k}p^k q^{n-k} \quad \text{(بنا به نامعادله‌ی (پ-۷))}\\ &= \left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}.\end{aligned}$$

■

تمرین‌های بخش پ – ۴

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

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

تمرین پ-۴-۳: نشان دهید که واریانس توزیع هندسی برابر با q/p^2 است (راهنمایی: از تمرین (الف-۱-۶) استفاده کنید).

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

تمرین پ-۴-۵: نشان دهید که مقدار بیشینه توزیع دوجمله‌ای $b(k;n,p)$ تقریباً $1/\sqrt{2\pi npq}$ است که در آن $q=1-p$.

★**تمرین پ-۴-۶:** نشان دهید که احتمال **هیچ موفقیت** در آزمایش برنولی، هر کدام با احتمال $p=1/n$، تقریباً $1/e$ است. نشان دهید که احتمال دقیقاً یک موفقیت نیز تقریباً $1/e$ است.

★ **تمرین پ-۴-۷:** پروفسور روزنکرانتس (Rosencrantz) یک سکه‌ی سالم را n بار پرتاب می‌کند و پروفسور گیلدنسترن (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 روزنکرانتس، شیر را موفقیت در نظر بگیرید؛ برای گیلدنسترن، خط را موفقیت در نظر بگیرید). با استفاده از بحث خود، اتحاد زیر را اثبات کنید:

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

★ **تمرین پ-۴-۸:** نشان دهید که برای $0\le k\le n$ داریم:

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

که $H(x)$ تابع آنتروپی است (پ-۸).

★ **تمرین پ-۴-۹:** تعداد n آزمایش برنولی را در نظر بگیرید که برای $i=1,2,\dots,n$، آزمایش i اُم دارای احتمال موفقیت p_i است و فرض کنید X یک متغیر تصادفی باشد که تعداد کل موفقیت‌ها را نشان می‌دهد. فرض کنید برای $i=1,2,\dots,n$ داریم $p\ge p_i$. ثابت کنید برای $1\le k\le n$ داریم:

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

★ **تمرین پ-۴-۱۰:** فرض کنید X یک متغیر تصادفی برای تعداد کل موفقیت‌ها در مجموعه‌ی A از n آزمایش برنولی باشد، که در آن، احتمال موفقیت آزمایش i اُم برابر با p_i و فرض کنید X' متغیر تصادفی برای تعداد کل موفقیت‌ها در مجموعه‌ی دوم A' از n آزمایش برنولی باشد که آزمایش i اُم دارای احتمال موفقیت $p'_i\ge p_i$ است. ثابت کنید برای $0\le k\le n$ داریم:

$$\Pr\{X'\ge k\}\ge\Pr\{X\ge k\}.$$

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

★پ-۵ دُم‌های توزیع دوجمله‌ای[1]

احتمال به دست آوردن حداقل، یا حداکثر k موفقیت در آزمایش برنولی، هر یک با احتمال موفقیت p، غالباً جالب‌تَر از احتمال دقیقاً k موفقیت است. در این بخش، **دُم‌های**[2] توزیع دوجمله‌ای را بررسی می‌کنیم: دو ناحیه از توزیع $b(k;n,p)$ که از میانگین np خیلی دور هستند. چندین کران مهم (مجموع تمام جملات در) یک دُم را اثبات خواهیم کرد.

ابتدا کرانی را روی دُم راست توزیع $b(k;n,p)$ فراهم می‌سازیم. کران‌های دُم چپ می‌تواند با وارون کردن نقش‌های موفقیت‌ها و شکست‌ها تعیین شود.

قضیه‌ی پ - ۲

دنباله‌ای از آزمایش‌های برنولی را در نظر بگیرید که در آن موفقیت با احتمال p اتفاق می‌افتد. فرض کنید X متغیر تصادفی باشد که تعداد کل موفقیت‌ها را نشان می‌دهد. آن‌گاه برای $0\le k\le n$، احتمال حداقل k موفقیت برابر است با:

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

1. Tails of the binomial distribution 2. Tails

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

$$\begin{aligned}
\Pr\{X \geq k\} &= \Pr\{\text{there exists } S \subseteq \{1, 2, \ldots, n\} : |S| = k \text{ and } A_S\} \\
&= \Pr\left\{\bigcup_{S \subseteq \{1,2,\ldots,n\}:|S|=k} A_S\right\} \\
&\leq \sum_{S \subseteq \{1,2,\ldots,n\}:|S|=k} \Pr\{A_S\} \quad \text{(بنا به نامعادله‌ی (پ-۲۱))} \\
&= \binom{n}{k} p^k . \quad \blacksquare
\end{aligned}$$

نتیجه‌ی زیر، این قضیه را برای دُم سمت چپ توزیع دوجمله‌ای بیان می‌کند. به طور کلی، اثبات آن را بر عهده‌ی خواننده واگذار می‌کنیم.

نتیجه‌ی پ-۳

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

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

کران بعدی که تعیین می‌کنیم به دُم چپ توزیع دوجمله‌ای مربوط می‌شود. نتیجه‌ی آن نشان می‌دهد که دور از میانگین، دُم چپ به طور نمایی نزول می‌کند.

قضیه‌ی پ-۴

دنباله‌ای از n آزمایش برنولی را در نظر بگیرید که در آن 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 وسیله‌ی سری هندسی، با استفاده از تکنیک بخش (الف – ۲) کران‌دار می‌کنیم. برای $i = 1, 2, \ldots, k$ ، از معادله‌ی (پ-۴۵) داریم:

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

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 $(k-i)$ باراستفاده‌ی مکرر از این نامعادله، به ازای $0\le i<k$ ، خواهیم داشت:

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

و در نتیجه داریم:

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

نتیجه‌ی پ-۵

دنباله‌ای از n آزمایش برنولی را در نظر بگیرید که در آن 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، کوچک‌تر از نصف احتمالِ کمتر از $k+1$ موفقیت است.

اثبات: چون $k\le np/2$، داریم:

$$\begin{aligned} \frac{kq}{np-k} &\le \frac{(np/2)q}{np-(np/2)} \\ &= \frac{(np/2)q}{np/2} \\ &\le 1' \end{aligned} \qquad \text{(پ -۴۶)}$$

زیرا $q\le 1$. فرض کنید X متغیر تصادفی باشد که تعداد موفقیت‌ها را نشان می‌دهد. از قضیه‌ی پ - ۴ و نامعادله‌ی (پ-۴۶) نتیجه می‌شود 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:

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

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

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

نتیجه‌ی پ-۶

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

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

■

نتیجه‌ی پ-۷

دنباله‌ای از n آزمایش برنولی را در نظر بگیرید 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. ■

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

نتیجه‌ی پ-۸

دنباله‌ای از n آزمایش برنولی را در نظر بگیرید 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 می‌کند و قرار دهید $\mu=E[X]$. آن‌گاه، برای $r>\mu$ داریم:

$$\Pr\{X-\mu\ge r\}\le\left(\frac{\mu e}{r}\right)^r.$$

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

$$\Pr\{X-\mu\ge r\}=\Pr\{e^{\alpha(X-\mu)}\ge e^{\alpha r}\},\qquad (\text{پ-۴۷})$$

که α بعداً تعیین می‌شود. با استفاده از نامعادله‌ی مارکوف (پ-۳۴)، خواهیم داشت:

$$\Pr\{e^{\alpha(X-\mu)}\ge e^{\alpha r}\}\le E[e^{\alpha(X-\mu)}]\,e^{-\alpha r}.\qquad (\text{پ-۴۸})$$

قسمت عمده‌ی اثبات، مربوط به کراندار کردن $E[e^{\alpha(X-\mu)}]$ و جایگزینی مقدار مناسبی برای α در نامعادله‌ی (پ-۴۸) است. ابتدا، $E[e^{\alpha(X-\mu)}]$ را ارزیابی می‌کنیم. با استفاده از نمادگذاری بخش ۵-۲، 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 این آزمایش شکست بخورد. بنابراین، داریم:

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

و بنا به خاصیت خطی بودن امید ریاضی، 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 نتیجه می‌شود:

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

برای ارزیابی $\mathrm{E}[e^{\alpha(X-\mu)}]$، به جای $X-\mu$ عبارت بالا را قرار می‌دهیم و به دست می‌آوریم:

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

که از معادله‌ی (پ-۲۷) به دست می‌آید، زیرا **استقلال متقابل** متغیرهای تصادفی X_i، **استقلال متقابل** متغیرهای تصادفی $e^{\alpha(X_i - p_i)}$ را نتیجه می‌دهد (تمرین پ-۳-۵ را ببینید). بنا به تعریف امید ریاضی، داریم:

$$\begin{aligned}
\mathrm{E}\left[e^{\alpha(X_i - p_i)}\right] &= e^{\alpha(1-p_i)} p_i + e^{\alpha(0-p_i)} q_i \\
&= p_i e^{\alpha q_i} + q_i e^{-\alpha p_i} \\
&\le p_i e^{\alpha} + 1 \\
&\le \exp(p_i e^{\alpha}) ,
\end{aligned} \tag{پ-۴۹}$$

که در آن $\exp(x)$ تابع نمایی است: $\exp(x) = e^x$ [نامعادله‌ی (پ-۴۹) 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 نامعادله‌ی (۳-۱۴) به دست می‌آید]. در نتیجه داریم:

$$\begin{aligned}
\mathrm{E}\left[e^{\alpha(X-\mu)}\right] &= \prod_{i=1}^{n} \mathrm{E}\left[e^{\alpha(X_i - p_i)}\right] \\
&\le \prod_{i=1}^{n} \exp(p_i e^{\alpha}) \\
&= \exp\left(\sum_{i=1}^{n} p_i e^{\alpha}\right) \\
&= \exp(\mu e^{\alpha}) ,
\end{aligned} \tag{پ-۵۰}$$

زیرا $\mu = \sum_{i=1}^{n} p_i$. بنابراین، از معادله‌ی (پ-۴۷) و نامعادلات (پ-۴۸) و (پ-۵۰) نتیجه می‌شود که:

$$\Pr\{X - \mu \ge r\} \le \exp(\mu e^{\alpha} - \alpha r). \tag{پ-۵۱}$$

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

$$\begin{aligned}\Pr\{X-\mu \geq r\} &\leq \exp(\mu e^{\ln(r/\mu)} - r\ln(r/\mu)) \\ &= \exp(r - r\ln(r/\mu)) \\ &= \frac{e^r}{(r/\mu)^r} \\ &= \left(\frac{\mu e}{r}\right)^r . \end{aligned} \quad \blacksquare$$

هنگام استفاده از آزمایش‌های برنولی که در آن، هر آزمایش احتمال موفقیت یکسانی دارد، قضیه‌ی پ-۸ منجر به نتیجه‌ی زیر می‌شود که دُم سمت راست توزیع دوجمله‌ای را محدود می‌کند.

نتیجه‌ی پ-۹

دنباله‌ای از n آزمایش برنولی را در نظر بگیرید که احتمال موفقیت آن 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 معادله‌ی (پ-۴۱)، داریم $\mu = E[X] = np$. ■

تمرین‌ها

★**تمرین پ-۵-۱:** احتمال کدام کمتر است: به وجود نیامدن هیچ شیر وقتی که سکه‌ی 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:** نتیجه‌های پ-۶ و پ-۷ را اثبات کنید.

★**تمرین پ-۵-۳:** برای تمام $a > 0$ و k به طوری که $0 < k < na/(a+1)$ 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 از شرایط قضیه‌ی پ-۸ برای $r > n - \mu$ نتیجه می‌شود که:

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

به طور مشابه، نشان دهید که از شرایط نتیجه‌ی پ-۹ برای $r > n - np$ نتیجه می‌شود که:

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

★**تمرین پ-۵-۶:** دنباله‌ای از n آزمایش برنولی را در نظر بگیرید 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 $p_i e^{\alpha q_i} + q_i e^{-\alpha p_i} \le e^{\alpha^2/2}$. سپس از طرح اثبات قضیه‌ی پ-۸ پیروی کنید، به طوری که به جای نامعادله‌ی (پ-۴۹) از این نامعادله استفاده کنید).

★**تمرین پ-۵-۷:** نشان دهید که سمت راست نامعادله‌ی (پ-۵۱) با انتخاب $\alpha = \ln(r/\mu)$ کمینه شود.

پ-۶ مسأله‌ها

مسأله‌ی پ-۱: توپ‌ها و سبدها.

در این مسأله، اثر فرض‌های مختلف بر روی تعداد روش‌های قراردادن n توپ در b سبد مختلف را بررسی می‌کنیم.

الف. فرض کنید n توپ متمایز هستند و ترتیب آن‌ها در یک سکه، مهم نیست. ثابت کنید تعداد روش‌های قراردادن توپ‌ها در سبدها، b^n است.

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

پ. فرض کنید توپ‌ها یکسان هستند و در نتیجه ترتیب آن‌ها در داخل سبد مهم نیست. نشان دهید که تعداد روش‌های قراردادن توپ‌ها در سبدها 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 سبدی 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 سبدی نباید خالی باشد. نشان دهید که تعداد روش‌های قراردادن توپ‌ها برابر با $\binom{n-1}{b-1}$ است.