به نام خدا

۱۳۰۷
دانشگاه صنعتی خواجه نصیرالدین طوسی

جزوه درس فرایندهای تصادفی

راضیه خودسیانی

فهرست مطالب

فصل ۱

مروری بر احتمال

همانطور که در درس مبانی احتمال دیدیم، نظریه احتمال یک مدل ریاضی برای مطالعه رخدادهای تصادفی (آزمایش های تصادفی) ارائه می کند. طبق تعریف، آزمایش تصادفی عملی با ویژگی های زیر است:

۱. قابلیت تکرار داشته باشد،

۲. در شرایط یکسان نتایج متفاوت داشته، و

۳. همه نتایج قبل از انجام آزمایش قابل پیش بینی باشد.

همچنین مجموعه تمام نتایج ممکن در یک آزمایش تصادفی را فضای نمونه گویند که آن را با نماد S یا Ω نمایش می دهند.

به هریک از عضوهای یک فضای نمونه که نتیجه ای از نتایج ممکن آزمایش تصادفی است، یک برآمد گفته می شود که آن را با ω نشان می دهیم.

هر زیرمجموعه از فضای نمونه را یک پیشامد گویند. پیشامدها را با حروف انگلیسی بزرگ نمایش می‌دهیم. هرگاه یک آزمایش تصادفی انجام شود نتیجه مشاهده شده در هر پیشامدی که باشد اصطلاحا می‌گوییم آن پیشامد رخ داده است. مثلا وقتی در آزمایش پرتاب تاس عدد ۲ مشاهده شود، پیشامد زوج بودن و پیشامد اول بودن عدد رو شده رخ داده است چون ۲ عضو هر دوی این پیشامدها است.

مهم ترین مساله در مورد یک آزمایش تصادفی، رخ دادن و یا ندادن پیشامدهای مورد توجه مرتبط با آن آزمایش تصادفی است.

انواع پیشامد:

- ساده: یک پیشامد تک عضوی است.
- مرکب: پیشامدی با تعداد اعضای بیشتر از یک است.
- محال یا تهی: پیشامد بدون عضو است که آن را به صورت $\phi = \{\}$ نمایش می دهند.
- حتمی: پیشامدی برابر با فضای نمونه است.

مثال ۱-۱. آزمایش پرتاب یک تاس را درنظر بگیرید. داریم $\Omega = \{۱, ۲, ۳, ۴, ۵, ۶\}$ و بنابراین $\omega_۱ = ۱$، $\omega_۲ = ۲$، ... و $\omega_۶ = ۶$ برآمدهای این آزمایش هستند. اگر A و B را به ترتیب پیشامد مشاهده عدد فرد و پیشامد مشاهده مضرب ۳ تعریف کنیم داریم

$$A = \{۱, ۳, ۵\} \qquad B = \{۳, ۶\}$$

پس اگر تاس را پرتاب کنیم و عدد ۳ مشاهده شود پیشامد A و B هر دو اتفاق افتاده اند. در حالیکه اگر عدد رو شده تاس ۵ باشد یعنی پیشامد A اتفاق افتاده اما B اتفاق نیافتاده است.

با استفاده از اعمال مجموعه ای می توان پیشامدهای جدید بدست آورد، برای مثال:

- پیشامد $A \cup B$ وقتی رخ می دهد که حداقل یکی از دو پیشامد A یا B رخ داده باشد و داریم
  $A \cup B = \{x \ : \ x \in A \ \text{ or } \ x \in B\}$
- پیشامد $A \cap B$ وقتی رخ می دهد که هر دو پیشامد A و B رخ داده باشد و داریم
  $A \cap B = \{x \ : \ x \in A \ \text{ and } \ x \in B\}$
- پیشامد $\bigcup_{n\in I} A_n$ وقتی رخ می دهد که حداقل یکی از پیشامدهای A_n رخ داده باشد و داریم
  $\bigcup_{n\in I} A_n = \{x \ : \ \exists n \in I \quad x \in A_n\}$
- پیشامد $\bigcap_{n\in I} A_n$ وقتی رخ می دهد که همه پیشامدهای A_n رخ داده باشد و داریم
  $\bigcap_{n\in I} A_n = \{x \ : \ \forall n \quad x \in A_n\}$
- پیشامد $A^c = A' = \{x \ : \ x \notin A\}$ وقتی رخ می دهد که پیشامد A رخ نداده باشد.
- پیشامد $A - B$ وقتی رخ می دهد که پیشامد A رخ داده اما پیشامد B رخ نداده باشد و داریم
  $A - B = \{x \ : \ x \in A \ \text{ and } \ x \notin B\} = A \cap B'$

- پیشامد $A\Delta B$ وقتی رخ می دهد که دقیقا یکی از پیشامدهای A و B رخ دهد و داریم $A\Delta B=\{x\ :\ (x\in A \ \text{ and } \ x\notin B) \text{ or } (x\notin A \ \text{ and } \ x\in B)\}$
- دو مجموعه A و B را مجزا یا ناسازگار گویند اگر این دو پیشامد نتوانند هیچگاه با هم رخ دهند، به عبارت دیگر داریم $A\cap B=\phi$
- اگر داشته باشیم $A\subset B$ آن گاه رخ دادن A رخ دادن B را نتیجه می دهد.

مروری بر نظریه مجموعه ها:

- $A=B \quad \Longleftrightarrow \quad A\subseteq B \ \text{ and } B\subseteq A$
- خاصیت جابجایی در مجموعه ها: $A\cup B=B\cup A$ و $A\cap B=B\cap A$
- خاصیت شرکت پذیری در مجموعه ها: $A\cup(B\cup C)=(A\cup B)\cup C$ و $A\cap(B\cap C)=(A\cap B)\cap C$
- خاصیت توزیع پذیری در مجموعه ها: $A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$ و $A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$
- قوانین دمورگان: $(A\cup B)'=A'\cap B'$ و $(A\cap B)'=A'\cup B'$

برای مثال پیشامد $(A'\cap B\cap C)\cup(A\cap B'\cap C)\cup(A\cap B\cap C')$ پیشامد این است که دقیقا دو تا از پیشامدهای A، B و C رخ دهد.

همانطور که گفتیم نظریه احتمال، مطالعه رویدادهای احتمالی از دیدگاه ریاضی است. مهم ترین بخش در ساختن یک مدل ریاضی برای آزمایش های تصادفی تخصیص یک عدد بین صفر و یک به هر پیشامد است که بیانگر شانس رخ دادن آن پیشامد باشد. این کار در سال ۱۹۳۳ توسط کولموگروف با تعریف سه اصل اولیه احتمال صورت گرفت.

تعریف ۱-۲. خانواده $\mathcal{F}$ از زیرمجموعه های Ω را یک σ-میدان (میدان سیگمایی) گویند هرگاه

۱. $\Omega\in\mathcal{F}$

۲. اگر $A\in\mathcal{F}$ آن گاه $A'\in\mathcal{F}$.

۳. اگر $\{A_n\}_{n\in\mathbb{N}}$ دنباله ای نامتناهی از اعضای $\mathcal{F}$ باشد، آن گاه $\bigcup_{n\in\mathbb{N}}A_n\in\mathcal{F}$.

به عبارت دیگر می توان گفت که σ-میدان نسبت به تمام اعمال نظریه مجموعه ها که به صورت متناهی یا شمارش پذیر اعمال شود بسته است.

برای مثال می توان دید که مجموعه توانی 2^Ω و $\{\emptyset,\Omega\}$، σ-میدان روی Ω هستند. همچنین برای هر پیشامد $A\in\Omega$، خانواده $\{A,A',\emptyset,\Omega\}$ نیز یک σ-میدان روی Ω است.

از این جا به بعد لفظ پیشامد را فقط برای اعضای یک σ-میدان $\mathcal{F}$ از زیرمجموعه های فضای نمونه Ω اطلاق می کنیم و فقط برای اعضای این σ-میدان احتمال را تعریف می کنیم. اما اینکه ''نحوه تخصیص احتمال باید چگونه باشد؟'' سوال مهمی است که لازم است اینجا به آن بپردازیم. تخصیص احتمال همانطور که در درس مبانی احتمال نیز گفته شد، ساختن تابعی است با دامنه $\mathcal{F}$ و بردی از زیرمجموعه $[0,1]$ که باید در اصول کلموگروف صدق کند. این تابع را اندازه احتمال گویند.

تعریف ۱-۳. هر اندازه احتمال روی فضای نمونه Ω تابعی چون $\mathrm{P}:\mathcal{F}\longrightarrow[0,1]$ است که در اصول زیر صدق کند:

۱. $\mathrm{P}(\Omega)=1$

۲. اگر $\{A_n\}_{n\in\mathbb{N}}$ دنباله ای نامتناهی از اعضای دو به دو مجزای $\mathcal{F}$ باشند (یعنی به ازای هر $i\neq j$، $A_i\cap A_j=\phi$) آن گاه

$$\mathrm{P}(\bigcup_{n=1}^{\infty}A_n)=\sum_{n=1}^{\infty}\mathrm{P}(A_n)$$

تعریف ۱-۴. به سه تایی مرتب $(\Omega,\mathcal{F},\mathrm{P})$ که Ω فضای نمونه در یک آزمایش تصادفی، $\mathcal{F}$ یک σ-میدان از زیرمجموعه های Ω (که گاهی اوقات به آن فضای پیشامد نیز گفته می شود) و P یک اندازه احتمال روی $\mathcal{F}$ است، فضای احتمال می گوییم.

توجه کنید که ساختن فضای احتمال برای یک آزمایش تصادفی کلی، کار آسانی نیست اما در حالتی که فضای نمونه آزمایش تصادفی متناهی یا نامتناهی شمارش پذیر باشد ساختن فضای احتمال راحت‌تر است که آن را فضای احتمال گسسته یا مدل احتمال گسسته می‌نامند. فرض کنید فضای نمونه Ω مربوط به یک آزمایش تصادفی متناهی شمارش‌پذیر ($\Omega=\{\omega_1,\omega_2,\cdots,\omega_m\}$) یا نامتناهی شمارش‌پذیر ($\Omega=\{\omega_1,\omega_2,\omega_3,\cdots\}$) باشد. در چنین شرایطی همیشه σ-میدان فضای پیشامد را کل زیرمجموعه‌های Ω درنظر می‌گیریم و داریم $\mathcal{F}=2^\Omega$ به عبارت دیگر در این حالت احتمال را برای تمام زیرمجموعه‌های فضای نمونه تعریف می‌کنیم.

در حالتی که Ω متناهی باشد، مقادیر $\mathrm{P}(\omega_i)=p_i$، $i=1,2,\cdots,m$، اعداد حقیقی نامنفی است که $\sum_{i=1}^m p_i=1$ و بنابراین تابع $\mathrm{P}\ :\ \mathcal{F}\longrightarrow[0,1]$ با ضابطه زیر یک اندازه احتمال روی $\mathcal{F}$ است،

$$\mathrm{P}(A)=\sum_{k:\ \omega_k\in A}p_k \qquad ;A\subset\Omega.$$

بطور مشابه در حالتی که Ω نامتناهی باشد، مقادیر $\mathrm{P}(\omega_i) = q_i$، $i = ۱, ۲, \cdots$، دنباله‌ای نامتناهی از اعداد حقیقی نامنفی است که $\sum_{i=۱}^{\infty} q_i = ۱$ و بنابراین تابع $\mathrm{P} : \mathcal{F} \longrightarrow [۰, ۱]$ با ضابطه زیر یک اندازه احتمال روی $\mathcal{F}$ است،

$$\mathrm{P}(A) = \sum_{k:\, \omega_k \in A} q_k \qquad ; A \subset \Omega.$$

پس در حالتی که فضای احتمال گسسته است، برای ساختن مدل احتمال آزمایش تصادفی کافی است احتمال پیشامدهای ساده (تک عضوی) بدست آورده شود. توجه کنید که در هر دو حالتی که Ω متناهی یا نامتناهی باشد داریم

$$\mathrm{P}(A) = \sum_{\omega \in A} \mathrm{P}(\omega).$$

اگر فضای نمونه متناهی باشد و $p_۱ = p_۲ = \cdots = p_m$، آن‌گاه

$$\mathrm{P}(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{m}$$

که در آن $|A|$ تعداد اعضای پیشامد A است. در چنین حالتی مدل احتمال گسسته به مدل احتمال گسسته یکنواخت معروف است.

برای بحث در مورد فضای احتمال یکنواخت پیوسته، فرض کنید Ω زیرمجموعه ای از فضای اقلیدسی $\mathbb{R}^d$ باشد بطوری که دارای حجم d-بعدی ناصفر و متناهی است. بطور شهودی می توان گفت که اگر A و B دو زیرمجموعه با حجم یکسان از فضای نمونه باشند انتظار داریم احتمال رخداد آن‌ها برابر باشد. پس احتمال هر پیشامد متناسب با حجم آن خواهد بود. چون احتمال کل فضای نمونه برابر با یک است پس احتمال هر پیشامد A از فضای نمونه برابر است با $\mathrm{P}(A) = \frac{|A|}{|\Omega|}$ یعنی نسبت حجم مجموعه A به حجم فضای نمونه.

لازم به ذکر است که در مدل احتمال یکنواخت پیوسته باید هم فضای نمونه و هم پیشامدها بورل باشند تا مفهوم حجم برای آن‌ها قابل تعریف باشد. همچنین منظور از حجم در فضای یک بعدی طول، در فضای دو بعدی مساحت، و در فضای سه بعدی حجم و در فضاهایی با بعد بالاتر که (البته ما به آن نمی‌پردازیم) ابرحجم است.

مثال ۱-۵. نقطه ای به تصادف از داخل دایره ای به مرکز مبدا مختصات و شعاع یک انتخاب میکنیم. احتمال اینکه عدد انتخاب شده در دایره ای به مرکز مختصات و شعاع $\frac{۱}{۲}$ باشد را بدست آورید.

حل. در این سوال داریم $\Omega = \{(x, y) \in \mathbb{R}^۲ : x^۲ + y^۲ < ۱\}$ و A پیشامد انتخاب عدد در دایره ای به مرکز مختصات و شعاع $\frac{۱}{۲}$ تعریف می شود. بنابراین

$$\mathrm{P}(A) = \frac{|A|}{|\Omega|} = \frac{\pi(\frac{۱}{۲})^۲}{\pi} = \frac{۱}{۴}.$$

هر اندازه احتمال P دارای خواص زیر است:

۱. اگر Aو B دو پیشامد دلخواه باشند آن گاه $\mathrm{P}(B-A)=\mathrm{P}(B)-\mathrm{P}(A\cap B)$ و $\mathrm{P}(A\cup B)=\mathrm{P}(A)+\mathrm{P}(B)-\mathrm{P}(A\cap B)$.

۲. اگر $A\subseteq B$ آن گاه $\mathrm{P}(A)\leq \mathrm{P}(B)$ یعنی اندازه احتمال یک تابع صعودی است.

۳. اگر $B_۱,B_۲,\dots B_n$ پیشامدهای دو به دو ناسازگار باشند آن گاه

$$\mathrm{P}(B_۱\cup B_۲\cup\dots B_n)=\mathrm{P}(B_۱)+\mathrm{P}(B_۲)+\dots+\mathrm{P}(B_n)$$

۴. برای هر n پیشامد دلخواه $A_۱,A_۲,\dots,A_n$ داریم

$$\mathrm{P}(\bigcup_{i=۱}^{n}A_i)=\sum_{i=۱}^{n}\mathrm{P}(A_i)-\sum_{۱<i<j<n}\mathrm{P}(A_i\cap A_j)+\cdots+(-۱)^{n-۱}\mathrm{P}(A_۱\cap\cdots\cap A_n).$$

۵. برای $I=\mathbb{N}$ یا $I=\{۱,۲,\cdots,n\}$ و هر دنباله $\{A_i\}_{i\in I}\subset\mathcal{F}$ از پیشامدها داریم

$$\mathrm{P}(\bigcup_{i\in I}A_i)\leq\sum_{i\in I}\mathrm{P}(A_i).$$

نکته ۱-۶. دنباله $\{A_n\}_{n\in\mathbb{N}}$ از اعضای $\mathcal{F}$ را درنظر بگیرید. اگر دنباله $\{B_n\}_{n\in\mathbb{N}}$ از زیرمجموعه های Ω را بصورت زیر تعریف کنیم

$$\begin{aligned}
B_۱&=A_۱\\
B_۲&=A_۲-A_۱\\
B_۳&=A_۳-(A_۱\cup A_۲)\\
&\vdots\\
B_n&=A_n-\bigcup_{k=۱}^{n-۱}A_k\\
&\vdots
\end{aligned}$$

آن گاه $\{B_n\}_{n\in\mathbb{N}}$ دنباله ای از اعضای دوبدو مجزای $\mathcal{F}$ است و داریم

$$\bigcup_{k=۱}^{n}B_k=\bigcup_{k=۱}^{n}A_k\qquad\forall\, n\in\mathbb{N}$$
$$\bigcup_{k=۱}^{\infty}B_k=\bigcup_{k=۱}^{\infty}A_k$$

تعریف ۱-۷. برای $I = \mathbb{N}$ یا $I = \{1, 2, \cdots, n\}$، دنباله $\{B_n\}_{n\in I} \subset \mathcal{F}$ یک افراز فضای نمونه نامیده می‌شود هرگاه

۱. B_n ها دو به دو مجزا باشند، یعنی برای هر $i, j \in I$ بطوریکه $i \neq j$، داشته باشیم $B_i \cap B_j = \emptyset$

۲. $\bigcup_{n\in I} B_n = \Omega$

تعریف ۱-۸. دنباله $\{A_n\}_{n\in\mathbb{N}}$ از اعضای $\mathcal{F}$ را صعودی گوییم هرگاه

$$\forall\, n \in \mathbb{N} \quad A_n \subset A_{n+1}$$

و بطور مشابه $\{A_n\}_{n\in\mathbb{N}}$ را نزولی گوییم هرگاه داشته باشیم

$$\forall\, n \in \mathbb{N} \quad A_{n+1} \subset A_n$$

قضیه ۱-۹. برای هر دنباله صعودی $\{A_n\}_{n\in\mathbb{N}}$ از اعضای $\mathcal{F}$ داریم

$$\mathrm{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mathrm{P}(A_n) \tag{۱-۱}$$

و بطور مشابه اگر $\{A_n\}_{n\in\mathbb{N}}$ یک دنباله نزولی باشد آن گاه

$$\mathrm{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mathrm{P}(A_n) \tag{۱-۲}$$

اثبات. در حالت صعودی فرض کنید $\{B_n\}_{n\in\mathbb{N}}$ دنباله ای تعریف شده بصورت نکته ۱-۶ باشد. در این صورت چون دنباله $\{A_n\}_{n\in\mathbb{N}}$ صعودی است داریم

$$\begin{aligned}
B_1 &= A_1 \\
B_2 &= A_2 - A_1 \\
B_3 &= A_3 - (A_1 \cup A_2) = A_3 - A_2 \\
&\vdots \\
B_n &= A_n - \bigcup_{k=1}^{n-1} A_k = A_n - A_{n-1} \\
&\vdots
\end{aligned}$$

و از آنجایی که $\{B_n\}_{n\in\mathbb{N}}$ دوبه دو مجزا هستند بدست می آوریم

$$\bigcup_{k=1}^{n} B_k = \bigcup_{k=1}^{n} A_k = A_n \qquad \forall\, n\in\mathbb{N}$$
$$\bigcup_{k=1}^{\infty} B_k = \bigcup_{k=1}^{\infty} A_k$$

در نتیجه بخاطر شمارجمعی بودن اندازه احتمال و تعریف یک سری نامتناهی داریم

$$\begin{aligned}
\mathrm{P}\left(\bigcup_{n=1}^{\infty} A_n\right) &= \mathrm{P}\left(\bigcup_{n=1}^{\infty} B_n\right)\\
&= \sum_{n=1}^{\infty} \mathrm{P}(B_n)\\
&= \lim_{n\to\infty}\sum_{k=1}^{n} \mathrm{P}(B_k)\\
&= \lim_{n\to\infty} \mathrm{P}\left(\bigcup_{k=1}^{n} B_k\right)\\
&= \lim_{n\to\infty} \mathrm{P}(A_n).
\end{aligned}$$

برای اثبات در حالت نزولی کافی است از این نکته استفاده کنیم که وقتی $\{A_n\}_{n\in\mathbb{N}}$ نزولی است آن گاه $\{A'_n\}_{n\in\mathbb{N}}$ صعودی است و بنابراین اثبات مشابه حالت صعودی کامل می شود. (تمرین) □

۱.۱ احتمال شرطی و استقلال پیشامدها

برای هر دو پیشامد A و B در $\mathcal{F}$، احتمال رخ دادن پیشامد A به شرط اینکه بدانیم پیشامد B رخ داده است را با $\mathrm{P}(A|B)$ نمایش داده و تعریف می کنیم

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(A\cap B)}{\mathrm{P}(B)} \qquad \text{if } \mathrm{P}(B) > 0$$

اگر $\mathrm{P}(B) = 0$ آن گاه احتمال شرطی فوق قابل تعریف نیست.

توجه داشته باشید که $\mathrm{P}(A|B)$ به عنوان تابعی از A یک اندازه احتمال است زیرا تابع $\mathrm{P}(.|B)\ :\ \mathcal{F}\longrightarrow [0,1]$ دارای خواص زیر است

۱) $\mathrm{P}(\Omega|B) = \frac{\mathrm{P}(\Omega\cap B)}{\mathrm{P}(B)} = \frac{\mathrm{P}(B)}{\mathrm{P}(B)} = 1$

۲) اگر $\{A_n\}_{n\in\mathbb{N}}$ دنباله ای نامتناهی از اعضای دو به دو مجزای $\mathcal{F}$ باشند آن‌گاه

$$\mathrm{P}\left(\bigcup_{n=1}^{\infty} A_n|B\right)=\sum_{n=1}^{\infty}\mathrm{P}(A_n|B)$$

و از اینرو احتمال شرطی تمام خواصی که برای یک اندازه احتمال برقرار بوده را دارا است. یعنی برای یک پیشامد B بطوریکه $\mathrm{P}(B)>0$ داریم

۱. اگر A و C دو پیشامد در $\mathcal{F}$ باشند بطوریکه $A\subseteq C$ آن‌گاه $\mathrm{P}(A|B)\leq \mathrm{P}(C|B)$.

۲. اگر A و C دو پیشامد دلخواه باشند آن‌گاه

$$\mathrm{P}(A-C|B)=\mathrm{P}(A|B)-\mathrm{P}(A\cap C|B)$$
$$\mathrm{P}(A\cup C|B)=\mathrm{P}(A|B)+\mathrm{P}(C|B)-\mathrm{P}(A\cap C|B)$$

۳. برای هر n پیشامد دلخواه $A_1,A_2,\ldots,A_n$ داریم

$$\mathrm{P}\left(\bigcup_{i=1}^{n} A_i|B\right)=\sum_{i=1}^{n}\mathrm{P}(A_i|B)-\sum_{1<i<j<n}\mathrm{P}(A_i\cap A_j|B)+\cdots+(-1)^{n-1}\mathrm{P}(A_1\cap\cdots\cap A_n|B).$$

۴. برای $I=\mathbb{N}$ یا $I=\{1,2,\cdots,n\}$ و هر دنباله $\{A_i\}_{i\in I}\subset\mathcal{F}$ از پیشامدها داریم

$$\mathrm{P}\left(\bigcup_{i\in I} A_i|B\right)\leq\sum_{i\in I}\mathrm{P}(A_i|B).$$

۵. برای هر دنباله صعودی $\{A_n\}_{n\in\mathbb{N}}$ از اعضای $\mathcal{F}$ داریم

$$\mathrm{P}\left(\bigcup_{n=1}^{\infty} A_n|B\right)=\lim_{n\to\infty}\mathrm{P}(A_n|B)$$

و بطور مشابه برای دنباله نزولی $\{A_n\}_{n\in\mathbb{N}}$ رابطه زیر همواره برقرار است

$$\mathrm{P}\left(\bigcap_{n=1}^{\infty} A_n|B\right)=\lim_{n\to\infty}\mathrm{P}(A_n|B).$$

مثال ۱-۱۰. آزمایش پرتاب یک تاس را در نظر بگیرید.

i. اگر A پیشامد مشاهده عدد فرد در این آزمایش باشد داریم

$$\mathrm{P}(A) = \frac{n(A)}{n(\Omega)} = \frac{۳}{۶}$$

ii. اگر بدانیم که عدد مشاهده شده اول است، برای محاسبه احتمال مشاهده عدد فرد داریم
پیشامد مشاهده عدد اول: B
پیشامد مشاهده عدد فرد: A
$A = \{۱, ۳, ۵\}$، $B = \{۲, ۳, ۵\}$ و $A \cap B = \{۳, ۵\}$ بنابراین

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \frac{\frac{۲}{۶}}{\frac{۳}{۶}} = \frac{۲}{۳}$$

iii. اگر عدد مشاهده شده بزرگتر از ۳ باشد، می خواهیم بررسی کنیم که با چه احتمالی عدد مشاهده شده فرد است. در این حالت داریم:
پیشامد مشاهده عدد بزرگتر از ۳: C
$C = \{۴, ۵, ۶\}$ و $A \cap C = \{۵\}$ و از اینرو

$$\mathrm{P}(A|C) = \frac{\mathrm{P}(A \cap C)}{\mathrm{P}(C)} = \frac{\frac{۱}{۶}}{\frac{۳}{۶}} = \frac{۱}{۳}$$

در حالتی که آزمایش تصادفی، چندمرحله ای باشد، احتمال شرطی ابزار مناسبی برای محاسبه احتمالات شرطی و غیرشرطی است. قوانین حاصلضربی احتمال، احتمال کل و فرمول بیز سه تا از مهم ترین فرمول ها در مبحث احتمال شرطی هستند.

قضیه ۱-۱۱. برای هر $A_۱, A_۲, \ldots, A_n \in \mathcal{F}$ که $\mathrm{P}(A_۱ \cap A_۲ \cap \ldots \cap A_n - ۱) > ۰$ داریم

$$\mathrm{P}(A_۱ \cap A_۲ \cap \ldots \cap A_n) = \mathrm{P}(A_۱)\mathrm{P}(A_۲|A_۱) \ldots \mathrm{P}(A_n|A_۱ \cap A_۲ \cap \ldots \cap A_n - ۱)$$

یا به عبارتی

$$\mathrm{P}\left(\bigcap_{k=۱}^{n} A_k\right) = \mathrm{P}(A_۱) \prod_{k=۲}^{n} \mathrm{P}(A_k|A_۱ \cap A_۲ \cap \ldots \cap A_k - ۱). \qquad (۱-۳)$$

اثبات.

$$\begin{aligned}
\mathrm{P}(A_۱ \cap A_۲ \cap \ldots \cap A_n) &= \mathrm{P}(A_۱ \cap A_۲ \cap \ldots \cap A_{n-۱})\mathrm{P}(A_n|A_۱ \cap A_۲ \cap \ldots \cap A_{n-۱}) \\
&= \mathrm{P}(A_۱ \cap A_۲ \cap \ldots \cap A_{n-۲})\mathrm{P}(A_{n-۱}|A_۱ \cap A_۲ \cap \ldots \cap A_{n-۲}) \\
&\quad \times \mathrm{P}(A_n|A_۱ \cap A_۲ \cap \ldots \cap A_{n-۱}) \\
&= \mathrm{P}(A_۱ \cap A_۲ \cap \ldots \cap A_{n-۳})\mathrm{P}(A_{n-۲}|A_۱ \cap A_۲ \cap \ldots \cap A_{n-۳}) \\
&\quad \times \mathrm{P}(A_{n-۱}|A_۱ \cap A_۲ \cap \ldots \cap A_{n-۲})\mathrm{P}(A_n|A_۱ \cap A_۲ \cap \ldots \cap A_{n-۱}) \\
&\quad \vdots \\
&= \mathrm{P}(A_۱)\mathrm{P}(A_۲|A_۱)\ldots\mathrm{P}(A_n|A_۱ \cap A_۲ \cap \ldots \cap A_n - ۱)
\end{aligned}$$

□

قضیه ۱-۱۲. (قانون احتمال کل) اگر $\{B_n\}_{n\in\mathbb{N}}$ افرازی از فضای نمونه باشد، آن گاه برای هر پیشامد A داریم

$$\mathrm{P}(A) = \sum_{n\in I} \mathrm{P}(A \cap B_n)$$

و در صورتی که برای هر $n \in I$ داشته باشیم $\mathrm{P}(B_n) > ۰$ آنگاه

$$\mathrm{P}(A) = \sum_{n\in I} \mathrm{P}(A \cap B_n) = \sum_{n\in I} \mathrm{P}(B_n)\mathrm{P}(A|B_n)$$

□

طبق تعریف احتمال شرطی، برای هر دو پیشامد دلخواه A و B داریم

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(A\cap B)}{P(B)} = \frac{\mathrm{P}(B|A)\mathrm{P}(A)}{P(B)}$$

همچنین چون $\mathrm{P}(B) = \mathrm{P}(B\cap A) + \mathrm{P}(B \cap A')$ (شکل ۱-۲ را ببینید)، می توان رابطه بالا را به صورت زیر بازنویسی کرد

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(B|A)\mathrm{P}(A)}{\mathrm{P}(B|A)\mathrm{P}(A) + \mathrm{P}(B|A')\mathrm{P}(A')}.$$

به طور کلی قضیه زیر را می توان ثابت کرد که به فرمول بیز معروف است.

قضیه ۱-۱۳. (فرمول بیز) اگر $\{B_n\}_{n\in\mathbb{N}}$ افرازی از فضای نمونه باشد، آن گاه برای هر پیشامد A که $\mathrm{P}(A) > ۰$ داریم

$$\mathrm{P}(B_k|A) = \frac{\mathrm{P}(A|B_k)\mathrm{P}(B_k)}{\sum_{n\in I}\mathrm{P}(A|B_n)\mathrm{P}(B_n)}.$$

شکل ۱-۱: نمایش دو پیشامد دلخواه A و B در یک فضای نمونه S

مثال ۱-۱۴. از جامعه ای که ۴۰ درصد آن ها را مردان و ۶۰ درصد را زنان تشکیل می دهند، یک نفر را انتخاب می کنیم. فرض کنید ۵۰ درصد از مردان و ۳۰ درصد از زنان سیگاری باشند. اگر فرد انتخاب شده سیگاری باشد، با چه احتمالی مرد است؟

تعریف می کنیم:

سیگاری بودن: A

مرد بودن: B

می خواهیم مقدار $\mathrm{P}(B|A)$ را بدست آوریم. طبق قانون احتمال بیز داریم

$$\begin{aligned}
\mathrm{P}(B|A) &= \frac{\mathrm{P}(A|B)\mathrm{P}(B)}{\mathrm{P}(A)} \\
&= \frac{\mathrm{P}(A|B)\mathrm{P}(B)}{\mathrm{P}(A|B)\mathrm{P}(B) + \mathrm{P}(A|B')\mathrm{P}(B')} \\
&= \frac{(0/5)(0/4)}{(0/5)(0/4) + (0/3)(0/6)} \\
&= \frac{0/2}{0/38} = 0/53
\end{aligned}$$

تمرین ۱-۱۵. شخصی به تصادف یکی از اعداد صحیح ۱، ۲ و ۳ را انتخاب می کند و سپس به تعداد عدد انتخاب شده یک تاس را پرتاب می کند.

الف. احتمال اینکه مجموع ۵ بیاورد را بیابید.

ب. اگر این شخص مجموع ۴ آورده باشد، احتمال اینکه عدد انتخاب شده ۲ باشد را بیابید.

نکته ۱-۱۶. چون احتمال شرطی همه خواص اندازه احتمال را دارد، می توان صورت های شرطی قانون احتمال کل و قانون حاصلضرب احتمال را برای احتمال شرطی بصورت زیر نوشت:

۱. برای هر $A_1, A_2, \ldots, A_n, B \in \mathcal{F}$ بطوریکه $\mathrm{P}(B \cap A_1 \cap A_2 \cap \ldots \cap A_{n-1}) > 0$، داریم

$$\mathrm{P}(A_1 \cap A_2 \cap \ldots \cap A_n|B) = \mathrm{P}(A_1|B)\mathrm{P}(A_2|B \cap A_1)\ldots\mathrm{P}(A_n|B \cap A_1 \cap A_2 \cap \ldots \cap A_{n-1})$$

یا

$$\mathrm{P}\left(\bigcap_{k=1}^{n} A_k|B\right) = \mathrm{P}(A_1|B)\prod_{k=2}^{n}\mathrm{P}(A_k|B \cap A_1 \cap A_2 \cap \ldots \cap A_{k-1})$$

۲. اگر $\{B_n\}_{n\in\mathbb{N}}$ یک افراز فضای نمونه باشد، آن گاه برای هر دو پیشامد A و C بطوریکه $\mathrm{P}(C) > 0$ داریم

$$\mathrm{P}(A|C) = \sum_{n\in I} \mathrm{P}(A \cap B_n|C)$$

و در صورتی که برای هر $n \in I$، داشته باشیم $\mathrm{P}(C \cap B_n) > 0$ آن گاه

$$\mathrm{P}(A|C) = \sum_{n\in I} \mathrm{P}(B_n|C)\mathrm{P}(A|C \cap B_n).$$

مثال ۱–۱۷. فرض کنید جعبه ای شامل b توپ سیاه و r توپ قرمز باشد. توپ ها را یک به یک و بدون جایگذاری از جعبه انتخاب کرده و هر بار رنگ توپ را ثبت می کنیم. احتمال اینکه توپ اول و دوم انتخاب شده هر دو سیاه باشند را بدست آورید. همچنین احتمال اینکه توپ دوم انتخاب شده سیاه باشد را بنویسید.
حل. برای $i \geq 1$، B_i را پیشامد اینکه توپ i–ام سیاه باشد تعریف می کنیم. در قسمت اول با استفاده از قانون حاصلضرب احتمال داریم

$$\mathrm{P}(B_1 \cap B_2) = \mathrm{P}(B_1)\mathrm{P}(B_2|B_1) = \frac{b}{b+r} \times \frac{b-1}{b+r-1}.$$

در قسمت دوم داریم

$$\mathrm{P}(B_2) = \mathrm{P}(B_2|B_1)\mathrm{P}(B_1)+\mathrm{P}(B_2|B_1')\mathrm{P}(B_1') = \frac{b-1}{b+r-1}\times\frac{b}{b+r}+\frac{b}{b+r-1}\times\frac{r}{b+r} = \frac{b}{b+r}.$$

که برابر با $\mathrm{P}(B_1)$ است. بطور کلی می توان ثابت کرد که برای هر $i = 1, 2, \ldots, b+r$ داریم $\mathrm{P}(B_i) = \frac{b}{b+r}$ (تمرین).

مثال ۱–۱۸. مشابه مثال ۱–۱۷ فرض کنید جعبه ای شامل b توپ سیاه و r توپ قرمز داشته باشیم. اگر توپ ها را یک به یک از جعبه انتخاب کنیم و در هر بار انتخاب، توپ انتخاب شده را به همراه c توپ اضافی همرنگ به جعبه بازگردانیم، مطلوب است،

الف. احتمال اینکه توپ اول و دوم انتخاب شده هر دو سیاه باشند را بدست آورید.

ب. احتمال اینکه توپ دوم انتخاب شده سیاه باشد را بدست آورید.

ج. اگر بدانیم که توپ اول قرمز انتخاب شده احتمال اینکه سه توپ انتخاب شده بعدی سیاه باشد را بدست آورید.

د. اگر بدانیم که توپ اول قرمز انتخاب شده احتمال اینکه توپ انتخاب شده سوم هم قرمز باشد را بدست آورید.

حل. برای $i \geq 1$، B_i را پیشامد اینکه توپ i-ام سیاه باشد تعریف می کنیم. داریم

$$\mathrm{P}(B_1 \cap B_2) = \mathrm{P}(B_1)\mathrm{P}(B_2|B_1) = \frac{b}{b+r} \times \frac{b+c}{b+r+c}$$

$$\begin{aligned}\mathrm{P}(B_2) &= \mathrm{P}(B_2|B_1)\mathrm{P}(B_1) + \mathrm{P}(B_2|B'_1)\mathrm{P}(B'_1) = \frac{b+c}{b+r+c} \times \frac{b}{b+r} + \frac{b}{b+r+c} \times \frac{r}{b+r}\\ &= \frac{b}{b+r} = \mathrm{P}(B_1)\end{aligned}$$

$$\begin{aligned}\mathrm{P}(B_2 \cap B_3 \cap B_4|B'_1) &= \mathrm{P}(B_2|B'_1)\mathrm{P}(B_3|B'_1 \cap B_2)\mathrm{P}(B_4|B'_1 \cap B_2 \cap B_3)\\ &= \frac{b}{b+r+c} \times \frac{b+c}{b+r+2c} \times \frac{b+2c}{b+r+3c}\end{aligned}$$

$$\begin{aligned}\mathrm{P}(B'_3|B'_1) &= \mathrm{P}(B'_3 \cap B_2|B'_1) + \mathrm{P}(B'_3 \cap B'_2|B'_1) = \mathrm{P}(B_2|B'_1)\mathrm{P}(B'_3|B'_1 \cap B_2) + \mathrm{P}(B'_2|B'_1)\mathrm{P}(B'_3|B'_1 \cap B'_2)\\ &= \frac{b}{b+r+c} \times \frac{r+c}{b+r+2c} + \frac{r+c}{b+r+c} \times \frac{r+2c}{b+r+2c}\\ &= \frac{r+c}{b+r+c} = \mathrm{P}(B'_2|B'_1).\end{aligned}$$

۲.۱ استقلال پیشامدها

در مثال های مربوط به احتمال شرطی، احتمال یک پیشامد مثل A مشروط بر هر یک از پیشامدهای B یا C گاهی متفاوت از احتمال A است. در چنین مواقعی گفته می شود که پیشامد A مستقل از این دو پیشامد نبوده است چون رخ دادن یا ندادن آن ها روی رخ دادن یا ندادن پیشامد A تاثیرگذار بوده است. به زبان ساده دو پیشامد A و B را مستقل گویند هرگاه $\mathrm{P}(A|B) = \mathrm{P}(A)$ و $\mathrm{P}(B|A) = \mathrm{P}(B)$ به شرط اینکه $\mathrm{P}(A) > 0$ و $\mathrm{P}(B) > 0$. بنابراین طبق تعریف احتمال شرطی، دو پیشامد A و B مستقل هستند اگر و تنها اگر

$$\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$$

تعریف ۱-۱۹. فرض کنید $(\Omega, \mathcal{F}, \mathrm{P})$ یک فضای احتمال باشد.

۱. دو پیشامد A و B در $\mathcal{F}$ را مستقل گویند هرگاه

$$\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B) \tag{۱-۴}$$

۲. پیشامدهای $A_1, A_2, \cdots, A_n \in \mathcal{F}$ را مستقل گویند هرگاه برای هر انتخاب $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ داشته باشیم

$$\mathrm{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = \mathrm{P}(A_{i_1})\mathrm{P}(A_{i_2}) \cdots \mathrm{P}(A_{i_k})$$

۳. دنباله $\{A_n\}_{n=1,2,\ldots}$ از پیشامدها در $\mathcal{F}$ را مستقل گویند هرگاه هر تعداد متناهی از این پیشامدها مستقل باشد، یا به عبارت دیگر برای هر $n \geq 1$ پیشامدهای $A_1, A_2, \cdots, A_n$ مستقل باشند.

۴. دو پیشامد A و B در $\mathcal{F}$ را به شرط پیشامد $C \in \mathcal{F}$ مستقل گویند هرگاه

$$\mathrm{P}(A \cap B | C) = \mathrm{P}(A|C)\mathrm{P}(B|C)$$

⋆ توجه کنید که پیشامدهای مستقل می توانند به طور همزمان رخ دهند اما رخ دادن یکی از آن ها تاثیری بر رخ دادن دیگری ندارد. در حالی که پیشامدهای مجزا یا ناسازگار، نمی توانند به طور همزمان اتفاق بیفتند.

نکته ۱-۲۰. اگر پیشامدهای A, B, C مستقل باشند آن گاه به سادگی می توان ثابت کرد که $A\Delta B$، $A-B$، $B-A$، $A\cup B$ و $A\cap B$ نیز از C مستقل هستند.

تمرین ۱-۲۱. فرض کنید نقطه ای به تصادف از داخل مربع واحد انتخاب کنیم. بررسی کنید که آیا دو پیشامد A و B تعریف شده در زیر مستقل هستند یا نه؟

$$A = \{(x,y) \in \Omega :\ 0 \leq x \leq 0/5\} = [0, 0/5] \times [0, 1]$$
$$B = \{(x,y) \in \Omega :\ 0 \leq y \leq 0/5\} = [0, 1] \times [0, 0/5]$$

در خیلی از مواقع آزمایش های تصادفی تکرار متناهی (یا حتی نامتناهی) شمارش‌پذیر از آزمایش های ساده تر است مثل آزمایش تصادفی مربوط به پرتاپ های متوالی یک سکه یا یک تاس. در این گونه موارد میتوان فرض کرد که برآمدهای مربوط به هر تکرار (مثلا تکرار i-ام) تاثیری بر روی برآمدهای مربوط به تکرارهای دیگر ندارد که در این صورت می گوییم آزمایش تصادفی اصلی، تکرار مستقل آزمایش ساده تر است. در حالتی که مدل احتمالی آزمایش ساده معلوم باشد میتوان مدل احتمالی آزمایش مربوط به تکرار مستقل آزمایش ساده تر را معین کرد.

اکنون به عنوان یک حالت خاص در این قسمت مدل احتمالی آزمایش تصادفی مربوط به n بار تکرار مستقل آزمایش برنولی را مشخص میکنیم. لازم است یادآوری کنیم که یک آزماشی برنولی، آزمایش تصادفی است که فقط شامل دو برآمد ممکن باشد که آن ها را موفقیت و شکست مینامیم و به ترتیب با s و f نشان میدهیم. چون فقط دو برامد ممکن وجود دارد و بنابراین فضای احتمال گسسته است، برای تعیین مدل احتمالی آزمایش برنولی کافی است که فقط احتمالات موفقیت و شکست را بدانیم. فرض میکنیم احتمال رخ دادن پیروزی p و بنابراین احتمال رخ دادن شکست $1-p$ خواهد بود که آن را با q نشان میدهیم. حال فرض کنید آزمایش برنولی را n بار بطور مستقل تکرار کنیم. فضای نمونه این آزمایش تصادفی را میتوان

بصورت $\Omega = \{s, f\}^n$ نمایش داد که مجموعه تمام n تایی‌های مرتب از s و f است و بنابراین ۲^n عضو دارد. به عبارت دیگر داریم

$$\begin{aligned}\Omega &= \{\boldsymbol{\omega} = (\omega_۱, \omega_۲, \cdots, \omega_n),\ \omega_i \in \{s, f\},\ i = ۱, ۲, \cdots, n\} \\ &= \{(s, s, \cdots, s), (f, s, s, \cdots, s), \cdots, (f, f, \cdots, f)\}\end{aligned}$$

چون Ω متناهی است پس مدل احتمالی این آزمایش گسسته است، یعنی فضای پیشامد $\mathcal{F} = ۲^\Omega$ است و کافیست احتمال هر برآمد را تعیین کنیم. چون آزمایش تصادفی تکرار مستقل آزمایش برنولی است پس معقول است که فرض کنیم احتمال برآمد $(s, s, \cdots, s)$ برابر با p^n است و بطور مشابه داریم

$$\begin{aligned}&\mathrm{P}(\{(f, f, \cdots, f)\}) = q^n \\ &\mathrm{P}(\{(\underbrace{f, f, \cdots, f}_{k}, \underbrace{s, s, \cdots, s}_{n-k})\}) = p^k q^{n-k}\end{aligned}$$

پس در حالت کلی $\mathrm{P}(\{(\omega_۱, \omega_۲, \cdots, \omega_n)\}) = p^k q^{n-k}$ به طوریکه s تا موفقیت در بردار $(\omega_۱, \omega_۲, \cdots, \omega_n)$ وجود دارد. حال اگر A را پیشامد مشاهده دقیقا k موفقیت در n تکرار آزمایش برنولی تعریف شود آن گاه

$$\begin{aligned}\mathrm{P}(A) &= \sum_{\boldsymbol{\omega} \in A} \mathrm{P}(\{\boldsymbol{\omega}\}) = \sum_{\boldsymbol{\omega} \in A} p^k q^{n-k} \\ &= |A| p^k q^{n-k} = \binom{n}{k} p^k q^{n-k}.\end{aligned}$$

نکته ۱-۲۲. اگر در آزمایش تصادفی مربوط به n بار تکرار مستقل آزمایش برنولی، A را پیشامد مشاهده موفقیت در اولین تکرار و B را پیشامد مشاهده شکست در دومین تکرار تعریف کنیم آن گاه می توان ثابت کرد که

$$\begin{aligned}\mathrm{P}(A) &= \mathrm{P}(\{\boldsymbol{\omega} = (\omega_۱, \omega_۲, \cdots, \omega_n)\ :\ \omega_۱ = s\}) = p \\ \mathrm{P}(B) &= \mathrm{P}(\{\boldsymbol{\omega} = (\omega_۱, \omega_۲, \cdots, \omega_n)\ :\ \omega_۲ = f\}) = q \\ \mathrm{P}(A \cap B) &= \mathrm{P}(\{\boldsymbol{\omega} = (\omega_۱, \omega_۲, \cdots, \omega_n)\ : \omega_۱ = s \text{ and } \omega_۲ = f\}) = pq\end{aligned}$$

پس داریم $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$ و درنتیجه دو پیشامد A و B مستقل هستند.

در واقع آنچه در نکته بالا اهمیت دارد این است که پیشامد A فقط به اولین تکرار و پیشامد B فقط به دومین تکرار بستگی دارد و چون آزمایش های برنولی را بطور مستقل تکرار می کنیم واضح است که

پیشامدهای A و B مستقل باشند. این نتیجه که در این درس بسیار مورد استفاده قرار می گیرد را می توان بصورت کلی به شکل زیر مطرح کرد:

$\star$ اگر یک آزمایش تصادفی تکرار مستقل (متناهی یا نامتناهی شمارش پذیر) از یک آزمایش ساده تر (یا اینکه انجام یک دنباله متناهی یا نامتناهی شمارش پذیر از آزمایش های مختلف بصورت مستقل) باشد و پیشامدهای A_i، $i = ۱, ۲, \cdots$، فقط به تکرار i-ام بستگی داشته باشد آنگاه پیشامدهای A_i ها مستقل هستند.

۳.۱ متغیرتصادفی و بردارتصادفی

گاهی اوقات بعضی از جنبه‌های عددی یا کمیت‌های عددی در ارتباط با یک آزمایش تصادفی مهم است. مثلا مجموع ارقام مشاهده شده در پرتاب دو تاس بصورت هم‌زمان، تعداد دفعات پرتاب یک تاس تا مشاهده وجه ۶. این کمیت‌های عددی، متغیر تصادفی نام دارند که با حروف بزرگ لاتین مثل X، Y، W و ... نمایش داده می‌شوند. در این بخش بطور مختصر به مفاهیمی چون متغیر تصادفی، بردار تصادفی، تابع توزیع، تابع توزیع توام، تابع احتمال، تابع احتمال توام، تابع چگالی و تابع چگالی توام میپردازیم. توجه کنید که از اینجا به بعد هرگاه از زیرمجموعه‌های $\mathbb{R}$ و یا $\mathbb{R}^n$ صحبت شود منظور زیرمجموعه‌های بورل آن‌هاست.

۱.۳.۱ متغیرهای تصادفی

تعریف ۱-۲۳. متغیر تصادفی X روی یک فضای احتمال $(\Omega, \mathcal{F}, \mathrm{P})$، یک تابع حقیقی به صورت $X : \Omega \longrightarrow \mathbb{R}$ است بطوریکه برای هر زیر مجموعه $A \subset \mathbb{R}$ داشته باشیم

$$(X \in A) = (\omega \ : \ X(\omega) \in A) \in \mathcal{F}$$

برد یک متغیر تصادفی X را مجموعه مقادیر یا تکیه گاه آن گوییم و آن را با S_X نشان می دهیم.

هرگاه مجموعه مقادیر S_X شمارش پذیر (متناهی یا نامتناهی) باشد، متغیرتصادفی X یک متغیرتصادفی گسسته نامیده می شود.

خواص احتمالی یا اماری یک متغیر تصادفی در توزیع آن نهفته است. توزیع یک متغیر تصادفی وقتی معلوم است که احتمال تمام پیشامدهای مربوط به آن معلوم باشد. درواقع احتمالات زیر معلوم باشد

$$\mathrm{P}(X \in A) \, , \quad \forall \, A \subset \mathbb{R}$$

در حالت کلی کار با احتمالات بالا ساده نیست و در محاسبات آن ها دچار چالش‌هایی خواهیم بود. به همین دلیل به جای بررسی احتمالات بالا، در حالتی که متغیر تصادفی داریم با **تابع توزیع** کار می کنیم. در واقع دانستن تابع توزیع معادل با دانستن توزیع متغیر تصادفی است.

تعریف ۱-۲۴. فرض کنید X یک متغیر تصادفی روی فضای احتمال $(\Omega, \mathcal{F}, \mathrm{P})$ باشد. **تابع توزیع** X را با F_X نشان داده و تعریف می کنیم

$$F_X: \mathbb{R} \longrightarrow \mathbb{R}, \qquad F_X(x) = \mathrm{P}(X \leq x) = \mathrm{P}(\{\omega: X(\omega) \leq x\})$$

که البته اندیس X در تعریف $F_X(x)$ صرفا جهت تاکید بر این است که تابع توزیع برای متغیر تصادفی X است.

قضیه ۱-۲۵. اگر F یک تابع توزیع برای یک متغیر تصادفی دلخواه باشد آن گاه

الف. F یک تابع غیرنزولی و از راست پیوسته است.

ب. $F(-\infty) = 0$ و $F(+\infty) = 1$.

ج. حد چپ F در هر نقطه وجود دارد و برای هر $x \in \mathbb{R}$ داریم $F(x^-) = \mathrm{P}(X < x)$.

د. مجموعه نقاط ناپیوستگی F شماراست و برای هر $x \in \mathbb{R}$ داریم

$$\mathrm{P}(X = x) = F(x) - F(x^-)$$

به عبارت دیگر ناپیوستگی F به صورت پرشی بوده و ارتفاع پرش در x برابر با $\mathrm{P}(X = x)$ است.

ه. برای هر $x < y$، $\mathrm{P}(x < X \leq y) = F(y) - F(x)$.

اثبات. با توجه به نامنفی بودن تابع احتمال، اثبات غیرنزولی بودن تابع توزیع بدیهی است و از غیر نزولی بودن تابع F و براساس ویژگی های تابع احتمال، حد چپ و راست تابع توزیع در هر نقطه وجود داشته و متناهی هستند و داریم $F(-\infty)$ و $F(+\infty)$.

با توجه به وجود حد راست، برای اثبات پیوستگی از راست تابع F، لازم است ثابت کنیم که برای هر نقطه دلخواه $x \in \mathbb{R}$ رابطه زیر برقرار است

$$\lim_{n\to\infty} F(x + \frac{1}{n}) = F(x).$$

دنباله A_n، $n \geq ۱$ را به صورت $A_n = (X \leq x + \frac{۱}{n})$ تعریف می کنیم. واضح است که برای هر $n \geq ۱$، $A_{n+۱} \subset A_n$ و بنابراین این دنباله نزولی است. پس طبق پیوستگی تابع احتمال و (۱-۲) داریم

$$\begin{aligned}\lim_{n\to\infty} F(x+\frac{۱}{n}) &= \lim_{n\to\infty} \mathrm{P}(A_n) = \mathrm{P}\left(\bigcap_{n=۱}^{\infty} A_n\right)\\ &= \mathrm{P}\left(\bigcap_{n=۱}^{\infty}\left(X \leq x+\frac{۱}{n}\right)\right)\\ &= \mathrm{P}(X \leq x) = F(x).\end{aligned}$$

برای اثبات $F(+\infty) = ۱$، چون $F(+\infty)$ وجود دارد، داریم $F(+\infty) = \lim_{n\to\infty} F(n)$ و با استفاده از (۱-۱) بدست می‌آوریم

$$\begin{aligned}\lim_{n\to\infty} F(n) &= \lim_{n\to\infty} \mathrm{P}(X \leq n)\\ &= \mathrm{P}\left(\bigcup_{n=۱}^{\infty}(X \leq n)\right)\\ &= \mathrm{P}(X \leq +\infty) = \mathrm{P}(\Omega) = ۱\end{aligned}$$

همچنین بطور مشابه $F(-\infty) = \lim_{n\to-\infty} F(n)$ و از (۱-۲) داریم

$$\begin{aligned}\lim_{n\to-\infty} F(n) &= \lim_{n\to-\infty} \mathrm{P}(X \leq n)\\ &= \lim_{t\to\infty} \mathrm{P}(X \leq -t) = \mathrm{P}\left(\bigcap_{t=۱}^{\infty}(X \leq -t)\right)\\ &= \mathrm{P}(X \leq -\infty) = \mathrm{P}(\emptyset) = ۰.\end{aligned}$$

با توجه به وجود حد چپ، برای اثبات قسمت ج لازم است ثابت کنیم که برای هر نقطه دلخواه $x \in \mathbb{R}$ رابطه زیر برقرار است

$$\lim_{n\to\infty} F(x-\frac{۱}{n}) = \mathrm{P}(X < x).$$

دنباله A_n، $n \geq ۱$ را به صورت $A_n = (X \leq x - \frac{۱}{n})$ تعریف می کنیم. واضح است که برای هر $n \geq ۱$،

$A_n \subset A_{n+1}$ و بنابراین این دنباله صعودی است. پس از (۱-۱) داریم

$$\begin{aligned} \lim_{n\to\infty} F(x - \frac{1}{n}) &= \lim_{n\to\infty} \mathrm{P}(A_n) = \mathrm{P}\left(\bigcup_{n=1}^{\infty} A_n\right) \\ &= \mathrm{P}\left(\bigcup_{n=1}^{\infty}\left(X \leq x - \frac{1}{n}\right)\right) \\ &= \mathrm{P}(X < x). \end{aligned}$$

اثبات قسمت د و ه بعنوان تمرین رها می شود. □

نکته ۱-۲۶. هر تابع غیرنزولی و از راست پیوسته $F : \mathbb{R} \longrightarrow \mathbb{R}$ که برای آن $F(+\infty) = 1$ و $F(-\infty) = 0$، یک تابع توزیع نامیده می شود. می توان ثابت کرد که برای هر تابع توزیع داده شده F یک فضای احتمال $(\Omega, \mathcal{F}, \mathrm{P})$ و یک متغیرتصادفی X روی این فضا وجود دارد بطوریکه $F_X = F$.

اثبات. اثبات این نکته در مطالب این درس نمی گنجد. □

مثال ۱-۲۷. فرض کنید از بین اعداد ۱ تا n، $n > 2$، یک عدد به تصادفی انتخاب کنیم. در این صورت داریم $\Omega = \{1, 2, \cdots, n\}$ و مدل احتمال یکنواخت است. متغیرتصادفی X را عدد انتخاب شده تعریف می‌کنیم یعنی

$$X :\ \Omega \longrightarrow \mathbb{R}, \qquad X(\omega) = \omega.$$

با توجه به اینکه برای $k \leq x < k + 1$

$$F_X(x) = \mathrm{P}(X \leq x) = \mathrm{P}(\{\omega \in \Omega\ :\ X(\omega) \leq x\}) = \mathrm{P}\{1, 2, \cdots, k\} = \frac{k}{n}$$

بدست می آوریم

$$\begin{aligned} F_X(x) &= \mathrm{P}(X \leq x) = \mathrm{P}(\{\omega \in \Omega\ :\ X(\omega) \leq x\}) \\ &= \begin{cases} 0 & x < 1 \\ \frac{[x]}{n} & 1 \leq x < n \\ 1 & x \geq n \end{cases} \end{aligned}$$

مثال ۱-۲۸. فرض کنید یک عدد حقیقی از بازه $(0, 1)$ به تصادف انتخاب کنیم. یدر این صورت داریم $\Omega = (0, 1)$ و مدل احتمال یکنواخت است. متغیرتصادفی X را عدد انتخاب شده تعریف می کنیم یعنی

$$X :\ \Omega \longrightarrow \mathbb{R}, \qquad X(\omega) = \omega.$$

داریم

$$\begin{aligned} F_X(x) &= \mathrm{P}(X \le x) = \mathrm{P}(\{\omega \in \Omega \ : \ X(\omega) \le x\}) = \mathrm{P}\left((0, x] \cap (0, 1)\right) \\ &= \begin{cases} 0 & x < 0 \\ \frac{x}{1} = x & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases} \end{aligned}$$

در حالتی که با متغیرهای تصادفی سروکار داریم، چه متغیرتصادفی گسسته باشد و یا پیوسته، گاهی بجای تابع توزیع، با تابع چگالی احتمال و یا تابع چگالی کار می‌کنیم چراکه دانستن آن معادل با دانستن تابع توزیع متغیرتصادفی است. تعریف تابع چگالی برای متغیرهای تصادفی گسسته و پیوسته متفاوت است.

تعریف ۱-۲۹. فرض کنید X یک متغیرتصادفی گسسته باشد، تابع احتمال یا تابع چگالی آن را با f_X نشان داده که به صورت زیر تعریف می‌شود

$$f_X : \ \mathbb{R} \longrightarrow \mathbb{R}, \qquad f_X(x) = \mathrm{P}(X = x) = \mathrm{P}(\{\omega : \ X(\omega) = x\}), \ x \in \mathbb{R}.$$

و باز تاکید می کنیم که اندیس X صرفا جهت تاکید بر نام متغیرتصادفی است و می توان آن را حذف کرد.

قضیه ۱-۳۰. فرض کنید f و F به ترتیب تابع چگالی و تابع توزیع یک متغیرتصادفی گسسته X با مجموعه مقادیر $S_X = \{x_1, x_2, \cdots\}$ یا $S_X = \{x_1, x_2, \cdots, x_n\}$ باشد. در این صورت داریم

۱. f یک تابع نامنفی است و برای هر $x \in S_X^c$، $f(x) = \mathrm{P}(X = x) = 0$. درواقع تابع f فقط روی S_X یعنی تعداد نقاط حداکثر شمارش‌پذیر می‌تواند مثبت باشد.

۲. $\sum_x f(x) = \sum_{x \in S_X} f(x) = \sum_i f(x_i) = 1$

۳. برای هر $A \subset \mathbb{R}$، $\mathrm{P}(x \in A) = \sum_{x \in A} f(x) = \sum_{x \in A \cap S_X} f(x) = \sum_{i : x_i \in A \cap S_X} f(x_i)$

۴. برای هر $x \in \mathbb{R}$ داریم

$$\begin{aligned} F(x) &= \sum_{u \le x} f(u) = \sum_{u \le x, u \in S_X} f(x) = \sum_{i : x_i \le x} f(x_i) \\ f(x) &= \mathrm{P}(X = x) = F(x) - F(x^-) \end{aligned}$$

□

نکته ۱-۳۱. هر تابع با دو خاصیت اول قضیه ۱-۳۰ را یک تابع چگالی احتمال یا تابع چگالی و یا حتی تابع احتمال گویند. برای هر تابع چگالی احتمال f یک فضای احتمال $(\Omega, \mathcal{F}, \mathrm{P})$ و یک متغیرتصادفی X روی این فضا وجود دارد بطوریکه $f_X = f$.

مثال ۱-۳۲. فرض کنید جعبه ای شامل n مهره بوده که از ۱ تا n شماره گذاری شده اند. از این جعبه r مهره، $r < n$، را با هم و بدون توجه به ترتیب آن ها انتخاب می کنیم. اگر متغیر تصادفی X را ماکزیمم عدد انتخاب شده تعریف کنیم داریم

$$S_X = \{r, r+1, \cdots, n\}, \qquad f_X(x) = \begin{cases} \dfrac{\binom{1}{1} \times \binom{x-1}{r-1}}{\binom{n}{r}} & x = r, r+1, \cdots, n \\ 0 & \text{o.w} \end{cases}$$

تمرین ۱-۳۳. در مثال ۱-۲۷ تابع چگالی احتمال را بدست آورید.

تعریف ۱-۳۴. متغیر تصادفی X را پیوسته گویند اگر یک تابع نامنفی $f_X : \mathbb{R} \longrightarrow \mathbb{R}$ وجود داشته باشد بطوریکه

$$\mathrm{P}(X \in A) = \int_A f_X(x)\mathrm{d}x \qquad A \subset \mathbb{R}$$

و f_X تابع چگالی احتمال X نامیده می شود.

قضیه ۱-۳۵. فرض کنید f و F به ترتیب تابع چگالی احتمال و تابع توزیع متغیر تصادفی پیوسته X باشد. در اینصورت

۱. برای هر $x \in \mathbb{R}$، $\mathrm{P}(X = x) = 0$

۲. $\int_{\mathbb{R}} f(x)\mathrm{d}x = 1$

۳. برای هر $x \in \mathbb{R}$،

$$F(x) = \mathrm{P}(X \leq x) = \int_{-\infty}^{x} f(t)\mathrm{d}t, \qquad x \in \mathbb{R}$$
$$f(x) = \frac{d}{dx}F(x) = F'(x), \qquad x \in \mathbb{R}.$$

مثال ۱-۳۶. متغیرتصادفی X تعریف شده در مثال ۱-۲۸ یک متغیرتصادفی پیوسته است که برای آن $S_X = (0, 1)$ و

$$f_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{o.w} \end{cases}$$

تعریف ۱-۳۷. بطورکلی هر تابع نامنفی $f : \mathbb{R} \longrightarrow \mathbb{R}$ بطوریکه $\int_{\mathbb{R}} f(x)\mathrm{dx} = 1$ را یک تابع چگالی احتمال گویند.

توجه داشته باشید که تابع احتمال یک متغیرتصادفی گسسته همچون X، در هر نقطه x برابر با احتمال پیشامد $X = x$ است و بنابراین همواره عددی بین صفر و یک خواهد بود. این درحالی است که اگر X یک متغیرتصادفی پیوسته باشد، تابع چگالی آن در هر نقطه دیگر برابر با احتمال پیشامد $X = x$ نیست و می‌تواند در بسیاری از نقاط عددی بزرگتر از یک باشد.

در جدول ۱.۱ برخی توزیع های شناخته شده نشان داده شده اند.

جدول ۱.۱: برخی توزیع های شناخته شده

توزیع	تابع چگالی	پارامترها
$X \sim N(\mu, \sigma^2)$	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{\frac{-1}{2\sigma^2}(x-\mu)^2\}; \quad x \in \mathbb{R}$	$\mu \in \mathbb{R}$ $\sigma > 0$
$X \sim E(\lambda)$	$f(x) = \lambda e^{-\lambda x}; \quad x > 0$	$\lambda > 0$
$X \sim Gamma(r, \lambda)$	$f(x) = \frac{r^\lambda}{\Gamma(r)} e^{-\lambda x} x^{r-1}; \quad x > 0$	$r, \lambda > 0$
$X \sim U(a, b)$	$f(x) = \frac{1}{b-a}; \quad a < x < b$	$-\infty < a < b < \infty$
$X \sim Beta(\alpha, \beta)$	$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(a)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}; \quad 0 < x < 1$	$\alpha, \beta > 0$
$X \sim B(n, p)$	$f(x) = \begin{pmatrix} n \\ x \end{pmatrix} p^x (1-p)^{n-x}; \quad x = 0, 1, \cdots, n$	$0 < p < 1$ $n \in \mathbb{N}$
$X \sim NB(\alpha, p)$	$f(x) = \begin{pmatrix} \alpha + x - 1 \\ x \end{pmatrix} p^\alpha (1-p)^x; \quad x = 0, 1, \cdots$	$0 < p < 1$ $\alpha > 0$
$X \sim G(p)$	$f(x) = p(1-p)^x; \quad x = 0, 1, \cdots$	$0 < p < 1$
$X \sim P(\lambda)$	$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}; \quad x = 0, 1, \cdots$	$\lambda > 0$

۲.۳.۱ بردارهای تصادفی

تعریف ۱-۳۸. اگر X_i، $i = ۱, ۲, \cdots, n$، متغیرهای تصادفی باشند آن گاه $\mathbf{X} = (X_۱, X_۲, \cdots, X_n)$ یک بردار تصادفی نامیده می شود. بردار تصادفی n بعدی $\mathbf{X}$ را می توان به صورت تابع $\mathbf{X} : \Omega \longrightarrow \mathbb{R}^n$ تعریف کرد بطوریکه برای هر زیرمجموعه $A \subset \mathbb{R}^n$ داشته باشیم

$$(\mathbf{X} \in A) = (\omega : \mathbf{X}(\omega) \in A) \in \mathcal{F}$$

برد یک بردار تصادفی $\mathbf{X}$ را مجموعه مقادیر یا تکیه گاه آن گوییم و آن را با $S_{\mathbf{X}}$ نشان می دهیم.

هرگاه مجموعه مقادیر $S_{\mathbf{X}}$ شمارش پذیر (متناهی یا نامتناهی) باشد، بردارتصادفی $\mathbf{X}$ یک بردارتصادفی گسسته نامیده می شود.

خواص احتمالی یا اماری یک بردارتصادفی در توزیع آن نهفته است. توزیع یک بردار تصادفی وقتی معلوم است که احتمال تمام پیشامدهای مربوط به آن معلوم باشد. درواقع احتمالات زیر معلوم باشد

$$\mathrm{P}(\mathbf{X} \in A), \quad \forall A \subset \mathbb{R}^n.$$

در حالت کلی کار با احتمالات بالا ساده نیست و در محاسبات آن ها دچار چالش‌هایی خواهیم بود. به همین دلیل به جای بررسی احتمالات بالا، در حالتی که بردار تصادفی داریم با **تابع توزیع توام** کار می کنیم. در واقع دانستن تابع توزیع توام معادل با دانستن توزیع بردار تصادفی است.

تعریف ۱-۳۹. فرض کنید $\mathbf{X} = (X_۱, X_۲, \cdots, X_n)$ یک بردار تصادفی روی فضای احتمال $(\Omega, \mathcal{F}, \mathrm{P})$ باشد. تابع توزیع توام این بردارتصادفی که با $F_{\mathbf{X}}$ (و یا بدون اندیس با F) نشان داده می شود بصورت زیر تعریف می شود

$$F : \mathbb{R}^n \longrightarrow \mathbb{R}, \qquad F(\mathbf{x}) = \mathrm{P}(\mathbf{X} \leq \mathbf{x}) = \mathrm{P}(X_۱ \leq x_۱, X_۲ \leq x_۲, \cdots, X_n \leq x_n).$$

قضیه ۱-۴۰. اگر F تابع توزیع توام یک بردار تصادفی باشد. آن‌گاه

۱. تابع F غیرنزولی است و درواقع اگر برای $i = ۱, ۲, \cdots, n$ داشته باشیم $x_i \leq y_i$ آن‌گاه $F(x_۱, \cdots, x_n) \leq F(y_۱, \cdots, y_n)$.

۲. تابع F از راست پیوسته است، یعنی

$$\lim_{h \to ۰^+} F(x_۱ + h, \cdots, x_n + h) = F(x_۱, \cdots, x_n).$$

۳. $\lim_{x_i\to\infty, i=1,\cdots,n} F(x_1,\cdots,x_n)=1$ و اگر حداقل یکی از متغیرهای x_i به $-\infty$ میل کند حد F صفر خواهد بود و داریم

$$\lim_{\exists i=1,\cdots,n;\ x_i\to-\infty} F(x_1,\cdots,x_n)=0$$

۴. حد چپ F در هر نقطه وجود دارد و برای هر $x\in\mathbb{R}$ داریم

$$\lim_{h\to 0^+} F(x_1-h,\cdots,x_n-h)=\mathrm{P}(X_1<x_1,\cdots,X_n<x_n)$$

با توجه به اینکه دانستن تابع احتمال توام و یا چگالی احتمال توام معادل با دانستن تابع توزیع توام است، در حالتی که در مورد بردارهای تصادفی گسسته و یا پیوسته بحث می کنیم به جای تابع توزیع توام، با تابع احتمال توام و تابع چگالی احتمال توام کار می کنیم. مطابق با تعریف، یک بردار تصادفی $\mathbf{X}$ گسسته است هرگاه مجموعه مقادیر آن شمارش پذیر باشد.

تعریف ۱-۴۱. فرض کنید $\mathbf{X}=(X_1,X_2,\cdots,X_n)$ یک بردار تصادفی گسسته باشد. تابع احتمال $\mathbf{X}$ را با f، $f:\mathbb{R}^n\longrightarrow\mathbb{R}$، نشان داده که به صورت زیر تعریف می شود

$$f=\mathrm{P}(\mathbf{X}=\mathbf{x})=\mathrm{P}(X_1=x_1,X_2=x_2,\cdots,X_n=x_n)\qquad \mathbf{x}=(x_1,x_2,\cdots,x_n)\in\mathbb{R}^n.$$

قضیه ۱-۴۲. فرض کنید f و F به ترتیب تابع احتمال توام و تابع توزیع توام یک بردار تصادفی همچون $\mathbf{X}=(X_1,X_2,\cdots,X_n)$ با مجموعه مقادیر $S_X=\{\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_n\}$ باشد. داریم

۱. f یک تابع نامنفی است و برای هر $\mathbf{x}\in S_X^c$ داریم $f(\mathbf{x})=\mathrm{P}(\mathbf{X}=\mathbf{x})=0$، به عبارت دیگر تابع f فقط روی S_X یعنی تعداد شمارش‌پذیری از نقاط می‌تواند مثبت باشد.

۲. $\sum_{\mathbf{x}} f(\mathbf{x})=\sum_{\mathbf{x}\in S_{\mathbf{X}}} f(\mathbf{x})=\sum_i f(\mathbf{x}_i)=1$

۳. برای هر $A\subset\mathbb{R}^n$ داریم

$$\mathrm{P}(\mathbf{X}\in A)=\sum_{\mathbf{x}\in A} f(\mathbf{x})=\sum_{\mathbf{x}\in A\cap S_{\mathbf{X}}} f(\mathbf{x})=\sum_{i:x_i\in A\cap S_{\mathbf{X}}} f(\mathbf{x}_i).$$

۴. برای هر $x\in\mathbb{R}$ داریم

$$\begin{aligned}F(\mathbf{x})=F(x_1,\cdots,x_n)&=\sum_{u_1\le x_1}\sum_{u_2\le x_2}\cdots\sum_{u_n\le x_n} f(u_1,u_2,\cdots,u_n)\\&=\sum\cdots\sum\nolimits_{u_1\le x_1,\cdots,u_n\le x_n,(u_1,\cdots,u_n)\in S_{\mathbf{X}}} f(u_1,u_2,\cdots,u_n)\end{aligned}$$

۵. برای هر $i = 1, 2, \cdots, n$ تابع احتمال X_i بصورت زیر بدست می‌آید

$$f_{X_i}(x) = \sum_{x_1}\sum_{x_2}\cdots\sum_{x_{i-1}}\sum_{x_{i+1}}\cdots\sum_{x_n} f(x_1, x_2, \cdots, x_{i-1}, x, x_{i+1}, \cdots, x_n).$$

رابطه مشابه برای زیربردارهای یک بردار تصادفی مثل $\mathbf{X}$ وجود دارد. بعنوان مثال برای زیربردار (X_1, X_3, X_4) داریم

$$f_{(X_1,X_3,X_4)}(x_1, x_3, x_4) = \sum_{x_2}\sum_{x_5}\cdots\sum_{x_n} f(x_1, x_2, \cdots, x_n)$$

یعنی روی هر تعداد آرگومان جمع ببندیم تابع احتمال توام بردارتصادفی متناظر با آرگومان‌های باقیمانده بدست می‌آید.

□

تعریف ۱-۴۳. بطورکلی هر تابع با دو خاصیت اول قضیه ۱-۴۲ را یک تابع احتمال توام می نامند.

مثال ۱-۴۴. فرض کنید جعبه ای شامل N_1 مهره سفید، N_2 مهره سیاه و N_3 مهره قرمز است. از این جعبه n مهره به تصادف انتخاب می‌کنیم. اگر X_1، X_2 و X_3 را به ترتیب تعداد مهره های سفید، سیاه و قرمز در n توپ انتخابی تعریف کنیم آن‌گاه تابع احتمال توام بردارتصادفی (X_1, X_2, X_3) به صورت زیر خواهد بود

$$f(x_1, x_2, x_3) = \frac{\binom{N_1}{x_1}\binom{N_2}{x_3}\binom{N_3}{x_3}}{\binom{N_1+N_2+N_3}{n}}, \quad 0 \le x_i \le N_i, \quad i = 1, 2, 3; \quad 0 \le x_1+x_2+x_3 \le n.$$

تعریف ۱-۴۵. بردارتصادفی $\mathbf{X}$ را پیوسته گویند هرگاه یک تابع نامنفی $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ وجود داشته باشد بطوری‌که

$$\mathrm{P}(\mathbf{X} \in A) = \int_A f(\mathbf{x})\mathbf{dx}, \qquad A \subset \mathbb{R}^n$$

به تابع f تابع چگالی توام یا تابع چگالی احتمال توام $\mathbf{X}$ گفته می‌شود.

قضیه ۱-۴۶. فرض کنید f و F به ترتیب تابع چگالی توام و تابع توزیع توام بردارتصادفی پیوسته $\mathbf{X}$ باشد. در اینصورت

۱. برای هر $\mathbf{x} \in \mathbb{R}^n$، $\mathrm{P}(\mathbf{X} = \mathbf{x}) = 0$.

۲. $\int_{\mathbb{R}^n} f(\mathbf{x})\mathbf{dx} = 1$.

۳. برای هر $x \in \mathbb{R}$،

$$F(\mathbf{x}) = F(x_1, x_2, \cdots, x_n) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_n} f(t_1, t_2, \cdots, t_n)\mathrm{d}t_1\mathrm{d}t_2\cdots\mathrm{d}t_n, \quad \mathbf{x} \in \mathbb{R}^n$$

$$f(\mathbf{x}) = f(x_1, x_2, \cdots, x_n) = \frac{\partial^n}{\partial x_1 \partial x_2 \ldots \partial x_n} F(x_1, x_2, \cdots, x_n), \qquad \mathbf{x} \in \mathbb{R}^n.$$

۴. برای هر $i = 1, 2, \ldots, n$ تابع چگالی X_i به صورت زیر بدست می‌آید

$$f_{X_i}(x) = \int_{\mathbb{R}}\int_{\mathbb{R}}\cdots\int_{\mathbb{R}} f(x_1, x_2, \cdots, x_{i-1}, x, x_{i+1}, \cdots, x_n)\mathrm{dx}_1\mathrm{dx}_2\ldots\mathrm{dx}_{i-1}\mathrm{dx}_{i+1}\ldots\mathrm{dx}_n.$$

رابطه مشابه برای زیربردارهای یک بردار تصادفی مثل $\mathbf{X}$ وجود دارد. بعنوان مثال برای زیربردار (X_1, X_3, X_4) داریم

$$f_{(X_1,X_3,X_4)}(x_1, x_3, x_4) = \int_{\mathbb{R}}\cdots\int_{\mathbb{R}} f(x_1, x_2, \cdots, x_n)\mathrm{dx}_2\mathrm{dx}_5\mathrm{dx}_6\ldots\mathrm{dx}_n.$$

یعنی روی هر تعداد آرگومان انتگرال بگیریم تابع چگالی توام بردارتصادفی متناظر با آرگومان‌های باقیمانده بدست می‌آید.

$\square$

تعریف ۱-۴۷. بطورکلی هر تابع نامنفی $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ با شرط $\int_{\mathbb{R}^n} f(\mathbf{x})\mathbf{dx} = 1$ را یک تابع چگالی احتمال توام یا تابع چگالی توام گویند.

مثال ۱-۴۸. فرض کنید نقطه ای به تصادف از داخل مربع واحد $[0, 1] \times [0, 1]$ انتخاب کنیم. اگر X و Y به ترتیب طول و عرض نقطه انتخاب شده تعریف شوند، آن‌گاه می‌توان نشان داد (تمرین؟) که (X, Y) یک بردارتصادفی پیوسته با تابع چگالی توام زیر است

$$f(x, y) = 1 \qquad 0 < x, y < 1.$$

تعریف ۱-۴۹. فرض کنید $\mathbf{X} = (X_1, X_2, \cdots, X_n)$ یک بردارتصادفی باشد. به تابع توزیع و تابع چگالی هر یک از متغیرهای تصادفی X_i، $i = 1, 2, \cdots, n$، به ترتیب تابع توزیع حاشیه‌ای و تابع چگالی حاشیه‌ای X_i گفته می‌شود.

دو قضیه ۱-۴۲ و ۱-۴۶ را می‌توان بصورت قضیه زیر تعمیم داد.

قضیه ۱-۵۰. یک بردارتصادفی $(\mathbf{X}, \mathbf{Y})$ را درنظر بگیرید که در آن $\mathbf{X}$ یک بردارتصادفی n بعدی و $\mathbf{Y}$ یک بردار تصادفی m بعدی است.

۱. فرض کنید $F(x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_m)$ تابع توزیع توام بردار تصادفی $(\mathbf{X}, \mathbf{Y})$ باشد، در اینصورت تابع توزیع توام بردار تصادفی $\mathbf{X}$ بصورت زیر محاسبه می‌شود

$$\begin{aligned} F_X(\mathbf{x}) &= F(x_1, x_2, \cdots, x_n, \infty, \infty, \cdots, \infty) \\ &= \lim_{y_1 \to \infty} \cdots \lim_{y_m \to \infty} F(x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_m). \end{aligned}$$

۲. اگر بردارتصادفی $(\mathbf{X}, \mathbf{Y})$ یک بردارتصادفی پیوسته با تابع چگالی توام $f(\mathbf{x}, \mathbf{y})$ باشد آن‌گاه بردارتصادفی $\mathbf{X}$ نیز یک بردار تصادفی پیوسته با تابع چگالی احتمال توام زیر است

$$f_X(\mathbf{x}) = \int_{\mathbb{R}^m} f(\mathbf{x}, \mathbf{y}) \mathbf{dy}; \qquad \mathbf{x} \in \mathbb{R}^n.$$

۲. اگر بردارتصادفی $(\mathbf{X}, \mathbf{Y})$ یک بردارتصادفی گسسته با تابع چگالی توام $f(\mathbf{x}, \mathbf{y})$ باشد آن‌گاه بردارتصادفی $\mathbf{X}$ نیز یک بردار تصادفی گسسته با تابع چگالی احتمال توام زیر بوده

$$f_X(\mathbf{x}) = \sum_{\mathbf{y}:(\mathbf{x},\mathbf{y}) \in S_{\mathbf{X},\mathbf{Y}}} f(\mathbf{x}, \mathbf{y}) \mathbf{dy}; \qquad \mathbf{x} \in \mathbb{R}^n.$$

که در آن $S_{\mathbf{X},\mathbf{Y}}$ مجموعه مقادیر بردارتصادفی $(\mathbf{X}, \mathbf{Y})$ است.

$\square$

نکته ۱-۵۱. یادآوری می کنیم که اگر $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ یک بردارتصادفی و $g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ یک تابع باشد آن‌گاه تابع $Y : \Omega \longrightarrow \mathbb{R}^m$ تعریف شده به شکل

$$\mathbf{Y} = (Y_1, Y_2, \cdots, Y_m) = g(\mathbf{X}) = g(X_1, X_2, \cdots, X_n)$$

نیز یک بردار تصادفی است. درواقع هر تابع از یک بردار یا متغیر تصادفی یک بردار یا متغیرتصادفی خواهد بود و داریم

$$\mathbf{Y}(\omega) = (Y_1(\omega), Y_2(\omega), \cdots, Y_m(\omega)) = g(X_1(\omega), X_2(\omega), \cdots, X_n(\omega)) \qquad \omega \in \Omega.$$

۴.۱ توزیع توابعی از متغیرهای تصادفی

یکی از مهمترین مسائل نظریه احتمال و آمار، تعیین توزیع تابعی معلوم از بردارهای تصادفی است. فرض کنید $\mathbf{X}$ یک بردارتصادفی با توزیع معلوم باشد و $\mathbf{Y} = g(\mathbf{X})$. داریم

$$\begin{aligned}
\mathrm{P}(\mathbf{Y} \in A) &= \mathrm{P}(g(\mathbf{X}) \in A) \\
&= \mathrm{P}(g(X_۱, X_۲, \cdots, X_n) \in A) \\
&= \mathrm{P}\left((X_۱, X_۲, \cdots, X_n) \in \{\mathbf{x} \in \mathbb{R}^n : \ g(\mathbf{x}) \in A\}\right) \\
&= \begin{cases} \sum_{\mathbf{x} \in S_{\mathbf{X}} : g(\mathbf{x}) \in A} f_{\mathbf{X}}(\mathbf{x}) \\ \int_{\{\mathbf{x} \in S_{\mathbf{X}} : g(\mathbf{x}) \in A\}} f_{\mathbf{X}}(\mathbf{x}) \mathbf{dx} \end{cases} \qquad (۵-۱)
\end{aligned}$$

در حالتی که $\mathbf{X}$ گسسته باشد آن‌گاه $\mathbf{Y}$ نیز بردارتصادفی گسسته است و تابع چگالی توام آن بصورت زیر بدست می‌آید

$$\mathrm{P}(\mathbf{Y} = \mathbf{y}) = \sum_{\mathbf{x} \in S_{\mathbf{X}} : g(\mathbf{x}) = \mathbf{y}} f_{\mathbf{X}}(\mathbf{x}).$$

از آنچایی که اغلب مجموعه $\{\mathbf{x} \in \mathbb{R}^n : \ g(\mathbf{x}) \in A\}$ خیلی پیچیده است، محاسبه انتگرال و مجموع موجود در رابطه (۱-۵) آسان نیست. اما در حالتی که تابع g یک به یک باشد، این محاسبات بسیار راحت خواهد بود. در ادامه این مسائل را ابتدا در حالت تک متغیر و سپس بردار تصادفی مطرح می‌کنیم.

قضیه ۱-۵۲. فرض کنید X یک متغیرتصادفی با تابع توزیع F_X باشد. همچنین فرض کنید $Y = g(X)$ و g یک تابع یک به یک با معکوس $g^{-۱}$ باشد. در این صورت

$$S_Y = g(S_X) = \{y \in \mathbb{R} : \ \exists x \in S_X \ \mathbf{s.t} \ y = g(x)\}$$

و داریم

۱. اگر تابع g اکیدا صعودی باشد آن‌گاه

$$F_Y(y) = F_X(g^{-۱}(y)), \qquad y \in \mathbb{R}$$

و اگر تابع g اکیدا نزولی باشد آن‌گاه

$$F_Y(y) = ۱ - F_X(g^{-۱}(y)^-), \qquad y \in \mathbb{R}$$

۲. اگر X یک متغیرتصادفی گسسته با تابع احتمال f_X باشد آن‌گاه Y نیز یک متغیرتصادفی گسسته با تابع احتمال زیر است

$$F_Y(y) = F_X(g^{-1}(y)), \qquad y \in S_y.$$

اگر X یک متغیرتصادفی پیوسته با تابع چگالی احتمال f_X و g یک تابع مشتق‌پذیر باشد و یک زیرمجموعه باز $U \subset \mathbb{R}$ وجود داشته باشد بطوری‌که $\mathrm{P}(X \in U) = 1$ و برای هر $x \in U$، $g'(x) > 0$. اگر

$$V = g(U) = \{y \in \mathbb{R} : \ \exists x \in U \text{ s.t } y = g(x)\}$$

آن‌گاه Y نیز یک متغیرتصادفی پیوسته با تابع چگالی زیر است

$$f_Y(y) = |\frac{d}{dy}g^{-1}(y)| f_X(g^{-1}(y)), \qquad y \in V.$$

□

حال فرض کنید $\mathbf{X} = (X_1, X_2, \cdots, X_n)$ یک بردارتصادفی n بعدی و $g : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ یک تابع یک به یک باشد.

برای بدست آوردن توزیع $\mathbf{Y} = g(\mathbf{X})$، در حالتی که $\mathbf{X}$ گسسته باشد آن‌گاه $\mathbf{Y}$ نیز گسسته است و

$$S_{\mathbf{Y}} = g(S_{\mathbf{X}}) = \{\mathbf{y} \in \mathbb{R}^n : \ \exists \mathbf{x} \in S_{\mathbf{X}} \text{ s.t } \mathbf{y} = g(\mathbf{x})\}$$

بنابراین بدست می‌آوریم

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(g^{-1}(\mathbf{y})) \qquad \mathbf{y} \in S_{\mathbf{Y}}.$$

در حالتی که $\mathbf{X}$ پیوسته باشد، فرض کنید $g : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ یک تابع یک به یک با مشتق پیوسته باشد. همچنین فرض کنید $U \subset \mathbb{R}^n$ یک مجموعه باز باشد بطوری که $\mathrm{P}(\mathbf{X} \in U) = 1$. در این صورت داریم

$$f_{\mathbf{Y}}(\mathbf{y}) = \left|\frac{\partial g^{-1}(\mathbf{x})}{\partial \mathbf{x}}\right| f_{\mathbf{X}}(g^{-1}(\mathbf{y}) \qquad \mathbf{y} \in V$$

بطوری که $V = g(U) = \{\mathbf{y} \in \mathbb{R}^n : \ \exists \mathbf{x} \in U \text{ s.t. } \mathbf{y} = g(\mathbf{x})\}$ و $\left|\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}}\right|$ قدرمطلق دترمینان ماتریس جاکوبی $g = (g_1, \ldots, g_n)$ در $\mathbf{x}$ است وقتی مولفه (i, j) ماتریس جاکوبی g برابر با $\frac{\partial g_i}{\partial x_i}$ است.

تمرین ۱-۵۳. اگر $X_1, X_2 \sim E(\lambda)$، تابع چگالی متغیر تصادفی $X_1 + X_2$ را بدست آورید.

۵.۱ امیدریاضی و خواص آن

فرض کنید X یک متغیر تصادفی باشد. امیدریاضی متغیرتصادفی X که با $E(X)$ نمایش داده می‌شود، بصورت زیر تعریف می‌شود

- اگر X گسسته باشد، $\mathrm{E}(X)=\sum_{x\in S_X} xf_X(x)$
- اگر X پیوسته باشد، $\mathrm{E}(X)=\int_{\mathbb{R}} xf_X(x)\mathrm{dx}$

توجه داشته باشید که شرط وجود امیدریاضی متغیر تصادفی X به شکل زیر است

- اگر X گسسته بوده لازم است نامساوی $\sum_{x\in S_X} |x|f_X(x)<\infty$ برقرار باشد
- اگر X پیوسته بوده لازم است نامساوی $\int_{\mathbb{R}} |x|f_X(x)\mathrm{dx}<\infty$ برقرار باشد

در بعضی مسائل نیاز به محاسبه امیدریاضی تابعی از X، مثل $g(X)$ داریم. واضح است که $Y=g(X)$ یک متغیر یا بردار تصادفی است و یک راه محاسبه $\mathrm{E}(Y)=\mathrm{E}(g(X))$ پیدا کردن توزیع Y و محاسبه $\mathrm{E}(Y)$ با استفاده از آن است. اما یک راه ساده‌تر استفاده از قضیه زیر است.

قضیه ۱-۵۴. فرض کنید $\mathbf{X}=(X_۱,X_۲,\cdots,X_n)$ یک بردار تصادفی و $g:\mathbb{R}^n\longrightarrow\mathbb{R}$ یک تابع دلخواه باشد. در این صورت

- اگر $\mathbf{X}$ گسسته بوده و $\sum_{x\in S_{\mathbf{X}}}|g(\mathbf{x})|f(\mathbf{x})<\infty$ آنگاه امیدریاضی $g(\mathbf{X})$ وجود دارد (و متناهی است) و داریم
$$\mathrm{E}(g(\mathbf{X}))=\sum_{x\in S_{\mathbf{X}}} g(\mathbf{x})f(\mathbf{x})$$
- اگر $\mathbf{X}$ پیوسته بوده و $\int_{\mathbb{R}^n}|g(\mathbf{x})|f(\mathbf{x})\mathbf{dx}<\infty$ آنگاه امیدریاضی $g(\mathbf{X})$ وجود دارد (و متناهی است) و داریم
$$\mathrm{E}(g(\mathbf{X}))=\int_{\mathbb{R}^n} g(\mathbf{x})f(\mathbf{x})\mathbf{dx}$$

نکته ۱-۵۵. توجه داشته باشید که امیدریاضی $g(\mathbf{X})$ وجود دارد اگر و تنها اگر $\mathrm{E}(|g(\mathbf{X}|)<\infty$. از این جا به بعد هرگاه خاصیتی برای امیدریاضی بیان شود منوط به وجود امیدریاضی ایت مگر آنکه متغیرتصادفی (و یا تابع $g(\mathbf{X})$) نامنفی باشد که در این حالت امیدریاضی همیشه تعریف می شود ولی ممکن است بینهایت شود.

قضیه ۱-۵۶. امیدریاضی دارای خواص زیر است:

۰.۱ اگر X یک متغیرتصادفی نامنفی باشد آن‌گاه $\mathrm{E}(X) \geq ۰$ و $\mathrm{E}(X) = ۰$ اگر و تنها اگر $\mathrm{P}(X = ۰) = ۱$.

۰.۲ اگر $\mathbf{X}$ و $\mathbf{Y}$ دو بردار تصادفی به ترتیب n و m بعدی با تابع چگالی احتمال توام $f(\mathbf{x}, \mathbf{y})$ باشد آن‌گاه برای هر تابع حقیقی $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ داریم

– برای $\mathbf{X}$ گسسته

$$\mathrm{E}(g(\mathbf{X})) = \sum_{(\mathbf{x},\mathbf{y}) \in S_{\mathbf{X},\mathbf{Y}}} g(\mathbf{x}) f(\mathbf{x}, \mathbf{y})$$

– برای $\mathbf{X}$ پیوسته

$$\mathrm{E}(g(\mathbf{X})) = \int_{\mathbb{R}^m} \int_{\mathbb{R}^n} g(\mathbf{x}) f(\mathbf{x}, \mathbf{y}) \mathbf{dxdy}$$

۰.۳ اگر $X_۱, \cdots, X_k$ متغیرهای تصادفی دلخواه بوده و برای $i = ۱, \cdots, k$، $g_i : \mathbb{R} \longrightarrow \mathbb{R}$ توابع حقیقی باشند، آن گاه برای هر $a_۱, \cdots, a_k \in \mathbb{R}$ داریم

$$\mathrm{E}(\sum_{i=۱}^{k} g_i(X_i)) = \sum_{i=۱}^{k} \mathrm{E}(g_i(X_i)).$$

۰.۴ اگر $X_۱, X_۲, \cdots$ دنباله‌ای از متغیرهای تصادفی نامنفی باشد آن‌گاه

$$\mathrm{E}(\sum_{n=۱}^{\infty} X_n) = \sum_{n=۱}^{\infty} \mathrm{E}(X_n).$$

۰.۵ اگر X یک متغیرتصادفی نامنفی باشد آن‌گاه

$$\mathrm{E}(X) = \int_۰^{\infty} \mathrm{P}(X > x) \mathrm{dx} = \int_۰^{\infty} \mathrm{P}(X \geq x) \mathrm{dx}.$$

۰.۶ اگر X یک متغیرتصادفی با مجموعه مقادیر $\{۰, ۱, \cdots\}$ باشد آن‌گاه

$$\mathrm{E}(X) = \sum_{x=۰}^{\infty} \mathrm{P}(X > x).$$

۰.۷ اگر X و Y دو متغیرتصادفی باشند، آن‌گاه

$$\mathrm{E}(|XY|) \leq \sqrt{\mathrm{E}(X^۲)} \sqrt{\mathrm{E}(Y^۲)}$$

۰.۸ اگر X یک متغیر تصادفی نامنفی باشد آن‌گاه برای هر $a > ۰$ داریم

$$\mathrm{P}(X > a) \leq \frac{\mathrm{E}(X)}{a}$$

فصل ۲

معرفی فرایندهای تصادفی

اغلب پدیده های تصادفی مرتبط با طبیعت، صنعت، اقتصاد و ...، در طی زمان جریان دارند. فرایندهای تصادفی یک مدل ریاضی برای توضیح رویدادهای این نوع از پدیده‌های تصادفی در طی زمان است. فرایندهای تصادفی مهم‌ترین ابزار در مدل‌سازی این پدیده‌ها در اقتصاد و مباحث مالی، زیست‌شناسی، نجوم و ... است.

براساس تعریف، یک فرآیند تصادفی دنباله ای از متغیرهای تصادفی X_t است که روی یک فضای احتمال تعریف شده و برحسب یک اندیس که معمولا اندیس زمان است مرتب شده است. اندیس زمان مقادیر خود را از مجموعه اندیس‌گذاری مانند T اختیار می‌کند. بنابراین

$$\{X_t\,;\quad t\in T\}$$

را یک فرایند تصادفی می‌نامند.

فضای احتمال یک پدیده تصادفی که طی زمان جریان دارد اغلب بسیار پیچیده بوده و فرایند تصادفی در حقیقت یکی از جنبه‌های عددی این پدیده تصادفی طی زمان است. درواقع، می‌توان گفت که فرایندهای تصادفی یک وضعیت یا حالت پدیده تصادفی را طی زمان مدل‌سازی می‌کند. مثلا می‌توان گفت که X_t وضعیت یا حالت فرایند در زمان t است. مقدار فرایند در لحظه t، مقدار مشاهده x_t است.

اگرچه ممکن است اندیس t ربطی به زمان نداشته باشد ولی بطور کلی به t اندیس زمان گفته می شود. مجموعه اندیس گذار T ممکن است شمارا باشد (یعنی حالت‌های پدیده تصادفی گسسته ثبت شده و یا اساسا طبیعت زمان بصورت گسسته است) که به آن مجموعه گسسته و به فرایند تعریف شده بر روی آن فرایند زمان

گسسته گفته می‌شود. در حالتی که T یک فاصله حقیقی یا اجتماعی از یک سری فاصله های حقیقی باشد T را پیوسته و فرایند متناظر با آن را فرایند زمان پیوسته می‌نامند. اغلب $T = [0, \infty)$، $T = \mathbb{N}$ و یا $T = \mathbb{N}_0 = \{0, 1, 2, \cdots\}$ است. در کتاب‌ها و جزوه‌های آماری گاهی در یک فرایند تصادفی با $T = \mathbb{N}$ یا $T = \mathbb{N}_0$ اندیس زمان را بجای t با n نمایش می‌دهند یعنی فرایند را به صورت $\{X_n;\ n \in T\}$ می‌نویسند.

توجه داشته باشید که در عمل متغیرهای تصادفی X_t مستقل نیستند و معمولا ارتباطی بین آن‌ها وجود دارد.

همانطور که گفته شد فرآیندهای تصادفی برای مدلسازی بسیاری از فرآیندها در علوم مختلف فیزیک، مهندسی، بیولوژی و غیره کاربرد گسترده دارد. چند مثال ببینیم.

مثال ۲-۱. فرض کنید یک سکه را پی در پی پرتاب کنیم و X_n نتیجه پرتاب n-ام تعریف شود. در این صورت $T = \{1, 2, 3, \cdots\}$ و بنابراین $\{X_n;\ n \in T\}$ یک فرایند زمان گسسته است.

مثال ۲-۲. در آزمایش پرتاب متوالی یک سکه فرض کنید X_n تعداد دفعاتی باشد که برای nامین مرتبه رو مشاهده شود. در این صورت داریم $T = \{1, 2, \cdots\}$ و بنابراین $\{X_n;\ n \in T\}$ یک فرایند زمان گسسته است. واضح است که در این مثال X_n ها از هم مستقل نیستند و با یکدیگر در ارتباط هستند.

مثال ۲-۳. فرض کنید X_t تعداد ماشین‌های پارک شده در لحظه t در یک پارکینگ را نشان دهد. در این صورت $T = [0, \infty)$ و بنابراین $\{X_t;\ t \in T\} = \{X_t;\ t \geq 0\}$ یک فرایند تصادفی زمان پیوسته است.

مثال ۲-۴. شخصی یک سکه که احتمال مشاهده شیر برای آن برابر با p است را بطور متوالی پرتاب می‌کند و هر بار اگر شیر مشاهده کند یک گام به جلو و اگر خط مشاهده کند یک گام به عقب برمی‌گردد. فرض کنید X_t مکان شخص پس از پرتاب t-ام سکه باشد. در این مثال که به قدم زدن تصادفی معروف است داریم $T = \{1, 2, \cdots\}$ و بنابراین $\{X_t;\ t \in T\}$ یک فرایند تصادفی زمان گسسته است.

در یک فرایند تصادفی $\{X_t;\ t \in T\}$، مجموعه همه مقادیری که متغیرهای تصادفی X_t اختیار می کنند را فضای وضعیت فرآیند می‌نامند که معمولا با S نمایش داده می‌شود و می‌توان نوشت

$$S_X = \bigcup_{t \in T} S_{X_t}.$$

اگر S گسسته باشد، یک فرایند با وضعیت گسسته خواهیم داشت.

- یک فرایند تصادفی زمان گسسته با وضعیت گسسته، یک زنجیر تصادفی نامیده می‌شود.
- یک فرایند تصادفی زمان پیوسته با وضعیت گسسته، یک فرایند نقطه‌ای نامیده می‌شود.

اگر $S=\mathbb{R}^k$ آن‌گاه X_t یک بردار k بعدی خواهد بود.

در مثال ۲-۲ و ۲-۴، به ترتیب داریم $S=\{۱,۲,۳,\cdots\}$ و $S=\{\cdots,-۱,۰,۱,\cdots\}$ بنابراین در هر دوی این مثال‌ها $\{X_n;\ n\in T\}$ یک زنجیر تصادفی است. همچنین در مثال ۲-۱، $S=\{۰,۱\}$ و بنابراین یک زنجیرتصادفی داریم. دقت کنید که در مثال ۲-۱، X_n یک متغیرتصادفی برنولی است و X_n ها مستقل و هم توزیع هستند. چنین فرایندهای تصادفی به فرایندهای برنولی معروف هستند.

تعریف ۲-۵. فرایندتصادفی $\{X_n;\ n=۱,۲,\cdots\}$، یک فرایند برنولی نامیده می‌شود هرگاه

۱. متغیرهای تصادفی X_n، $n=۱,۲,\cdots$ مستقل از هم باشند.

۲. برای هر مقدار $n=۱,۲,\cdots$، $\mathrm{P}(X_n=x)=p^x(۱-p)^{۱-x}$ بطوریکه $x=۰,۱$.

پیشامد $X_n=x$ به این معنی است که فرایند در لحظه n در وضعیت x قرار دارد. برای مثال در مثال قدم زدن تصادفی (مثال ۲-۴) داریم

$$\mathrm{P}(X_۳=-۱)=\binom{۳}{۱}p(۱-p)^۲$$
$$\mathrm{P}(X_۳=۳)=p^۳$$
$$\mathrm{P}(X_۳=۰)=\mathrm{P}(X_۳=۲)=۰$$
$$\mathrm{P}(X_۳=۱)=\binom{۳}{۲}p^۲(۱-p)$$

تمرین ۲-۶. فرض کنید شخصی یک تاس را بطور متوالی پرتاب می‌کند و هر بار اگر عدد فرد مشاهده کند یک گام به جلو و اگر عدد زوج غیراول مشاهده کند یک گام به عقب برمی‌گردد و در صورتی که عدد ۲ را مشاهده کند گامی برنمی‌دارد. فرض کنید X_t مکان شخص پس از پرتاب t-ام تاس باشد. در این مثال فرایند و ویژگی های آن را مشخص کرده و تابع احتمال $X_۳$ را بدست آورید.

تعریف ۲-۷. در فرایند تصادفی $\{X_t;\ t\in T\}$، اگر x_t یک مقدار ممکن برای متغیر تصادفی X_t باشد آن‌گاه $\{x_t;\ t\in T\}$ را یک مسیر نمونه‌ای گفته می‌شود. مسیر نمونه‌ای ارتباطی است که به هر $t\in T$ مقدار x_t را مربوط می‌کند. مثلا در مثال ۲-۱، ۱۰۱۱۰۰۱۰۱۰۰۰۰۱۰ قسمتی از یک مسیر نمونه‌ای است.

همانطورکه در فصل اول هم گفتیم، خواص آماری و احتمالی هر بردار تصادفی بطورکامل توسط توزیع توام و یا تابع چگالی توام تعیین و مشخص می‌شود. در حالتی که با یک فرایند تصادفی $\{X_t;\ t\in T\}$ کار می‌کنیم (بخصوص وقتی T ناشماراست) توزیع توام فرایند تصادفی ممکن است قابل تعریف نباشد. در نظریه

فرایندهای تصادفی طبق قرارداد توزیع یک فرایند تصادفی وقتی معلوم و مشخص است که توزیع هر مقطع متناهی از این فرایند معلوم باشد. درواقع، فرایند تصادفی $\{X_t;\ t \in T\}$ در صورتی از نظر احتمالاتی کاملا معلوم است که به ازای هر $n \in \mathbb{N}$ و هر انتخاب دلخواه $t_0, t_1, \cdots, t_n \in T$، تابع توزیع توام بردارتصادفی $(X_{t_0}, X_{t_1}, \cdots, X_{t_n})$ یعنی

$$F_{t_0,t_1,\cdots,t_n}(x_1, \cdots, x_n) = \mathrm{P}(X_{t_0} \leq x_0, X_{t_1} \leq x_1, \cdots, X_{t_n} \leq x_n) \qquad x_0, \cdots, x_n \in \mathbb{R}$$

معلوم باشد.

همچنین دو فرایند تصادفی $\{X_t;\ t \in T\}$ و $\{Y_t;\ t \in T\}$ از نظر احتمالاتی برابر (یا هم توزیع) هستند اگر و تنها اگر به ازای هر n و هر $t_0, t_1, \cdots, t_n \in T$ داشته باشیم

$$(X_{t_0}, X_{t_1}, \cdots, X_{t_n}) =^d (Y_{t_0}, X_{t_1}, \cdots, Y_{t_n}).$$

تعریف ۲-۸. به خانواده توزیع‌های

$$\{F_{t_0,\cdots,t_n} \ :\quad t_0, \cdots, t_n \in T,\ t_0 \leq \cdots \leq t_n,\ n \in \mathbb{N}\}$$

توزیع‌های متناهی البعد فرایندتصادفی $\{X_t;\ t \in T\}$ گفته می‌شود.

پس توزیع یک فرایند تصادفی توسط توزیع‌های متناهی البعد تعیین می‌گردد.
گاهی ممکن است به جای بدست آوردن توابع توزیع توام $F_{t_0,\cdots,t_n}$، توابع چگالی احتمال توام (یعنی $f_{t_0,\cdots,t_n}$) را بدست آوریم که به آن توابع چگالی احتمال متناهی البعد گفته می‌شود.

مثال ۲-۹. در مثال ۲-۱ برای بدست آوردن توابع چگالی احتمال متناهی البعد، به ازای هر $n \in \mathbb{N}$ و هر $x_1, \cdots, x_n \in \mathbb{N}$ داریم

$$\begin{aligned} f_{X_{t_1},\cdots,X_{t_n}}(x_1, \cdots, x_n) &= \mathrm{P}(X_{t_1} = x_1, \cdots, X_{t_n} = x_n) \qquad x_1, \cdots, x_n \in \{0, 1\} \\ &= \prod_{i=1}^{n} \mathrm{P}(X_{t_i} = x_i) \\ &= \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \\ &= p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}. \end{aligned}$$

توجه داشته باشید که در حالتی که $T = \mathbb{N}$، توزیع‌های متناهی البعد فرایند تصادفی وقتی معلوم هستند که خانواده

$$\{F_{1,2,\cdots,n};\ n \in \mathbb{N}\}$$

معلوم باشد و در حالتی که $T = \mathbb{N}_0$، توزیع‌های متناهی البعد فرایند تصادفی وقتی معلوم هستند که خانواده

$$\{F_{0,1,\cdots,n};\ n \in \mathbb{N}_0\}$$

معلوم باشد.

تمرین ۲-۱۰. در مثال ۲-۲ توابع چگالی احتمال متناهی البعد را بدست آورید.

۱.۲ زنجیر مارکف

تعریف ۲-۱۱. یک فرآیند مارکف، فرآیندی است که اثر گذشته فرآیند بر آینده آن برابر با اثر آخرین گذشته است. به زبان ریاضی فرایند تصادفی $\{X_t;\ t \in T\}$ که روی فضای احتمال $(\Omega, \mathcal{F}, \mathrm{P})$ تعریف شده است فرایند مارکف است اگر به ازای هر پیشامد A و $s < t \in T$ داشته باشیم

$$\mathrm{P}(X_t \in A | X_u,\ \forall u \leq s) = \mathrm{P}(X_t \in A | X_s).$$

زنجیر مارکف یک فرایند مارکف زمان گسسته با وضعیت شمارا است. بنابراین بطور خاص اگر S یک مجموعه شمارش پذیر باشد آنگاه فرایند تصادفی $\{X_n;\ n = 1, 2, \cdots\}$ که برای آن به ازای هر $i_0, i_1, \cdots, i_{n-1}, i, j \in S$ داشته باشیم

$$\mathrm{P}(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \cdots, X_{n-1} = i_{n-1}, X_n = i) = \mathrm{P}(X_{n+1} = j | X_n = i)$$

یک زنجیر مارکف روی S نامیده می‌شود.

زنجیر مارکف زمان گسسته را می توان به صورت حرکت ذره ای مجسم کرد که اطلاعات مربوط به مکان ذره در لحظه $n+1$ به شرط دانستن مسیر ذره تا انتقال n-ام برابر با این اطلاعات به شرط دانستن وضعیت آن ذره تنها در انتقال n-ام است.

تعریف ۲-۱۲. $\mathrm{P}(X_{n+1} = j | X_n = i)$ احتمال انتقال از حالت i به j طی یک مرحله است که با $p_{ij}^{n\ n+1}$ تعریف می‌شود. در صورتی که احتمال انتقال یک مرحله‌ای به n وابسته نباشد، زنجیر مارکف را زمان-همگن (یا ایستا) نامیده و برای سادگی بجای $p_{ij}^{n\ n+1}$ از نماد p_{ij} استفاده می‌شود.

از اینجا به بعد فقط به زنجیره‌های مارکف زمان-همگن می‌پردازیم و p_{ij} احتمال انتقال از وضعیت i به j طی یک مرحله را نشان می‌دهد.

گزاره ۲-۱۳. به راحتي میتوان نشان داد (با استفاده از رابطه (۱-۳)) هر زنجیر مارکف زمان همگن با معلوم بودن احتمالات انتقال یك مرحله ای و تابع توزیع متغیر لحظه شروع، از نظر احتمالاتی معلوم است. در واقع به سادگی میتوان نشان داد برای هر $n = 0, 1, 2, \cdots$ و وضعیت‌های دلخواه $i_0, i_1, \cdots, i_n$،

$$\mathrm{P}(X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n) = \mathrm{P}(X_0 = i_0) p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} \qquad (۲-۱)$$

ماتریس $\mathbf{P} = (p_{ij})$، $i, j \in S$، ماتریس احتمالات انتقال یک مرحله‌ای زنجیرمارکف با فضای وضعیت S نامیده می‌شود. درواقع داریم

$$\mathbf{P} = \begin{pmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ \vdots & \ddots & \ddots & \vdots \end{pmatrix}$$

واضح است که مولفه‌های این ماتریس همه نامنفی است و برای هر $i \in S$ داریم $\sum_{j \in S} p_{ij} = 1$. درواقع جمع هر سطر از این ماتریس برابر با یک است زیرا زنجیر در هر وضعیت که باشد در مرحله بعد به یک وضعیت در S خواهد رفت.

برای مثال ماتریس احتمال انتقال یک مرحله‌ای یک زنجیر مارکف با دو وضعیت ۰ و ۱ برابر است با

$$\mathbf{P} = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix} = \begin{pmatrix} p & 1-p \\ q & 1-q \end{pmatrix}$$

که در آن $p, q \in [0, 1]$.

مثال ۲-۱۴. فرایند مارکف $\{X_t;\ t = 0, 1, 2, \ldots\}$ را با فضای وضعیت $\{0, 1, 2\}$ و ماتریس احتمال انتقال زیر درنظر بگیرید

$$\mathbf{P} = \begin{pmatrix} 0/6 & 0/3 & 0/1 \\ 0/3 & 0/3 & 0/4 \\ 0/4 & 0/1 & 0/5 \end{pmatrix}$$

اگر بدانیم $\mathrm{P}(X_0 = 0) = \mathrm{P}(X_0 = 1) = \frac{1}{2}$، مقدار $\mathrm{P}(X_0 = 1, X_1 = 0, X_2 = 2)$ را محاسبه کنید.

حل. با استفاده از (۲-۱) داریم

$$\begin{aligned} \mathrm{P}(X_0 = 1, X_1 = 0, X_2 = 2) &= \mathrm{P}(X_0 = 1)\mathrm{P}(X_1 = 0 | X_0 = 1)\mathrm{P}(X_2 = 2 | X_1 = 0) \\ &= \mathrm{P}(X_0 = 1) p_{10} p_{02} \\ &= \frac{1}{2} \times 0/3 \times 0/1 = 0/015. \end{aligned}$$

با معلوم بودن ماتریس احتمال انتقال و توزیع متغیر زمان شروع در یک زنجیر مارکف، فرآیند از نظر احتمالی معلوم است. تعریف می کنیم

$$\mathrm{P}(X_\circ = i) = \pi_i \qquad \forall\, i \in S$$

در نتیجه طبق رابطه ۲−۱ داریم

$$\begin{aligned}&\mathrm{P}(X_\circ = i, X_1 = i_1, \cdots, X_n = i_n)\\ &\quad = \mathrm{P}(X_n = i_n | X_{n-1} = i_{n-1})\mathrm{P}(X_{n-1} = i_{n-1} | X_{n-2} = i_{n-2}) \cdots \mathrm{P}(X_1 = i_1 | X_\circ = i)\mathrm{P}(X_\circ = i)\\ &\quad = p_{i_{n-1}i_n} \times p_{i_{n-2}i_{n-1}} \times \cdots p_{ii_1}\pi_i.\end{aligned}$$

مثال ۲−۱۵. فرض کنید $\{X_n;\ n = \circ, 1, \cdots\}$ دنباله‌ای iid از متغیرهای تصادفی با توزیع بصورت زیر باشد

$$\mathrm{P}(X_\circ = k) = a_k \qquad k = \circ, 1, 2, \cdots, m.$$

داریم

$$\begin{aligned}\mathrm{P}(X_{n+1} = j | X_\circ = i_\circ, \cdots, X_{n-1} = i_{n-1}, X_n = i) &= \mathrm{P}(X_{n+1} = j) = a_j\\ &= \mathrm{P}(X_{n+1} = j | X_n = i)\end{aligned}$$

و بنابراین این زنجیر تصادفی یک زنجیر مارکف بوده که ماتریس احتمال انتقال آن به شکل $\mathbf{P} = (p_{ij} = a_j)$ است.

مثال ۲−۱۶. فرض کنید $\{X_n;\ n \in \mathbb{N}_\circ\}$ یک زنجیر مارکف با فضای حالت $S = \{\circ, 1, \cdots, m\}$ و ماتریس احتمال انتقال با مولفه‌های زیر باشد

$$p_{ij} = \begin{cases} q_i & \text{if} \quad j = i - 1 \\ p_i & \text{if} \quad j = i + 1 \end{cases} \text{ if } i = 1, 2, \cdots, m - 1 \qquad p_{\circ\circ} = p_{mm} = 1.$$

این زنجیر یک مدل ورشکستگی قمارباز است به این صورت که دو بازیکن با مجموع سرمایه m با هم بازی می‌کنند با این قانون که در هر بازی بازیکن بازنده یک واحد به بازیکن برنده می‌پردازد. اگر X_n سرمایه یکی از بازیکن‌ها درنظر گرفته شود آن‌گاه $\{X_n;\ n \in \mathbb{N}_\circ\}$ یک زنجیر مارکف خواهد بود.

مثال ۲−۱۷. فرض کنید d مهره به شماره‌های $1, 2, \cdots, d$ و دو جعبه با شماره های ۱ و ۲ داریم. در ابتدا i مهره در جعبه ۱ و بقیه در جعبه ۲ قرار دارند. هربار شماره‌ای از $1, 2, \cdots, d$ به تصادف انتخاب شده و مهره

نظیر آن شماره از جعبه‌ای که در آن قرار دارد خارج و به جعبه دیگر انداخته می‌شود. فرض کنید X_n تعداد مهره‌های موجود در جعبه ۱ پس از n بار تکرار مستقل آزمایش باشد. در این صورت $\{X_n;\ n \in \mathbb{N}_0\}$ یک زنجیر مارکف با احتمال انتقالات زیر است

$$p_{ij} = \begin{cases} \dfrac{i}{d} & \text{if} \quad j = i - 1 \\ \frac{d-i}{d} & \text{if} \quad j = i + 1 \end{cases} \qquad i = 1, 2, \cdots, d-1$$

$$p_{01} = p_{m,m-1} = 1.$$

مثال ۲-۱۸. وضعیتی را در نظر بگیرید که برای پاسخگویی به تقاضای موجود، یک کالا انبار می شود. فرض کنید که کل تقاضا در دوره n-ام متغیر تصادفی ξ_n است که توزیع آن به دوره زمانی n بستگی ندارد.

$$\mathrm{P}(\xi_n = k) = a_k \qquad k = 0, 1, \cdots \quad a_k \geq 0 \quad \sum_{k=0}^{\infty} a_k = 1$$

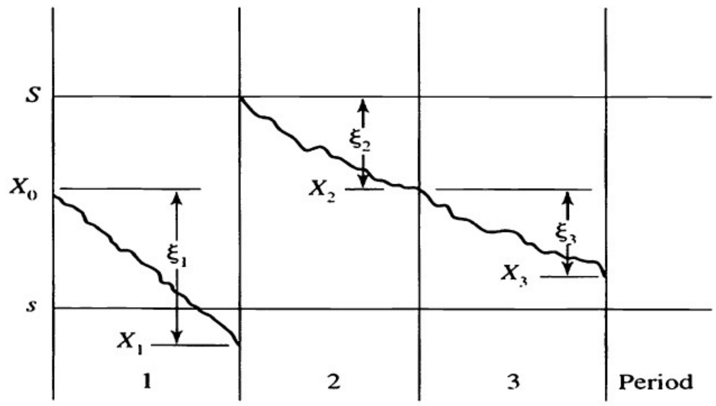
انبار فقط در پایان هر دوره‌ی $n = 0, 1, \cdots$ پر می شود بدین صورت که سطح ذخیره انبار در پایان هر دوره بررسی می‌شود و اگر موجودی از s کمتر باشد آنرا تا S پر می کنند بطوریکه $s < S$. همچنین اگر در پایان دوره موجودی بیشتر یا مساوی s باشد جایگزینی انجام نمی‌گیرد. فرض کنید X_n تعداد کالاهای ذخیره شده در پایان دوره n دقیقا قبل از پر کردن مجدد باشد. بنابراین

$$X_n \in \{S, S-1, \cdots, 1, 0, -1, -2, \cdots\}$$

که مقادیر منفی به این معناست که تقاضایی در n-امین دوره وجود داشته که پاسخ داده نشده و به محض پر کردن انبار پاسخ داده می شود. یک مسیر نمونه ای از این فرآیند در شکل ۲-۱ نشان داده شده است.

احتمالات انتقال یک مرحله‌ای این زنجیر به فرم زیر است

$$p_{ij} = \mathrm{P}(X_{n+1} = j | X_n = i) = \begin{cases} a_{i-j} & s < i \leq S,\ j \leq i \\ a_{S-j} & i \leq s,\ j \leq S \\ 0 & \text{o.w} \end{cases}$$



شکل ۲-۱: قسمتی از مسیر نمونه‌ای یک فرایند مدل انبارداری

و بنابراین داریم

$$\mathbf{P} = \begin{array}{c} \\ S \\ S-۱ \\ S-۲ \\ \vdots \\ s \\ s-۱ \\ \vdots \end{array} \begin{array}{c} \begin{array}{cccccc} S & S-۱ & S-۲ & \ldots & s & \ldots \end{array} \\ \begin{pmatrix} a_۰ & a_۱ & a_۲ & \cdots & \cdots & \cdots \\ ۰ & a_۰ & a_۱ & \cdots & \cdots & \cdots \\ ۰ & a_۰ & a_۱ & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ldots & \ldots & \cdots \\ a_۰ & a_۱ & a_۲ & \cdots & \cdots & \cdots \\ a_۰ & a_۱ & a_۲ & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ldots & \ldots & \cdots \end{pmatrix} \end{array}$$

مثال ۲-۱۹. فرض کنید یک فرد از مبدأ خود روی یک خط راست در هر واحد زمانی یک واحد به جلو یا به عقب می رود. اگر وضعیت فرد در قدم n-ام نسبت به مبدا را با X_n نمایش دهیم آن‌گاه $\{X_n;\ n \in \mathbb{N}\}$ زنجیر قدم‌زدن تصادفی ساده نامیده می‌شود و داریم

$$S = \{۰, \pm ۱, \pm ۲, \cdots\}$$

$$\mathrm{P}(X_n = i + ۱ | X_{n-۱} = i) = \mathrm{P}(X_n = i - ۱ | X_{n-۱} = i) = \frac{۱}{۲} \qquad \forall i \in \mathbb{N}$$

$$\mathrm{P}(X_n = i | X_{n-۱} = i) = ۰ \qquad \forall i \in \mathbb{N}$$

و بنابراین ماتریس انتقال $\mathbf{P} = (p_{i,j})$ بطوریکه $\mathrm{P}(X_n = j | X_{n-۱} = i) = p_{i,j}$ با مولفه‌های زیر تعریف

می‌شود

$$p_{ij} = \begin{cases} \frac{1}{2} & \text{if} \quad j = i+1 \text{ or } j = i-1 \\ 0 & \text{o.w} \end{cases}$$

حالت دیگر وقتی است که فقط مقادیر صحیح نامنفی عضو فضای وضعیت هستند. در این حالت فرض کنید وقتی فرد به وضعیت صفر می‌رسد، برمی‌گردد ولی به هر حال زنجیری است که در هر انتقال یا یک واحد افزایش و یا یک واحد کاهش می‌یابد. میتوان فرض دیگری را نیز اضافه کرد و آن هم اینکه در یک انتقال بتواند بدون تغییر در جای خود بماند و همچنین شانس تغییر وضعیت به مکان شخص در آن لحظه نیز بستگی داشته باشد. در این حالت داریم

$$S = \{0, 1, 2, \cdots\} = \mathbb{N}_0$$

$$p_{ij} = \begin{cases} p_i & \text{if} \quad j = i+1 \\ q_i & \text{if} \quad j = i-1 \\ r_i & \text{if} \quad j = i \\ 0 & \text{o.w} \end{cases}$$

بطوری‌که $p_0 + r_0 = 1$، $p_i, q_i, r_i \geq 0$ و برای $i = 1, 2, \cdots$، $p_i + q_i + r_i = 1$. زنجیر تصادفی تعریف شده در این حالت یک زنجیر قدم زدن تصادفی نامیده می‌شود.

مثال ۲-۲۰. سرویس دهنده ای را در نظر بگیرید که در هر واحد زمانی به یک مشتری سرویس می دهد و کار مشتری در یک واحد زمانی به پایان می‌رسد. در هر واحد زمانی تعدادی تصادفی مشتری وارد صف می‌شوند و سرویس دهنده به ترتیب ورود آن‌ها به آن‌ها سرویس خواهد داد. فرض کنید تعداد مشتریانی که در واحد زمانی n وارد صف می‌شوند، مستقل از سایر زمان‌ها باشد و توزیع احتمال آن نیز به n وابستگی نداشته باشد. اگر ξ_n تعداد مشتریانی که در واحد زمانی n وارد صف می‌شوند تعریف شود آن‌گاه داریم

$$\mathrm{P}(\xi_n = k) = \mathrm{P}(\xi = k) = a_k \qquad k = 0, 1, 2 \cdots \quad \sum_{k=0}^{\infty} a_k = 1.$$

حال اگر X_n تعداد مشتریان داخل صف در پایان زمان n-ام باشد آن‌گاه

$$X_{n+1} = \begin{cases} \xi_n & \text{if} \quad X_n = 0 \\ X_n - 1 + \xi_n & \text{if} \quad X_n > 0 \end{cases}$$

و

$$\mathrm{P}(X_{n+1}=j|X_n=i)=\mathrm{P}(X_n-1+\xi_n=j|X_n=i)=\mathrm{P}(\xi_n=j-i+1)=a_{j-i+1}$$
$$\mathrm{P}(X_{n+1}=j|X_n=0)=\mathrm{P}(\xi_n=j)=a_j$$

بنابراین ماتریس انتقال $\mathbf{P}=(p_{ij})$ دارای مولفه‌های زیر است

$$p_{ij}=\begin{cases} a_{j-i+1} & \text{if} \quad j-i+1\geq 0 \\ 0 & \text{o.w} \end{cases}$$

دقت داشته باشید که اگر فضای وضعیت یک زنجیر تصادفی متناهی باشد، ماتریس احتمال انتقال یک مرحله ای آن مربعی با بعد متناهی است، در غیر این صورت تعداد سطرها و ستون‌های ماتریس انتقال نامتناهی خواهد بود.

تعریف ۲-۲۱. ماتریس $\mathbf{A}=(a_{ij})$ یک ماتریس تصادفی نامیده می‌شود اگر

۱. برای هر $i,j\in S$، $a_{ij}\geq 0$

۲. $\sum_{i\in S}a_{ij}=1$.

این دو خاصیت علاوه بر اینکه در مورد ماتریس انتقال یک مرحله ای برقرار است در جبر ماتریس‌ها دارای اهمیت ویژه است. بنابراین ماتریس احتمالات انتقال یک مرحله‌ای در زنجیرهای مارکف یک ماتریس تصادفی است. هر زنجیر مارکف دارای ماتریس احتمال انتقال تصادفی است و برای هر ماتریس تصادفی می‌توان یک زنجیر مارکف تعریف کرد.

مثال ۲-۲۲. فرض کنید $Y_0,Y_1,\cdots$ دنباله‌ای از متغیرهای تصادفی مستقل و هم‌توزیع با متغیر Y باشد و داشته باشیم $\mathrm{P}(Y=i)=a_i$. در این صورت در مورد فرایند تصادفی $\{Y_n;\ n\in\mathbb{N}_0\}$ داریم

$$\mathrm{P}(Y_{n+1}=j|Y_0=i_0,Y_1=i_1,\cdots,Y_n=i)=\mathrm{P}(Y_{n+1}=j)=a_j=\mathrm{P}(Y_{n+1}=j|Y_n=i)$$

و بنابراین این فرایند یک فرایند مارکف (و البته یک زنجیر مارکف) است. ماتریس احتمال انتقال یک مرحله‌ای این زنجیر بصورت زیر است

$$\mathbf{P}=\begin{array}{c c} & \begin{array}{cccc} 0 & 1 & 2 & \dots \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \end{array} & \begin{pmatrix} a_0 & a_1 & a_2 & \cdots \\ a_0 & a_1 & a_2 & \cdots \\ a_0 & a_1 & a_2 & \cdots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix} \end{array}$$

دقت کنید که در مثال بالا استقلال متغیرهای تصادفی فرایند موجب یکسان بودن مولفه‌های سطرهای مختلف ماتریس انتقال است. می توان ثابت کرد که عکس این مطلب هم درست است یعنی برابر بودن سطرهای مختلف ماتریس انتقال یک مرحله‌ای در یک فرایند تصادفی بیان کننده استقلال بین متغیرهای تصادفی آن فرایند است.

مثال ۲-۲۳. در مثال ۲-۲۲ فرض کنید که برای $i = 0, 1, 2, \cdots$، تعریف کنیم $X_n = Y_0 + Y_1 + \cdots + Y_n$. در این صورت داریم

$$\begin{aligned}
\mathrm{P}(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \cdots, X_n = i) \\
&= \mathrm{P}(Y_0 + Y_1 + \cdots + Y_{n+1} = j | Y_0 = i_0, Y_0 + Y_1 = i_1, \cdots, Y_0 + Y_1 + \cdots + Y_n = i) \\
&= \mathrm{P}(Y_{n+1} = j - i) = a_{j-i} \qquad j - i \geq 0
\end{aligned}$$

و بنابراین

$$p_{ij} = \begin{cases} a_{j-i} & \text{if} \quad j - i \geq 0 \\ 0 & \text{o.w} \end{cases}$$

درواقع داریم

$$\mathbf{P} = \begin{array}{c c} & \begin{array}{cccc} 0 & 1 & 2 & \cdots \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \end{array} & \begin{pmatrix} a_0 & a_1 & a_2 & \cdots \\ 0 & a_0 & a_1 & \cdots \\ 0 & 0 & a_0 & \cdots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix} \end{array}$$

در یک فرایند زمان- همگن، احتمال انتقال در n مرحله یا گام که برابر است با $\mathrm{P}(X_{m+n} = j | X_m = i)$، به m، یعنی شروع n گام بستگی ندارد و تعریف می‌کنیم

$$p_{ij}^{(n)} = \mathrm{P}(X_{m+n} = j | X_m = i)$$

همچنین ماتریس احتمال انتقال n مرحله‌ای به صورت $\mathbf{P}^{(n)} = (p_{ij}^{(n)})$ نشان داده می‌شود.

قضیه ۲-۲۴. احتمالات انتقال n مرحله‌ای در رابطه زیر که به نام رابطه چپمن-کلموگروف معروف است، صدق می کند

$$p_{ij}^{(n)} = \sum_{k \in S} p_{ik} p_{kj}^{(n-1)}$$

بطوریکه

$$p_{ij}^{(\circ)} = \begin{cases} ۱ & \text{if} \quad i=j \\ \circ & \quad \text{o.w} \end{cases}$$

اثبات.

$$\begin{aligned} p_{ij}^{(n)} &= \mathrm{P}(X_n = j | X_\circ = i) \\ &= \sum_{k\in S} \mathrm{P}(X_n = j, X_۱ = k | X_\circ = i) \\ &= \sum_{k\in S} \mathrm{P}(X_n = j | X_۱ = k, X_\circ = i)\mathrm{P}(X_۱ = k | X_\circ = i) \\ &= \sum_{k\in S} \mathrm{P}(X_n = j | X_۱ = k)\mathrm{P}(X_۱ = k | X_\circ = i) \\ &= \sum_{k\in S} p_{kj}^{(n-۱)} p_{ik} \end{aligned}$$

□

براساس قضیه ۲-۲۴، داریم

$$\mathbf{P}^{(n)} = \mathbf{P}\mathbf{P}^{(n-۱)} = \mathbf{P}\mathbf{P}\mathbf{P}^{(n-۲)} = \cdots = \mathbf{P}^n.$$

به جهت سادگی در نوشتار از اینجا به بعد برای نوشتن احتمال انتقال n مرحله‌ای از وضعیت i به j ممکن است بجای نماد $p_{ij}^{(n)}$ از نماد p_{ij}^n نیز استفاده شود.

مثال ۲-۲۵. فرض کنید $\{X_n;\ n\in\mathbb{N}_\circ\}$ یک زنجیر مارکف با فضای حالت $S=\{a,b,c\}$ و ماتریس انتقال زیر باشد

$$\mathbf{P} = \begin{array}{c c} & \begin{array}{ccc} a & b & c \end{array} \\ \begin{array}{c} a \\ b \\ c \end{array} & \begin{pmatrix} \frac{۱}{۲} & \frac{۱}{۴} & \frac{۱}{۴} \\ \frac{۲}{۳} & \circ & \frac{۱}{۳} \\ \frac{۳}{۵} & \frac{۲}{۵} & \circ \end{pmatrix} \end{array}$$

در این صورت با توجه به اینکه

$$\mathbf{P}^۲ = \begin{pmatrix} \frac{۱۷}{۳\circ} & \frac{۹}{۴\circ} & \frac{۵}{۲۴} \\ \frac{۸}{۱۵} & \frac{۳}{۱\circ} & \frac{۱}{۶} \\ \frac{۱۷}{۳\circ} & \frac{۳}{۲\circ} & \frac{۱۷}{۶\circ} \end{pmatrix}$$

بدست می‌آوریم

$$
\begin{aligned}
\mathrm{P}(X_1 = b, X_2 = c, X_4 = c, X_5 = a, X_6 = c | X_0 = c) &= p_{cb}\, p_{bc}\, p_{cc}^2\, p_{ca}\, p_{ac} \\
&= \frac{2}{5} \times \frac{1}{3} \times \frac{17}{60} \times \frac{3}{5} \times \frac{1}{4}
\end{aligned}
$$

مثال ۲-۲۶. فرض کنید $\{X_n;\ n \in \mathbb{N}_0\}$ یک فرایند برنولی و T_n زمان n-امین موفقیت باشد. برای $\{T_n;\ n \in \mathbb{N}_0\}$ داریم $T_0 = 0$ و $S = \{0, 1, \cdots\}$. همچنین چون

$$
\begin{aligned}
\mathrm{P}(T_{n+1} = j | T_0 = i_0, T_1 = i_1, \cdots, T_n = i) &= \mathrm{P}(T_{n+1} = j | T_n = i) \\
&= p(1-p)^{j-i-1} \qquad \text{if } j = i+1, i+2, \cdots
\end{aligned}
$$

پس $\{T_n;\ n \in \mathbb{N}_0\}$ یک زنجیر مارکف است و داریم

$$
p_{ij} = \begin{cases} p(1-p)^{j-i-1} & \text{if} \quad j \geq i+1 \\ 0 & \text{o.w} \end{cases}
$$

$$
p_{ij}^m = \begin{cases} \binom{j-i-1}{m-1} p^m (1-p)^{j-i-m} & \text{if} \quad j \geq i+m \\ 0 & \text{o.w} \end{cases}
$$

مثال ۲-۲۷. زنجیره قدم زدن تصادفی با ماتریس احتمال انتقال زیر را در نظر بگیرید که فضای وضعیت آن به صورت $S = \{0, 1, 2, \cdots, a\}$ درنظر گرفته شده باشد

$$
\mathbf{P} = \begin{pmatrix}
1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
q_1 & r_1 & p_1 & 0 & \cdots & 0 & 0 & 0 \\
0 & q_2 & r_2 & p_2 & 0 & \cdots & 0 & 0 \\
\vdots & 0 & \ddots & \ddots & \ddots & 0 & \cdots & 0 \\
0 & 0 & \ddots & 0 & 0 & q_{a-1} & r_{a-1} & p_{a-1} \\
0 & 0 & \ddots & \cdots & \cdots & 0 & q_a & r_a
\end{pmatrix}
$$

بطوریکه $r_a + q_a = 1$ و برای $i = 1, 2, \cdots, a-1$، داریم $p_i + q_i + r_i = 1$. دقت کنید که در این زنجیر تصادفی، اگر وضعیت ۰ ملاقات شود تا ابد در همان وضعیت خواهیم ماند. در واقع برای هر $n = 1, 2, \cdots$، $p_{00}^n = 1$ و برای هر وضعیت $j \neq 0$، $p_{0j}^n = 0$. به وضعیت صفر در این زنجیر تصادفی اصطلاحا وضعیت جاذب گفته می‌شود.

تعریف ۲-۲۸. در یک زنجیر مارکف وضعیت یا حالت i، $i \in S$، را جاذب گویند هرگاه $p_{ii} = ۱$.

تعریف ۲-۲۹. در یک زنجیر مارکف، وضعیت یا حالت j را دردسترس i گویند و آن را با نماد $i \longrightarrow j$ نمایش می‌دهند، هرگاه

$$\exists\, n \geq ۰; \quad p_{ij}^{n} > ۰$$

به عبارت دیگر زنجیر با شروع از وضعیت i با احتمال مثبت پس از تعداد متناهی انتقال به وضعیت j می‌رسد.

تعریف ۲-۳۰. در یک زنجیر مارکف وضعیت یا حالت j و i را دردسترس هم یا مرتبط با هم می‌نامند و آن را با نماد $j \longleftrightarrow i$ نمایش می‌دهند، هرگاه

$$\exists\, n \geq ۰;\ p_{ij}^{n} > ۰ \quad \text{and} \quad \exists\, m \geq ۰;\ p_{ji}^{m} > ۰.$$

برای مثال به کمک رسم نمودار انتقالات یک مرحله‌ای در مثال ۲-۲۰ و با فرض اینکه به ازای هر $k \geq ۰$، $a_k > ۰$، مرتبط بودن تمام وضعیت‌ها را می‌توان دید. پس می‌توان گفت که در این مثال تمام وضعیت‌ها در دسترس یکدیگر هستند.

توجه کنید که مرتبط بودن یک رابطه هم‌ارزی است چون

- هر وضعیت با خودش در ارتباط است یعنی برای هر $i \in S$، $i \longleftrightarrow i$ (زیرا $p_{ii}^{۰} = ۱$).
- اگر $j \longleftrightarrow i$ آن‌گاه $i \longleftrightarrow j$.
- اگر $j \longleftrightarrow i$ و $k \longleftrightarrow j$، آن‌گاه $k \longleftrightarrow i$. (اثبات با استفاده از تساوی چپمن-کلموگروف انجام می‌گردد).

با توجه به اینکه مرتبط بودن یک رابطه هم‌ارزی است، بنابراین در هر فرایند تصادفی می‌توان وضعیت‌ها را به کلاس‌های هم‌ارزی تقسیم کرد بطوریکه در هر رده یا کلاس هم‌ارزی وضعیت‌ها با هم مرتبط هستند.

مثال ۲-۳۱. فرض کنید یک زنجیر مارکف با فضای وضعیت $S = \{۱, ۲, ۳, ۴, ۵\}$ و ماتریس احتمال انتقال به شکل زیر باشد

$$\mathbf{P} = \begin{pmatrix} \frac{۱}{۲} & \frac{۱}{۲} & ۰ & ۰ & ۰ \\ \frac{۱}{۳} & \frac{۲}{۳} & ۰ & ۰ & ۰ \\ ۰ & ۰ & \frac{۱}{۴} & \frac{۱}{۴} & \frac{۱}{۲} \\ ۰ & ۰ & \frac{۱}{۳} & \frac{۱}{۳} & \frac{۱}{۳} \\ ۰ & ۰ & ۱ & ۰ & ۰ \end{pmatrix}$$

واضح است که در این زنجیر دو کلاس هم‌ارزی به صورت $\{1,2\}$ و $\{3,4,5\}$ وجود دارد.

مثال ۲-۳۲. زنجیر تصادفی مربوط به قدم زدن تصادفی با فضای وضعیت $S=\{0,1,2,\cdots,a\}$ و ماتریس احتمال انتقال زیر را درنظر بگیرید

$$\mathbf{P}=\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ q & 0 & p & 0 & \cdots & 0 & 0 \\ 0 & q & 0 & p & \cdots & 0 & 0 \\ \vdots & 0 & \ddots & \ddots & 0 & \cdots & 0 \\ 0 & 0 & \ddots & 0 & q & 0 & p \\ 0 & 0 & \ddots & \cdots & 0 & 0 & 1 \end{pmatrix}$$

کلاس‌های هم‌ارزی این زنجیر به صورت $\{0\},\{1,2,\cdots,a-1\},\{a\}$ است.

اگر یک زنجیر فقط یک رده هم‌ارزی داشته باشد آن زنجیر را تحویل‌ناپذیر گویند.

مثال ۲-۳۳. زنجیر مارکف با ماتریس احتمال انتقال زیر را درنظر بگیرید

$$\mathbf{P}=\begin{pmatrix} p_0 & q_0 & 0 & 0 & \cdots \\ p_1 & 0 & q_1 & 0 & \cdots \\ p_2 & 0 & 0 & q_2 & \cdots \\ \vdots & \vdots & \ddots & \ddots & \cdots \end{pmatrix}.$$

درواقع در این زنجیر در هر مرحله یا یک واحد افزایش داریم و یا به وضعیت صفر کاهش می‌یابد. در این مثال X_n می‌تواند سن افراد باشد به این صورت که فرد در سال n-ام یا یک سال بزرگتر می‌شود و یا می‌میرد.

در این مثال می‌توان ثابت کرد که اگر برای هر $i=0,1,2,\cdots$، $p_i,q_i>0$ آن‌گاه برای هر دو وضعیت i و k، وجود دارد یک $n\geq 0$ بطوریکه $p_{ik}^n>0$.

گزاره ۲-۳۴. اگر تعداد وضعیت‌های یک زنجیر مارکف m باشد و وضعیت j در دسترس i باشد آن‌گاه وضعیت j در کمتر یا مساوی $m-1$ مرحله دست یافتنی است.

اثبات. با توجه به اینکه j در دسترس i است پس یک $n\geq 0$ وجود دارد بطوریکه $p_{ij}^n>0$. پس

$$\exists k_1,k_2,\ldots,k_{n-1}\in S;\quad p_{ij}^{(n)}>p_{ik_1}p_{k_1k_2}p_{k_2k_3}\ldots p_{k_{n-2}k_{n-1}}p_{k_{n-1}j}>0$$

دقت کنید که می‌توان k_r ها را طوری درنظر گرفت که برای $r \neq r'$، $k_r \neq k_{r'}$. حال اگر $k_r = k_{r'}$، آن‌گاه طول مسیر بالا به اندازه $|r - r'|$ کوتاهتر می‌شود چون داریم

$$p_{ik_1}p_{k_1k_2}\cdots p_{k_{r-1}k_r}p_{k_{r'}k_{r'+1}}\cdots p_{k_{n-2}k_{n-1}}p_{k_{n-1}j} > 0.$$

درواقع در حالتی که برای هر $r \neq r'$، $k_r \neq k_{r'}$ چون k_iها از یک مجموعه $m-2$ عضوی انتخاب می‌شوند و بیش از $m-2$ عضو متمایز نمی‌توان اختیار کرد داریم $n-1 \leq m-2$ و بنابراین $n \leq m-1$. $\square$

تعریف ۲-۳۵. دوره تناوب وضعیت i برابر است با $\gcd\{n \geq 1;\ p_{ii}^n > 0\}$ که آن را با $d(i)$ نمایش می‌دهند. اگر برای هر $n \geq 1$ داشته باشیم $p_{ii}^n = 0$ آن‌گاه $d(i) = 0$.

مثال ۲-۳۶. در مثال ۲-۳۲ با توجه به ماتریس احتمال انتقال داریم

$$\begin{aligned} &d(0) = 1 \qquad d(a) = 1 \\ &d(i) = \gcd\{2, 4, \cdots\} = 2 \quad i = 1, 2, \cdots, a-1. \end{aligned}$$

مثال ۲-۳۷. فرض کنید ماتریس احتمال انتقال یک زنجیر مارکف با فضای وضعیت $S = \{0, 1, \cdots, n\}$ به صورت زیر تعریف شده باشد

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \cdots \\ 0 & 0 & \cdots & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

در این مثال برای هر $i = 0, 1, \cdots, n$، داریم $d(i) = n$

مثال ۲-۳۸. در مثال ۲-۳۳ داریم

$$\begin{aligned} d(0) &= \gcd\{1, 2, 3, \cdots\} = 1 \\ d(1) &= \gcd\{2, 3, 4, \cdots\} = 1 \\ &\vdots \\ d(i) &= 1 \qquad \forall i \geq 0 \end{aligned}$$

در قضیه زیر اثبات می‌شود که دوره تناوب تمام وضعیت‌های یک کلاس هم‌ارزی برابرند و بنابراین اگر یک زنجیر تحویل‌ناپذیر باشد آن‌گاه دوره تناوب همه وضعیت‌های آن برابر است.

قضیه ۲-۳۹. اگر $i \longleftrightarrow j$ آن‌گاه $d(i) = d(j)$.

اثبات. برای دو وضعیت i, j بطوریکه $i \longleftrightarrow j$ طبق تعریف اعداد r, m وجود دارند بطوریکه $p_{ij}^m, p_{ji}^r > 0$. به ازای هر $n \geq 0$ که $p_{ii}^n > 0$ داریم

$$p_{jj}^{m+r+n} \geq p_{ji}^r \, p_{ii}^n \, p_{ij}^m > 0$$

و بنابراین $d(j)$ مقدار $m + r + n$ را می‌شمارد. چون $d(j)$ مقدار $m + r$ را می‌شمارد، n را نیز می‌شمارد. از آنجایی که $d(i)$ بزرگترین شمارنده مجموعه nهایی است که $p_{ii}^n > 0$، $d(j) \leq d(i)$. به استدلال مشابه نیز بدست می‌آوریم $d(i) \leq d(j)$ و از اینرو $d(i) = d(j)$. □

تعریف ۲-۴۰. زنجیر مارکفی که دوره تناوب همه وضعیت‌های آن برابر با یک باشد (یعنی برای هر $i \in S$، $d(i) = 1$)، نامتناوب نامیده می‌شود. (مثل مثال ۲-۳۸)

قضیه ۲-۴۱. اگر $d(i)$ دوره تناوب وضعیت i در یک زنجیر مارکف باشد آن‌گاه

$$\exists N_i;\ \forall n > N_i; \quad p_{ii}^{nd(i)} > 0.$$

تعریف ۲-۴۲. نماد f_{ii}^n احتمال این است که اولین بازگشت به i در انتقال n-ام رخ دهد به شرط اینکه شروع فرایند از وضعیت i بوده باشد. درواقع داریم

$$\begin{aligned} f_{ii}^n &= \mathrm{P}(X_n = i, X_{n-1} \neq i, \cdots, X_1 \neq i | X_0 = i) \\ f_{11}^1 &= p_{ii} \\ f_{ii}^0 &= 0 \end{aligned}$$

گزاره ۲-۴۳. تساوی زیر همواره برقرار است

$$p_{ii}^n = \sum_{k=0}^{n} f_{ii}^k \, p_{ii}^{n-k} \qquad n \geq 1 \quad i \in S. \tag{۲-۲}$$

اثبات. فرض کنید E_k پیشامد این باشد که $X_n = i$ و اولین بازگشت به حالت i در لحظه k رخ دهد مشروط

بر اینکه $X_0 = i$. در این صورت داریم

$$\begin{aligned}
p_{ii}^n &= \mathrm{P}(X_n = i|X_0 = i) = \mathrm{P}(\bigcup_{k=0}^{n} E_k) = \sum_{k=0}^{n} \mathrm{P}(E_k) \\
&= \sum_{k=0}^{n} \mathrm{P}(X_n = i, X_k = i, X_{k-1} \neq i, X_{k-2} \neq i, \cdots, X_1 \neq i|X_0 = i) \\
&= \sum_{k=0}^{n} \mathrm{P}(X_n = i|X_k = i, X_{k-1} \neq i, X_{k-2} \neq i, \cdots, X_1 \neq i, X_0 = i) \\
&\qquad \times \mathrm{P}(X_k = i, X_{k-1} \neq i, X_{k-2} \neq i, \cdots, X_1 \neq i|X_0 = i) \\
&= \sum_{k=0}^{n} \mathrm{P}(X_n = i|X_k = i) f_{ii}^k \\
&= \sum_{k=0}^{n} p_{ii}^{n-k}\ f_{ii}^k
\end{aligned}$$

□

تعریف ۲-۴۴. تابع P_{ij} به عنوان تابع مولد دنباله $\{p_{ij}^n\}$ بصورت زیر تعریف می‌شود و بر روی فاصله $(-1, 1)$ متناهی و پیوسته است

$$\mathrm{P}_{ij}(s) = \sum_{n=0}^{\infty} s^n\ p_{ij}^n \qquad |s| \leq 1.$$

تعریف ۲-۴۵. تابع F_{ij} به عنوان تابع مولد دنباله $\{f_{ij}^n\}$ بصورت زیر تعریف می‌شود و بر روی فاصله $[-1, 1]$ متناهی و پیوسته است

$$\mathrm{F}_{ij}(s) = \sum_{n=0}^{\infty} s^n\ f_{ij}^n \qquad |s| \leq 1. \tag{۲-۳}$$

گزاره ۲-۴۶. به ازای هر وضعیت i، رابطه زیر همواره برقرار است

$$\mathrm{P}_{ii}(s) = \frac{1}{1 - \mathrm{F}_{ii}(s)} \tag{۲-۴}$$

اثبات. با استفاده از تعریف تابع P_{ij} و رابطه (۲-۲) و براساس این نکته که رابطه (۲-۲) برای $n = 0$ برقرار

نیست، بدست می‌آوریم

$$\begin{aligned}
\mathrm{P}_{ii}(s) &= \sum_{n=0}^{\infty} s^n\, p_{ii}^n \\
&= 1 + \sum_{n=1}^{\infty} s^n\, p_{ii}^n \qquad\qquad (\text{Since } s^0 p_{ii}^0 = 1) \\
&= 1 + \sum_{n=1}^{\infty} s^n \sum_{k=1}^{n} f_{ii}^k\, p_{ii}^{n-k} \\
&= 1 + \sum_{n=1}^{\infty} \sum_{k=1}^{n} s^n\, f_{ii}^k\, p_{ii}^{n-k} \\
&= 1 + \sum_{k=1}^{\infty} f_{ii}^k\, s^k \sum_{n=k}^{\infty} p_{ii}^{n-k}\, s^{n-k} \\
&= 1 + \sum_{k=0}^{\infty} f_{ii}^k\, s^k \sum_{n=0}^{\infty} p_{ii}^{n}\, s^{n} \qquad (\text{Since } f_{ii}^0 = 0) \\
&= 1 + \mathrm{F}_{ii}(s)\mathrm{P}_{ii}(s)
\end{aligned}$$

□

لازم به ذکر است که به ازای $|s| < 1$، تابع $\mathrm{F}_{ii}(s)$ اکیدا کوچکتر از ۱ است.

مشابه اثبات رابطه (۲-۴)، می‌توان ثابت کرد که برای هر $i \neq j$ رابطه زیر همواره برقرار است

$$\mathrm{P}_{ij}(s) = \mathrm{F}_{ij}(s)\mathrm{P}_{jj}(s) \tag{۲-۵}$$

تعریف ۲-۴۷. وضعیت i را بازگشتی نامند اگر فرایند با شروع از i با احتمال یک بعد از تعداد متناهی بار انتقال به وضعیت i بازگردد. به عبارت دیگر اگر

$$\sum_{n=1}^{\infty} f_{ii}^n = 1$$

آن‌گاه وضعیت i بازگشتی نامیده می‌شود. وضعیتی که بازگشتی نباشد گذرا نامیده می‌شود.

مثال ۲-۴۸. بررسی کنید که در زنجیرمارکف با فضای وضعیت $S = \{0, 1, 2, 3\}$ و ماتریس انتقال زیر وضعیت صفر گذرا است یا بازگشتی؟

$$\mathrm{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & \frac{2}{3} & 0 & 0 \\ 0 & 0 & \frac{3}{4} & \frac{1}{4} \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

حل. برای وضعیت صفر داریم

$$f_{00}^{1} = \frac{1}{2} \qquad f_{00}^{2} = \frac{1}{2} \times \frac{1}{3} = \frac{1}{6}$$
$$\vdots$$
$$f_{00}^{n} = \frac{1}{2} \times (\frac{2}{3})^{n-2} \times \frac{1}{3} \quad n = 2, 3, \cdots$$
$$\Longrightarrow \quad \sum_{n=1}^{\infty} f_{00}^{n} = \frac{1}{2} + \sum_{n=2}^{\infty} \frac{1}{2} \times (\frac{2}{3})^{n-2} \times \frac{1}{3}$$
$$= \frac{1}{2} + \frac{\frac{1}{6}}{1 - \frac{2}{3}} = 1$$

در نتیجه وضعیت صفر بازگشتی است.

وضعیت بازگشتی را با توجه به مقادیر p_{ii}^{n} نیز می‌توان تشخیص داد.

قضیه ۲-۴۹. وضعیت i بازگشتی است اگر و تنها اگر

$$\sum_{n=1}^{\infty} p_{ii}^{n} = \infty$$

□

گزاره ۲-۵۰. اگر $j \longleftrightarrow i$ و i بازگشتی باشد، آن‌گاه j نیز بازگشتی است.

بنابراین مانند دوره تناوب، خاصیت بازگشتی یک وضعیت نیز یک خاصیت رده‌ای است. یعنی در هر کلاس هم‌ارزی یا همه وضعیت‌ها بازگشتی و یا همه گذرا هستند.

مثال ۲-۵۱. در یک زنجیر قدم زدن تصادفی با فضای وضعیت اعداد صحیح، اگر در هر انتقال ذره با احتمال p یک قدم به راست (جلو) و با احتمال $q = 1 - p$ یک قدم به چپ (عقب) برود، وضعیت‌ها را از نظر بازگشتی یا گذرا بودن بررسی کنید.

حل. با توجه به اینکه این زنجیر تنها یک رده هم‌ارزی دارد پس تحویل ناپذیر است. بنابراین فقط کافی است یکی از وضعیت‌ها مثلا وضعیت صفر را از نظر گذرا را بازگشتی بودن بررسی کنیم. داریم

$$p_{00}^{2n+1} = 0 \qquad p_{00}^{2n} = \binom{2n}{n} p^{n} q^{n} = \frac{(2n)!}{n!n!}(pq)^{n}.$$

با استفاده از تقریب استرلینگ می دانیم $n! \simeq n^{n+0/5}e^{-n}\sqrt{2\pi}$ و بنابراین

$$p_{00}^{2n} \simeq \frac{(4pq)^n}{\sqrt{n\pi}}$$

می دانیم $pq \leq 0/25$ و تساوی وقتی برقرار است که $p = q = 0/5$. در چنین شرایطی داریم

$$\sum_{n=0}^{\infty} p_{00}^n \simeq \sum_{n=0}^{\infty} \frac{1}{\sqrt{n\pi}} = \infty.$$

حال اگر $pq < 0/25$ آن‌گاه

$$\sum_{n=0}^{\infty} p_{00}^n \simeq \sum_{n=0}^{\infty} \frac{(4pq)^n}{\sqrt{n\pi}} < \infty.$$

بنابراین در این زنجیر قدم زدن تصادفی تمام وضعیت‌ها بازگشتی هستند فقط اگر $p = q = 0/5$ و در غیراینصورت همه گذرا هستند.

تمرین ۲-۵۲. ماتریس انتقال یک مرحله‌ای زیر را برای یک زنجیر مارکف با فضای حالت‌های $S = \{0, 1, 2, 3, 4, 5\}$ درنظر بگیرید.

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 \\ 0 & \frac{1}{5} & \frac{2}{5} & \frac{1}{5} & 0 & \frac{1}{5} \\ 0 & 0 & 0 & \frac{1}{6} & \frac{1}{3} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{4} & 0 & \frac{3}{4} \end{pmatrix}$$

کلاس‌های این زنجیر را مشخص کرده و گذرا یا بازگشتی بودن هر یک از وضعیت‌ها را نیز تعیین نمایید.

گزاره ۲-۵۳. اگر وضعیت i بازگشتی بوده و با وضعیت j در ارتباط نباشد آن‌گاه به ازای هر $n \geq 0$، $p_{ij}^n = 0$. درواقع زمانی که فرآیند وارد یک کلاس بازگشتی از وضعیت‌ها شد هرگز آن را ترک نمی‌کند. **به همین دلیل یک کلاس بازگشتی یک کلاس بسته نامیده می شود.**

بنابراین در ماتریس احتمالات انتقال، همه درایه‌های موجود در یک سطر متناظر با یک وضعیت بازگشتی برابر با صفر هستند مگر درایه‌های مربوط به اعضا همان کلاس هم‌ارزی.

برای بدست آوردن احتمالات ملاقات براي اولين بار بررسی تمام مسیرهای ممکن شانس اشتباه بالایی دارد و اغلب نیز پیچیده است. در مثال بعد روش ساده‌ای برای این منظور مطرح می‌شود.

مثال ۲-۵۴. یک زنجیر مارکف با فضای وضعیت $S=\{۱,۲,۳\}$ و ماتریس احتمال انتقال زیر را درنظر بگیرید

$$\mathbf{P}=\begin{pmatrix} ۱ & ۰ & ۰ \\ \frac{۱}{۲} & \frac{۱}{۶} & \frac{۱}{۳} \\ \frac{۱}{۳} & \frac{۳}{۵} & \frac{۱}{۱۵} \end{pmatrix}$$

کلاس‌های این زنجیر را مشخص کرده و گذرا یا بازگشتی بودن هر یک از وضعیت‌ها را نیز تعیین نمایید.

حل. طبق تعریف f_{ij}^k احتمال این است که با شروع از وضعیت i بعد از k مرحله برای اولین بار j را ملاقات کنیم. طبق تعریف ۲-۴۷، برای بررسی بازگشتی بودن هر وضعیت مثل i کافی است درستی رابطه $\sum_{n=۱}^{\infty} f_{ii}^n=۱$ را چک کنیم. اما چطور f_{ii}^n را محاسبه کنیم.

فرض کنید $i=۳$. برای بدست آوردن مقادیر $f_{۳۳}^n$ تعریف می کنیم

$$\mathbf{f}_۳^{(k)}=(f_{۱۳}^k,f_{۲۳}^k,f_{۳۳}^k)'$$

بردار فوق را می‌توان علاوه بر محاسبه درایه به درایه مستقیم به صورت زیر نیز محاسبه کرد. به این صورت که ماتریس $\mathbf{Q}$ را برابر با ماتریس $\mathbf{P}$ قرار می‌دهیم با این تفاوت که ستون سوم آن برابر با صفر است یعنی

$$\mathbf{Q}=\begin{pmatrix} ۱ & ۰ & ۰ \\ \frac{۱}{۲} & \frac{۱}{۶} & ۰ \\ \frac{۱}{۳} & \frac{۳}{۵} & ۰ \end{pmatrix}$$

در این صورت می توان نشان داد که

$$\mathbf{f}_۳^{(k)}=\mathbf{Q}\mathbf{f}_۳^{(k-۱)},\qquad \mathbf{f}_۳^{(۱)}=\begin{pmatrix} ۰ \\ \frac{۱}{۳} \\ \frac{۱}{۱۵} \end{pmatrix}.$$

بنابراین

$$\mathbf{f}_۳^{(۲)}=\begin{pmatrix} ۰ \\ \frac{۱}{۱۸} \\ \frac{۱}{۵} \end{pmatrix},\quad \mathbf{f}_۳^{(۳)}=\mathbf{Q}\mathbf{f}_۳^{(۲)}=\begin{pmatrix} ۰ \\ \frac{۱}{۱۰۸} \\ \frac{۱}{۳۰} \end{pmatrix},\quad \cdots \Longrightarrow \begin{cases} f_{۲۳}^k=\frac{۱}{۳}(\frac{۱}{۶})^{k-۱} & k=۱,۲,\cdots \\ f_{۳۳}^k=\begin{cases} \frac{۱}{۱۵} & \text{if}\quad k=۱ \\ \frac{۳}{۵}(\frac{۱}{۶})^{k-۲}(\frac{۱}{۳}) & \text{if}\quad k>۱ \end{cases} \end{cases}$$

$$\Longrightarrow \sum_{n=۱}^{\infty} f_{۳۳}^n=\frac{۱}{۱۵}+\sum_{n=۲}^{\infty}\frac{۳}{۵}(\frac{۱}{۶})^{n-۲}(\frac{۱}{۳})=\frac{۱}{۱۵}+\frac{\frac{۱}{۵}}{۱-\frac{۱}{۶}}=۰/۳\neq ۱$$

حال متغیر جدیدی را در یک فرایند تصادفی تعریف می کنیم. فرض کنید T_j مدت زمانی باشد که از لحظه صفر طی شده تا برای اولین بار وضعیت j را ملاقات کنیم. درواقع داریم

$$\mathrm{P}_i(T_۳ = k) = f_{i۳}^k.$$

بنابراین داریم

$$\mathrm{P}_i(T_۳ < \infty) = \sum_{k=۱}^{\infty} f_{i۳}^k.$$

برای مثال در مثال ۲-۵۴ بدست می آوریم

$$\mathrm{P}_۲(T_۳ < \infty) = \sum_{k=۱}^{\infty} f_{۲۳}^k = \sum_{k=۱}^{\infty} \frac{۱}{۳}(\frac{۱}{۶})^{k-۱} = \frac{\frac{۱}{۳}}{۱-\frac{۱}{۶}} = \frac{۲}{۵}$$
$$\mathrm{P}_۳(T_۳ < \infty) = \sum_{k=۱}^{\infty} f_{۳۳}^k = \frac{۱}{۱۵} + \frac{۳}{۵} \times \frac{۱}{۳} \sum_{k=۱}^{\infty} (\frac{۱}{۶})^{k-۲} = \frac{۱}{۱۵} + \frac{۳}{۱۵} \times \frac{۱}{۱-\frac{۱}{۶}} = \frac{۲۳}{۶۵} < ۱$$

قضیه ۲-۵۵. فرض کنید $\{X_۰, X_۱, \cdots\}$ یک زنجیر مارکف و Y یک تابع کراندار از متغیرهای تصادفی $X_n, X_{n+۱}, \cdots$ باشد. در این صورت داریم

$$\mathrm{E}(Y|X_۰, X_۱, \cdots, X_n) = \mathrm{E}(Y|X_n)$$

(درواقع نوعی خاصیت مارکفی در امیدریاضی برقرار است)

تعریف ۲-۵۶. متغیر تصادفی T با فضای مقادیر $\{۰, ۱, ۲, \cdots\}$ برای فرایند $\{X_n;\ n \in \mathbb{N}_۰\}$ زمان توقف نامیده می‌شود اگر به ازای هر $n \geq ۰$ رخداد یا عدم رخداد پیشامد $T = n$ با مشاهده $X_۰, X_۱, \cdots, X_n$ مشخص باشد.

می‌توان نشان داد که مجموع چند زمان توقف یک زمان توقف است و مینیمم چند زمان توقف یک زمان توقف است.

تعریف ۲-۵۷. مدت زمانی که طول می‌کشد تا برای اولین بار در زنجیر تصادفی یک وضعیت خاص ملاقات شود یک **زمان توقف** است.

قضیه ۲-۵۸. در یک زنجیر مارکف اگر T زمان توقف باشد برای هر تابع کراندار f روی $\mathbb{R}^\infty$ رابطه زیر برقرار است

$$\mathrm{E}(f(X_T, X_{T+۱}, \cdots)|X_n,\ n \leq T) = \mathrm{E}(f(X_T, X_{T+۱}, \cdots)|X_T)$$
$$\mathrm{P}(X_{T+m} = j|X_n = i_n,\ n \leq T) = \mathrm{P}(X_{T+m} = j|X_T = i_T) = \mathrm{P}_{i_T,j}^m$$

یکی دیگر از شاخص های مفید در مساله زنجیر مارکف $N(j)$ است که طبق تعریف تعداد دفعاتی را می‌شمارد که زنجیر وضعیت j را ملاقات می‌کند. براساس رابطه‌ای شاخص $N(j)$ به صورت زیر تعریف می‌شود

$$N(j) = \sum_{n=0}^{\infty} I_j(X_n)$$

بطوریکه

$$I_j(x) = \begin{cases} 1 & \text{if} \quad x = j \\ 0 & \text{if} \quad x \neq j \end{cases}$$

تابع نشانگر است. وقتی $N(j)$ متناهی است به این معنا است که زنجیر بلاخره زمانی j را برای آخرین بار ملاقات می‌کند و بعد از آن هرگز به j بازنمی‌گردد. در حالی که وقتی $N(j)$ بینهایت باشد، زنجیر بارها و بارها وضعیت j را ملاقات خواهد کرد.

گزاره ۲-۵۹. اگر $N(j)$ تعداد دفعاتی باشد که زنجیر وضعیت j را ملاقات می‌کند، آن‌گاه

$$\mathrm{P}(N(j) = m | X_0 = j) = F_{jj}^{m-1}(1 - F_{jj}) \qquad m = 1, 2, \cdots$$

و برای $i \neq j$،

$$\mathrm{P}(N(j) = m | X_0 = i) = \begin{cases} 1 - F_{ij} & \text{if} \quad m = 0 \\ F_{ij} F_{jj}^{m-1}(1 - Fjj) & \text{if} \quad m = 1, 2, \cdots \end{cases}$$

نتیجه ۲-۶۰. در هر زنجیر مارکف داریم

$$\mathrm{P}(N(j) < \infty | X_0 = j) = \begin{cases} 1 & \text{if} \quad F_{jj} < 1 \\ 0 & \text{if} \quad F_{jj} = 1 \end{cases}$$

و بنابراین اگر با احتمال ۱، $N(j) < \infty$ آن‌گاه $N(j)$ دارای توزیع هندسی با پارامتر $1 - F_jj$ است. در نتیجه $\mathrm{E}_j(N(j)) = \dfrac{1}{1 - F_{jj}}$.

گزاره ۲-۶۱. وضعیت j بازگشتی است اگر

$$\mathrm{P}(T_j < \infty | X_0 = j) = 1$$

یا بطور معادل

$$\mathrm{P}(N(j) = \infty | X_0 = j) = 1.$$

همچنین وضعیت j گذراست اگر

$$\mathrm{P}(T_j = \infty | X_0 = j) > 0$$

یا بطور معادل

$$\mathrm{P}(N(j) = \infty | X_0 = j) = 0$$

که در این حالت به شرط اینکه $X_0 = j$ داریم $N(j)$ و $G(1 - F_{jj})$ هم توزیع هستند.

گزاره ۲-۶۲. اگر i و j دو وضعیت از زنجیر مارکف $\{X_n;\ n \geq 0\}$ باشند آن‌گاه

$$\mathrm{E}_i(N(j)) = \sum_{n=0}^{\infty} p_{ij}^n. \tag{۲-۶}$$

بنابراین می‌توان گفت وضعیت i بازگشتی است اگر امید تعداد ملاقات‌های i به شرط شروع از وضعیت i بینهایت باشد. همچنین اگر امید تعداد دیدارها از وضعیت i متناهی باشد بدین معناست که با احتمال ۱، فرایند سرانجام i را ترک می‌کند و هرگز به آن بازنمی‌گردد.

تعریف ۲-۶۳. ماتریس $\mathbf{R}$ ماتریس پتانسیل زنجیر تصادفی نامیده می‌شود بطوریکه مولفه (i, j) آن بصورت زیر تعریف شده باشد

$$R(i, j) = \mathrm{E}_i(N(j)). \tag{۲-۷}$$

می توان ثابت کرد که $R(i, j) = F_{ij} R(j, i)$. دقت کنید که $F_{ij} = \sum_{n=0}^{\infty} f_{ij}^n$ احتمال این است که زنجیر با شروع از حالت i بلاخره به j برسد. و $R(i, i) = \frac{1}{1 - F_{ii}}$ واورن احتمال این است که زنجیر با شروع از وضعیت i هرگز به i بازنگردد.
لازم به ذکر است که برای محاسبه F_{ij} در اکثر مواقع راحتتر است که ابتدا $R(i, j)$ را بدست آورده و سپس F_{ij} را بدست آوریم.
با توجه به آنچه تاکنون گفته شد، وضعیت j بازگشتی است اگر $R(i, j) = \infty$ و گذراست اگر $R(i, j) < \infty$. دقت کنید که در حالتی که j گذراست داریم

$$R(i, j) = F_{ij} R(j, i) \leq R(j, j) < \infty$$

و برای یک وضعیت بازگشتی j داریم

$$R(i, j) = \begin{cases} 0 & \text{if} \quad F_{ij} = 0 \\ \infty & \text{if} \quad F_{ij} > 0 \end{cases}$$

اگر j گذرا باشد آن‌گاه

$$\sum_{n=0}^{\infty} p_{jj}^{n} = R(j,j) < \infty$$

و در نتیجه

$$\lim_{n\longrightarrow\infty} p_{jj}^{n} = 0.$$

ولی اگر j بازگشتی باشد ممکن است $\lim_{n\longrightarrow\infty} p_{jj}^{n}$ برابر با صفر باشد یا نباشد. در مورد $\lim_{n\longrightarrow\infty} p_{jj}^{n}$ در حالتی که j بازگشتی است تا اینجا نمی‌توان نتیجه‌ای گرفت.

لم ۲-۶۴. اگر j بازگشتی باشد و $k \longrightarrow j$، آن‌گاه $j \longrightarrow k$ و $F_{kj} = 1$. یعنی بین دو وضعیت بازگشتی در یک کلاس بینهایت رفت و برگشت رخ می‌دهد.

با توجه به آنچه تا کنون گفته شد، اگر $k \longrightarrow j$، دو حالت ممکن است داشته باشیم: (۱) j و k در یک کلاس هم ارزی قرار دارند؛ (۲) j گذراست. یعنی از وضعیت‌های بازگشتی فقط به حالت‌های بازگشتی همان کلاس می‌توان رفت.

تمرین ۲-۶۵. زنجیر مارکفی با ماتریس انتقال زیر را درنظر بگیرید.

$$\mathbf{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{5} & \frac{4}{5} & 0 \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \end{pmatrix}$$

کلاس‌های هم ارزی این زنجیر را مشخص کرده و حالت‌های گذرا و بازگشتی را در آن تعیین نمایید.

با توجه به روابط (۲-۶) و (۲-۷)، براساس نماد ماتریسی می توان نوشت

$$\mathbf{R} = \mathbf{I} + \mathbf{P} + \mathbf{P}^{2} + \cdots$$

و بنابراین داریم

$$\mathbf{RP} = \mathbf{P} + \mathbf{P}^{2} + \mathbf{P}^{3} + \cdots = \mathbf{R} - \mathbf{I}$$
$$\Longrightarrow \quad \mathbf{R}(\mathbf{I} - \mathbf{P}) = \mathbf{I}$$

همچنین بطور مشابه بدست می‌آوریم

$$\mathbf{PR} = \mathbf{P} + \mathbf{P}^{2} + \mathbf{P}^{3} + \cdots = \mathbf{R} - \mathbf{I}$$
$$\Longrightarrow \quad (\mathbf{I} - \mathbf{P})\mathbf{R} = \mathbf{I}.$$

در این صورت اگر در یک زنجیر مارکف تعداد متناهی وضعیت داشته باشیم و $\mathbf{I}-\mathbf{P}$ معکوس‌پذیر باشد آن‌گاه زنجیر گذراست.

فرض کنید در یک زنجیر مارکف تعداد متناهی وضعیت گذرا وجود دارد و بقیه وضعیت‌ها بازگشتی باشند. اگر احتمالات انتقال بین وضعیت‌های گذرا را درایه‌های بالای سمت چپ ماتریس قرار دهیم آن‌گاه ماتریس احتمالات انتقال این زنجیر را می‌توان به فرم ماتریس بلوک بندی زیر نشان داد

$$\mathbf{P}=\begin{pmatrix}\mathbf{K} & \mathbf{0}\\ \mathbf{L} & \mathbf{Q}\end{pmatrix}$$

که $\mathbf{Q}$ احتمالات انتقال بین وضعیت های گذراست. طبق ویژگی ماتریس‌های بلوک بندی داریم

$$\mathbf{P}^m=\begin{pmatrix}\mathbf{K}^m & \mathbf{0}\\ \mathbf{L}_m & \mathbf{Q}^m\end{pmatrix}$$

و در نتیجه ماتریس $\mathbf{R}$ به شکل زیر خواهد بود

$$\mathbf{R}=\begin{pmatrix}\sum_{m=0}^{\infty}\mathbf{K}^m & \mathbf{0}\\ \sum_{m=0}^{\infty}\mathbf{L}_m & \sum_{m=0}^{\infty}\mathbf{Q}^m\end{pmatrix}$$

اگر تعریف کنیم $\mathbf{S}=\sum_{m=0}^{\infty}\mathbf{Q}^m$ آن‌گاه

$$\mathbf{S}(\mathbf{I}-\mathbf{Q})=(\mathbf{I}-\mathbf{Q})\mathbf{S}=\mathbf{I}$$

و بنابراین

$$\mathbf{S}=(\mathbf{I}-\mathbf{Q})^{-1}$$

به عبارت دیگر درایه‌های نظیر وضعیت‌های گذرا در ماتریس $\mathbf{R}$ براساس رابطه بالا قابل محاسبه هستند.

تمرین ۲-۶۶. زنجیر مارکفی با ماتریس انتقال زیر را درنظر بگیرید.

$$\mathbf{P}=\left(\begin{array}{ccc|cc|ccc}
0.4 & 0.3 & 0.3 & 0 & 0 & 0 & 0 & 0\\
0 & 0.6 & 0.4 & 0 & 0 & 0 & 0 & 0\\
0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0\\
\hline
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0\\
0 & 0 & 0 & 0.8 & 0.2 & 0 & 0 & 0\\
\hline
0 & 0 & 0 & 0 & 0 & 0.4 & 0.6 & 0\\
0.4 & 0.4 & 0 & 0 & 0 & 0 & 0 & 0.2\\
0.1 & 0 & 0.3 & 0 & 0 & 0.6 & 0 & 0
\end{array}\right)$$

کلاس‌های هم ارزی این زنجیر را مشخص کرده و حالت‌های گذرا و بازگشتی را در آن تعیین نمایید.

۲.۲ مساله جذب شدن در یک کلاس

فرض کنید زنجیر مارکفی با ماتریس احتمال انتقال زیر داشته باشیم

$$\mathbf{P}=\begin{array}{c} \\ 0\\ 1\\ 2\end{array}\begin{array}{c} \begin{array}{ccc} 0 & 1 & 2\end{array}\\ \begin{pmatrix} 1 & 0 & 0\\ \alpha & \beta & \gamma\\ 0 & 0 & 1\end{pmatrix}\end{array}$$

بطوریکه $\alpha,\beta,\gamma>0$ و $\alpha+\beta+\gamma=1$. واضح است که اگر فرایند در وضعیت ۱ باشد تا زمانی که به وضعیت ۰ یا ۲ نرود در ۱ می‌ماند ولی به محض وارد شدن به وضعیت ۰ تا ابد در آن می‌ماند (جذب می‌شود). همچنین اگر از ۱ به ۲ هم برود جذب ۲ شده است. تعریف می‌کنیم

$$T=\min\{n\geq 0;\ X_n=0 \text{ or } X_n=2\}$$

که درواقع T مدت زمان لازم برای جذب وضعیت‌های ۰ یا ۲ شدن است. حال می‌خواهیم احتمال جذب شدن را محاسبه کنیم.

احتمال اینکه بلاخره جذب ۰ شویم را به صورت $u = \mathrm{P}(X_T = 0, T < \infty | X_0 = 1)$ تعریف می کنیم و $v = \mathrm{E}(T|X_0 = 1)$ متوسط زمان جذب شدن را نشان می‌دهد. داریم

$$\begin{aligned}
u &= \mathrm{P}(X_T = 0, T < \infty | X_0 = 1)\\
&= \sum_{k=0}^{2} \mathrm{P}(X_T = 0, T < \infty | X_0 = 1, X_1 = k)\mathrm{P}(X_1 = k | X_0 = 1)\\
&= 1 \times \alpha + u\beta + 0 \times \gamma\\
&= \alpha + \beta u\\
&\Longrightarrow u = \frac{\alpha}{1-\beta}\\
&\Longrightarrow v = \mathrm{E}(T|X_0 = 1) = \mathrm{E}(\mathrm{E}(T|X_0 = 1)|X_1) = \alpha \times 1 + \beta(1 + v) + \gamma \times 1\\
&\Longrightarrow v = \frac{1}{1-\beta}
\end{aligned}$$

محاسبات بالا ساده بدست آمد چون ماتریس انتقال درنظر گرفته شده ماتریس ساده‌ای بود. حال مساله را در حالتی که بیش از یک وضعیت گذرا وجود دارد بررسی می‌کنیم. فرض کنید

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ p_{10} & p_{11} & p_{12} & p_{13} \\ p_{20} & p_{21} & p_{22} & p_{23} \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

ماتریس انتقال احتمال برای یک زنجیر مارکف باشد. با توجه به ماتریس انتقال بالا، جذب در وضعیت ۰ یا ۳ اتفاق می‌افتد. می‌توان نشان داد که وضعیت‌های ۱ و ۲ گذرا هستند(**چرا؟**). داریم

$$T = \min\{n \geq 0;\ X_n = 0 \text{ or } X_n = 3\}$$

$$u_i = \mathrm{P}(X_T = 0 | X_0 = i) \quad \Longrightarrow \begin{cases} u_0 = 1 \\ u_1 = p_{10} + p_{11}u_1 + p_{12}u_2 \\ u_3 = 0 \\ u_2 = p_{20} + p_{21}u_1 + p_{22}u_2 \end{cases}$$

$$v_i = \mathrm{E}(T | X_0 = i) \quad \Longrightarrow \begin{cases} v_0 = v_3 = 0 \\ v_1 = 1 + p_{11}v_1 + p_{12}v_2 \\ v_2 = 1 + p_{21}v_1 + p_{22}v_2 \end{cases}$$

تمرین ۲-۶۷. فرض کنید ماتریس انتقال یک زنجیر مارکف به صورت زیر باشد

$$\mathbf{P} = \begin{array}{c} \\ ۰ \\ ۱ \\ ۲ \\ ۳ \end{array} \begin{array}{c} \begin{array}{cccc} ۰ & ۱ & ۲ & ۳ \end{array} \\ \begin{pmatrix} ۱ & ۰ & ۰ & ۰ \\ ۰/۴ & ۰/۳ & ۰/۲ & ۰/۱ \\ ۰/۳ & ۰/۳ & ۰/۳ & ۰/۱ \\ ۰ & ۰ & ۰ & ۱ \end{pmatrix} \end{array}$$

مقادیر u_i و v_i را به ازای $i = ۱, ۲$ بدست آورید.

در حالت کلی فرض کنید $\{X_n;\ n = ۰, ۱, \cdots\}$ یک زنجیر مارکف با وضعیت‌های متناهی $\{۰, ۱, \cdots, N\}$ باشد بطوریکه $۰, ۱, \cdots, r - ۱$ وضعیت‌های گذرا و $r, r + ۱, \cdots, N$ وضعیت‌های جاذب باشند. ماتریس احتمال انتقال این زنجیر فرمی به شکل زیر خواهد داشت

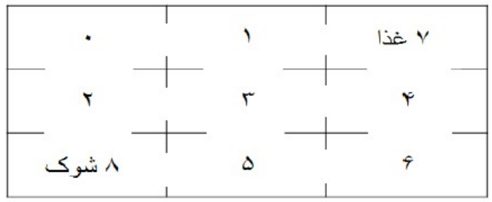
$$\mathbf{P} = \begin{array}{c} \\ ۰ \\ \vdots \\ r-۱ \\ r \\ \vdots \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} ۰ & \cdots & r-۱ & r & \cdots & N \end{array} \\ \begin{pmatrix} \mathbf{Q} & \mathbf{R} \\ & \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \end{array}$$

فرایند با شروع از یک حالت گذرا تا مدت کوتاهی در آن وضعیت‌ها قرار می‌گیرد ولی بلاخره جذب یکی از وضعیت‌های جاذب می‌شود. احتمال اینکه با شروع از حالت گذرای i بلاخره جذب حالت k شود را با نماد u_{ik} نشان می‌دهیم. در این صورت

$$\begin{aligned} u_{ik} &= p_{ik} + \sum_{j=r\ \&\ j\neq k}^{N} p_{ij} \times ۰ + \sum_{j=۰}^{r-۱} p_{ij} u_{jk} \\ &= p_{ik} + \sum_{j=۰}^{r-۱} p_{ij} u_{jk} \qquad i = ۰, ۱, \cdots, r - ۱ \end{aligned}$$

با حل سیستم معادلات بالا احتمال جذب k شدن به شرط شروع از i بدست می‌آید. اگر v_i را متوسط زمان تا جذب به شرط شروع از وضعیت i درنظر بگیریم آن‌گاه

$$v_i = \sum_{k=۰}^{r-۱} p_{ik}(v_k + ۱) + \sum_{k=r}^{N} p_{ik} \qquad i = ۰, ۱, \cdots, r - ۱$$



شکل ۲-۲: شکل قرارگیری اتاق‌ها در تونل

که با حل دستگاه معادلات بالا امیدریاضی مدت زمان جذب با شروع از هر حالت گذرا بدست می‌آید.

تمرین ۲-۶۸. یک موش در داخل یک تونل با اتاق‌های مختلف رها شده است که شکل تونل به صورت نشان داده شده در شکل ۲-۲ است فرض کنید موش در هر اتاق از این تونل k اتاق مجاور خود در جهات مختلف را با احتمال $\frac{1}{k}$ انتخاب می‌کند و می‌رود تا نهایتا یا غذا را پیدا کند و یا با شوک مواجه شود. احتمال اینکه موش غذا را قبل از مواجه شدن با شوک پیدا کند چقدر است؟

۳.۲ زنجیر شاخه‌ای

موجوداتی را درنظر بگیرید که می‌توانند موجودات دیگر از نوع خود را تولید کنند. تعداد موجودات اولیه نسل اولیه (نسل صفر) نامیده و زاده‌های آن‌ها نسل اول گفته می‌شوند. نسل اول هم زاده‌هایی خواهند داشت که نسل دوم نام دارند و همین طور نسل‌های سوم و ... نامگذاری می‌شوند. تعداد زاده‌های i امین فرد از نسل n-ام را با نماد $N(n,i)$ نشان می‌دهیم. حال فرض کنید $N(n,1), N(n,2), \cdots \sim \xi$، $i = 1, 2, \cdots$، و

$$\mathrm{P}(\xi = k) = p_k \qquad k = 0, 1, \cdots$$

اگر X_n جمعیت در نسل n-ام باشد داریم

$$X_{n+1} = \begin{cases} 0 & \text{if} \quad X_n = 0 \\ N(n,1) + N(n,2) + \cdots + N(n,X_n) & \text{if} \quad X_n > 0 \end{cases}$$

و بنابراین $\{X_n;\ n = 0, 1, \cdots\}$ یک زنجیر مارکف با فضای حالت $S = \{0, 1, 2, \cdots\}$ است.

با توجه به تعریف، وضعیت صفر یک حالت جاذب است. در چنین فرایندی، $T = \min\{n;\ X_n = 0\}$ زمان انقراض نامیده می‌شود. اگر $T < \infty$، آنگاه جمعیت بعد از تعداد متناهی از نسل‌ها منقرض می‌شود. درواقع احتمال اینکه زنجیر بعد از مدتی بلاخره جذب صفر شود برابر است با احتمال اینکه بلاخره منقرض

شود. این احتمال جذب شدن به راحتی با مطالبی که در بخش قبل گفته شد قابل محاسبه است بخصوص حالت‌های خاصی در مورد نسل‌ها وجود دارد که محاسبات را حتی راحتتر نیز می‌کند. مثلا اگر $p_0 = 0$، آن‌گاه با شروع از حالت $i \geq 1$ داریم

$$i = X_0 \leq X_1 \leq \cdots$$

در نتیجه $\mathrm{P}(T = \infty) = 1$ و بنابراین احتمال انقراض صفر است.

حال اگر $p_0 > 0$ و $p_0 + p_1 = 1$ آن‌گاه

$$i = X_0 \geq X_1 \geq \cdots$$

و بنابراین چون

$$\mathrm{P}_i(T \leq n) = (1 - \mathrm{P}_1^n)^i$$

بدست می‌آوریم

$$\mathrm{P}_i(T < \infty) = 1.$$

اما اگر $p_0 > 0$، $p_0 + p_1 < 1$ و $m = \sum_k kp_k < \infty$، آن‌گاه $\mathrm{P}_i(T < \infty) = \eta^i$ بطوریکه η احتمال انقراض با شروع از یک موجود است. می‌توان نشان داد که اگر $m \leq 1$ آن‌گاه $\eta = 1$ و اگر $m > 1$ آن‌گاه $0 < \eta < 1$.

فصل ۳

فرایندهای تجدید

در این فصل به بررسی فرایندهای تصادفی زمان پیوسته می‌پردازیم. قاعدتا در ابتدا لازم است هر یک از مفاهیم کلی که برای زنجیرهای تصادفی تعریف شد برای فرایندهای تصادفی زمان پیوسته نیز تعریف شوند.

در فرایند زمان پیوسته $\{X_t;\ t \geq 0\}$ منظور از نمو فرایند میزان تغییر X_t در دو زمان پشت سر هم است یعنی $X_{t_1} - X_{t_0}$. یک فرایند را با نموهای مستقل نامند هرگاه برای زمان‌های بدون اشتراک نموها مستقل باشند درواقع

$$X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \cdots, X_{t_n} - X_{t_{n-1}}$$

مستقل باشند.

یک فرایند را ایستا گویند اگر به ازای هر $(X_{t_0}, X_{t_1}, \cdots, X_{t_n})$ و هر $h > 0$، داشته باشیم

$$(X_{t_0}, X_{t_1}, \cdots, X_{t_n}) \ =^d \ (X_{t_0+h}, X_{t_1+h}, \cdots, X_{t_n+h})$$

و فرایند دارای نموهای ایستا است اگر برای هر $h > 0$

$$(X_{t_1}-X_{t_0}, X_{t_2}-X_{t_1}, \cdots, X_{t_n}-X_{t_{n-1}}) \ =^d \ (X_{t_1+h}-X_{t_0+h}, X_{t_2+h}-X_{t_1+h}, \cdots, X_{t_n+h}-X_{t_{n-1}+h})$$

نکته ۳-۱. اگر فرایند زمان پیوسته دارای نمو ایستا باشد و $\mathrm{E}(X_t) < \infty$ آن‌گاه

$$\mathrm{E}(X_t) = \mathrm{E}(X_0) + (\mathrm{E}(X_1) - \mathrm{E}(X_0))t.$$

۱.۳ فرایند پواسون

یک فرایند پواسون، فرایند مارکف زمان پیوسته $\{N_t;\ t \in [0, \infty)\}$ با احتمالات انتقال ایستا بصورت زیر است

$$p_{ij}(t) = \mathrm{P}(N_{t+u} = j | N_u = i)$$

می توان اینطور تعبیر کرد که N_t تعداد اتفاقاتی را نشان می‌دهد که در فاصله $[0, t]$ رخ داده است. تعداد اتفاقات رخ داده در فاصله $(t, t+s]$ یعنی $N_{t+s} - N_t$ تنها وابسته به s است نه به t. بعلاوه این تعداد به N_t یعنی تعداد اتفاقات تا لحظه t بستگی ندارد. اگر در هر لحظه حداکثر جهشی به اندازه واحد رخ دهد فرایند شمارشی خواهیم داشت.

فرایند پواسون را به زبان ریاضی به صورت زیر تعریف می‌کنیم.

تعریف ۳-۲. فرایند $\{N_t;\ t \geq 0\}$ فرایند پواسون گفته می‌شود هرگاه

I با احتمال یک $N_0 = 0$؛

II به ازای هر $t, s \geq 0$، $N_{t+s} - N_t$ مستقل از $\{N_u;\ u \leq t\}$ است؛

III به ازای هر s و t، $N_{t+s} - N_t$ فقط به s بستگی دارد؛

IV به احتمال یک هر جهش نگاشت $t \longrightarrow N_t$ به اندازه واحد است.

با توجه به خاصیت دوم مطرح شده در تعریف بالا، برای $t_1 \leq t_2 \leq \cdots \leq t_n \leq t$ و $s > 0$، $N_{t+s} - N_t$ مستقل از $N_{t_1}, N_{t_2}, \cdots, N_{t_n}$ است. بنابراین $N_{t+s} - N_t$ مستقل از

$$N_{t_1}, N_{t_2} - N_{t_1}, \cdots, N_{t_{n-1}} - N_{t_{n-2}}, N_{t_n} - N_{t_{n-1}}$$

است. با این استدلال نتیجه می‌شود این فرایند دارای نموهای مستقل است یعنی $N_{t_1}, N_{t_2} - N_{t_1}, \cdots$ از هم مستقل هستند. به عبارت دیگر تعداد اتفاقات در فواصل زمانی مجزا از هم مستقل هستند.

می توان ثابت کرد که سه خاصیت آخر در تعریف بالا معادل است با اینکه

$$\forall\, t \geq s \geq 0 \qquad N_t - N_s \sim P(\lambda(t-s))$$

و $\lambda > 0$ متوسط تعداد اتفاقات در یک واحد زمان است. در نتیجه

$$\mathrm{E}(N_t) = \lambda t \qquad \forall\, t > 0 \tag{۳-۱}$$

و $\{N_t;\ t \geq 0\}$ را اصطلاحا یک فرایند پواسون با نرخ λ گویند.

نتیجه ۳-۳. در یک فرایند پواسون $\{N_t;\ t \geq 0\}$ با نرخ λ،

$$\mathrm{P}(N_{t+s} - N_t = k | N_u\ u \leq t) = \mathrm{P}(N_{t+s} - N_t = k) = \frac{e^{-s\lambda}(s\lambda)^k}{k!} \quad k = 0, 1, \cdots$$

مثال ۳-۴. در یک فرایند پواسون $\{N_t;\ t \geq 0\}$ با نرخ $\lambda = 8$، مقدار

$$\mathrm{P}(N_{2/5} = 17, N_{3/7} = 22, N_{4/3} = 36)$$

را بدست آورید.

حل.

$$\begin{aligned} P(N_{2/5} = 17, N_{3/7} = 22, N_{4/3} = 36) &= P(N_{2/5} = 17, N_{3/7} - N_{2/5} = 5, N_{4/3} - N_{3/7} = 14) \\ &= \mathrm{P}(N_{2/5} = 17)\mathrm{P}(N_{1/2} = 5)\mathrm{P}(N_{0/6} = 14) \\ &= \frac{e^{-8\times 2/5}(2/5 \times 8)^{17}}{17!} \times \frac{e^{-8\times 1/2}(1/2 \times 8)^{5}}{5!} \\ &\quad \times \frac{e^{-8\times 0/6}(0/6 \times 8)^{14}}{14!} \end{aligned}$$

یکی از مهم ترین ویژگی فرایندهای پواسون در مورد توزیع زمان بین رخدادهای آن است.

فرض کنید S_n به صورت زمان n-امین رخداد در فرایند پواسون با نرخ λ تعریف شود. واضح است که زمان بین رخدادها یک متغیر تصادفی مثبت است (یعنی همواره $S_{n+1} - S_n > 0$). حال برای بررسی توزیع این متغیر تصادفی داریم

$$\begin{aligned} \mathrm{P}(S_{n+1} - S_n > t | S_1, S_2, \cdots, S_n) &= \mathrm{P}(S_{n+1} - S_n > t) \\ &= \mathrm{P}(N_t = 0) = e^{-\lambda t} \end{aligned}$$

در واقع در یک فرایند پواسون $\{N_t;\ t \geq 0\}$ با نرخ λ زمان بین رخدادها از توزیع نمایی با پارامتر λ پیروی می کند.

قضیه ۳-۵. فرض کنید $\{N_t;\ t \geq 0\}$ یک فرایند تصادفی باشد که تعداد رخدادها تا لحظه t را نشان می‌دهد و $S_1, S_2, \cdots$ زمان رخدادهای متوالی باشد. در این صورت این فرایند یک فرایند پواسون با نرخ λ است اگر و تنها اگر $S_1, S_2 - S_1, S_3 - S_2, \cdots$ متغیرهای تصادفی مستقل با توزیع نمایی با پارامتر λ باشند.

گزاره ۳-۶. اگر $\{X_t;\ t \geq 0\}$ و $\{Y_t;\ t \geq 0\}$ دو فرایند پواسون مستقل به ترتیب با نرخ‌های λ و μ باشد آن‌گاه

$$N_t = X_t + Y_t \qquad t \geq 0$$

نیز یک فرایند پواسون با نرخ $\lambda + \mu$ است.

۲.۳ فرایند تجدید

در فرایند پواسون دیدیم که زمان‌های بین دو رخداد مستقل و هم‌توزیع با توزیع نمایی است. فرایند تجدید یک فرایند شمارشی کلی‌تر از فرایند پواسون است چرا که در آن محدودیت توزیع نمایی وجود ندارد.

فرض کنید $X_۰, X_۱, \cdots$ دنباله‌ای از متغیرهای مستقل و هم‌توزیع نامنفی با توزیع F باشد و

$$F(۰) = \mathrm{P}(X_n \leq ۰) = \mathrm{P}(X_n = ۰) < ۱.$$

همچنین فرض کنید X_n زمان بین رخداد $(n-۱)$-ام و n-ام بوده و

$$S_۰ = ۰ \qquad S_n = \sum_{i=۱}^{n} X_i \quad n \geq ۱.$$

در واقع S_n زمان n-امین رخداد است و N_t تعداد رخدادها تا زمان t است یعنی

$$N_t = \max\{n;\ S_n \leq t\}.$$

فرایند شمارشی $\{N_t;\ t \geq ۰\}$ را یک فرایند تجدید و هر رخداد را یک تجدید می‌نامند. چون زمان‌های بین دو تجدید مستقل و هم‌توزیع هستند، در هر تجدید فرایند از لحاظ احتمالاتی از نو شروع می‌شود. توزیع هر متغیر فرایند تجدید به صورت زیر است

$$\begin{aligned} \mathrm{P}(N_t = n) &= \mathrm{P}(N_t \geq n) - \mathrm{P}(N_t \geq n+۱) \\ &= \mathrm{P}(S_n \leq t) - \mathrm{P}(S_{n+۱} \leq t) \\ &= F^{(n)}(t) - F^{(n+۱)}(t) \end{aligned}$$

که در آن منظور از $F^{(n)}(t)$ تابع توزیع S_n در نقطه t است که بر اساس فرم تابع توزیع F بدست خواهد آمد.

برای مثال یک جز از یک دستگاه را درنظر بگیرید که همواره مورد استفاده قرار می‌گیرد و قابل تعویض است. اگر X طول عمر این جز تعریف شود و دارای تابع توزیع F باشد به طوری که هر بار که این جز خراب می‌شود آنرا با یک جز نو عوض کنند آنگاه هر تعویض یک تجدید است و تعداد تجدیدها فرایند تجدید خواهد بود.

وقوع سیل در یک شهر، تعداد تصادفات جاده ای در یک منطقه جغرافیایی، تعداد مراجعه افراد به یک سایت مشخص، تعداد حملات سایبری به یک برنامه بانکی همه میتوانند بعنوان مثال‌هایی از فرایند تجدید و حتی فرایند پواسون مطرح شوند.

تعریف ۳-۷. امیدریاضی تعداد تجدیدها تا لحظه t، یعنی $\mathrm{E}(N_t)$، را تابع تجدید می‌نامند. تابع تجدید که در واقع تابعی از t است با نماد $m(t)$ تعریف می‌شود و داریم

$$m(t) = \mathrm{E}(N_t)$$

قضیه ۳-۸. تابع تجدید در رابطه زیر صدق می کند

$$m(t) = \sum_{n=1}^{\infty} F^{(n)}(t)$$

اثبات. تابع نشانگر $I_n(t)$ را روی بازه $[0, t]$ تعریف می کنیم به این صورت که این تابع اگر n-امین تجدید در $[0, t]$ رخ دهد برابر با یک است و در غیر این صورت صفر خواهد بود. با این تعریف داریم $N_t = \sum_{n=1}^{\infty} I_n(t)$. در نتیجه

$$\begin{aligned}
\mathrm{E}(N_t) &= \mathrm{E}(\sum_{n=1}^{\infty} I_n(t)) \\
&= \sum_{n=1}^{\infty} \mathrm{E}(I_n(t)) \\
&= \sum_{n=1}^{\infty} \mathrm{P}(I_n(t) = 1) \\
&= \sum_{n=1}^{\infty} \mathrm{P}(S_n \leq t) = \sum_{n=1}^{\infty} F^{(n)}(t)
\end{aligned}$$

□

تمرین ۳-۹. اگر در یک فرایند تجدید زمان بین دو تجدید از توزیع یکنواخت $(0, 1)$ پیروی کند، تابع تجدید را بدست آورید.

ثابت می‌شود که تابع تجدید برای $0 \leq t \leq \infty$ به صورت منحصر بفرد توزیع بین رخدادها (یعنی تابع توزیع F) را مشخص می‌کند. برای مثال $m(t) = t\lambda$ متناظر با توزیع نمایی با میانگین $\frac{1}{\lambda}$ است (رابطه (۳-۱) را ببینید).

برای بررسی خواص تابع تجدید به مفهوم زمان توقف بازمی‌گردیم. در اینجا زمان توقف را با N نمایش می‌دهیم.

همانطور که در فصل قبل نیز مطرح شد، متغیر تصادفی N را یک زمان توقف برای دنباله متغیرهای تصادفی مستقل $X_۱, X_۲, \cdots$ گوییم اگر به ازای $n = ۱, ۲, \cdots$، پیشامد $N = n$ مستقل از $X_{n+۱}, X_{n+۲}, \cdots$ باشد. مثلا اگر $X_i \sim Be(p)$، $i = ۱, ۲, \cdots$، آن‌گاه

$$N = \min\{n;\ X_۱ + X_۲ + \cdots + X_n = ۱۰\}$$

زمان توقف آزمایش در جمع پیروزی برابر با ۱۰ است.

قضیه ۳-۱۰. اگر $X_۱, X_۲, \cdots$ متغیرهای تصادفی مستقل و هم‌توزیع با امیدریاضی متناهی بوده و N یک زمان توقف برای X_i ها تعریف شود بطوریکه $\mathrm{E}(N) < \infty$، آن‌گاه

$$\mathrm{E}(\sum_{n=۱}^{N} X_n) = \mathrm{E}(N)\mathrm{E}(X)$$

مثال ۳-۱۱. ثابت کنید که در هر فرایند تجدید، $N_t + ۱$ یک زمان توقف برای زمان‌های بین تجدید است.
حل. برای اثبات کافی است نشان دهیم پیشامد $N_t + ۱ = n$ به $X_{n+۱}, X_{n+۲}, \cdots$ بستگی ندارد. داریم

$$\begin{aligned} N_t + ۱ = n \quad &\Longrightarrow N_t = n - ۱ \\ &\Longrightarrow X_۱ + X_۲ + \cdots + X_{n-۱} \leq t \quad \text{and} \quad X_۱ + X_۲ + \cdots + X_n > t \end{aligned}$$

پس پیشامد $N_t + ۱ = n$ فقط به $X_۱, X_۲, \cdots, X_n$ بستگی دارد و مستقل از $X_{n+۱}, X_{n+۲}, \cdots$ است. بنابراین $N_t + ۱$ یک زمان توقف است و طبق قضیه ۳-۱۰ داریم

$$\mathrm{E}(X_۱ + X_۲ + \cdots + X_{N_t+۱}) = \mathrm{E}(X)\mathrm{E}(N_t + ۱)$$

نتیجه ۳-۱۲. اگر $\mathrm{E}(X) = \mu < \infty$، آن‌گاه

$$\mathrm{E}(S_{N_t+۱}) = \mu(m(t) + ۱)$$

قضیه ۳-۱۳. نسبت متوسط تعداد تجدید در واحد زمان به $\frac{۱}{\mu}$ یعنی عکس امیدریاضی زمان بین دو تجدید، میل می‌کند و داریم

$$\lim_{t \longrightarrow \infty} \frac{m(t)}{t} = \frac{۱}{\mu}.$$

توجه داشته باشید که فرایند پواسون یک فرایند تصادفی با نموهای مستقل و ایستا است و داریم

$$\mathrm{E}(X_0) = 0 \qquad \mathrm{E}(X_t) = \lambda t.$$

در حالیکه فرایند تجدید لزوما دارای نموهای مستقل نیست و لزوما نمو ایستا ندارد. می‌توان نشان داد که **تنها فرایند تجدیدی که هم نموهای مستقل دارد و هم نموهای ایستا، فرایند پواسون است.**

مثال ۳-۱۴. در صورتی که $\{X_t;\ t \geq 0\}$ یک فرایند پواسون با نرخ λ باشد، مقدار $\mathrm{P}(X_1 = 2 | X_3 = 5)$ را بدست آورید.

$$\begin{aligned}
\mathrm{P}(X_1 = 2 | X_3 = 5) &= \frac{\mathrm{P}(X_1 = 2, X_3 = 5)}{\mathrm{P}(X_3 = 5)} \\
&= \frac{\mathrm{P}(X_1 = 2, X_3 - X_1 = 3)}{\mathrm{P}(X_3 = 5)} \\
&= \frac{\mathrm{P}(X_1 = 2)\mathrm{P}(X_2 = 3)}{\mathrm{P}(X_3 = 5)} \\
&= \frac{\frac{e^{-\lambda}\lambda^2}{2!} \times \frac{e^{-2\lambda}(2\lambda)^3}{3!}}{\frac{e^{-3\lambda}(3\lambda)^5}{5!}} = 0/33
\end{aligned}$$

مثال ۳-۱۵. فرض کنید $\{N_t;\ t \geq 0\}$ یک فرایند پواسون با نرخ λ باشد. مقدار $\mathrm{E}(N_t N_{t+s})$ و $\mathrm{P}(N_t = 5 | N_{t+s} = 20)$ را بدست آورید.

حل.

$$\begin{aligned}
\mathrm{E}(N_t N_{t+s}) &= \mathrm{E}(N_t(N_t + (N_{t+s} - N_t))) \\
&= \mathrm{E}(N_t)^2 + \mathrm{E}(N_t)\mathrm{E}(N_{t+s} - N_t) \\
&= (\mathrm{Var}(N_t) + \mathrm{E}^2(N_t)) + \mathrm{E}(N_t)\mathrm{E}(N_s) \\
&= (\lambda t + (\lambda t)^2) + (\lambda t)(\lambda s) \\
&= \lambda t(1 + \lambda t + \lambda s)
\end{aligned}$$

$$\begin{aligned}
\mathrm{P}(N_t = 5 | N_{t+s} = 20) &= \frac{\mathrm{P}(N_t = 5, N_{t+s} = 20)}{\mathrm{P}(N_{t+s} = 20)} = \frac{\mathrm{P}(N_t = 5, N_{t+s} - N_t = 15)}{\mathrm{P}(N_{t+s} = 20)} \\
&= \frac{\mathrm{P}(N_t = 5)\mathrm{P}(N_s = 15)}{\mathrm{P}(N_{t+s} = 20)} \\
&= \frac{\frac{e^{\lambda t}(\lambda t)^5}{5!} \times \frac{e^{\lambda s}(\lambda s)^{15}}{15!}}{\frac{e^{\lambda(t+s)}(\lambda(t+s))^{20}}{20!}} = \frac{20!\ t^5\ s^{15}}{5!\ 15!\ (t+s)^{20}}
\end{aligned}$$

بطور کلی ارتباطی بین فرایند پواسون و توزیع دوجمله‌ای وجود دارد که بصورت قضیه زیر بدون اثبات مطرح می‌شود.

قضیه ۳-۱۶. اگر $\{N_t;\ t \geq 0\}$ یک فرایند پواسون با نرخ λ باشد آن‌گاه برای $0 < s < t$ و $n \geq 0$ داریم

$$(N_s|N_t = n) \sim bin(n, \frac{s}{t})$$

مثال ۳-۱۷. فرض کنید $\{N_1(t);\ t \geq 0\}$ و $\{N_2(t);\ t \geq 0\}$ دو فرایند پواسون مستقل به ترتیب با نرخ‌های λ_1 و λ_2 باشد. احتمال اینکه اولین رخداد در فرایند ترکیبی $\{N_1(t) + N_2(t);\ t \geq 0\}$ از نوع فرایند $\{N_1(t);\ t \geq 0\}$ باشد را بدست آورید.

حل. با توجه به گزاره ۳-۶، $\{N_1(t) + N_2(t);\ t \geq 0\}$ یک فرایند پواسون با نرخ $\lambda_1 + \lambda_2$ است و بنابراین داریم

$$\begin{aligned}
\mathrm{P}(N_1(t) = 1|N_1(t) + N_2(t) = 1) &= \frac{\mathrm{P}(N_1(t) = 1, N_1(t) + N_2(t) = 1)}{\mathrm{P}(N_1(t) + N_2(t) = 1)} \\
&= \frac{\mathrm{P}(N_1(t) = 1, N_2(t) = 0)}{\mathrm{P}(N_1(t) + N_2(t) = 1)} = \frac{\mathrm{P}(N_1(t) = 1)\mathrm{P}(N_2(t) = 0)}{\mathrm{P}(N_1(t) + N_2(t) = 1)} \\
&= \frac{\frac{e^{-\lambda_1 t}(\lambda_1 t)^1}{1!} \times \frac{e^{-\lambda_2 t}(\lambda_2 t)^0}{0!}}{\frac{e^{-(\lambda_1+\lambda_2)t}((\lambda_1+\lambda_2)t))^1}{1!}} \\
&= \frac{\lambda_1}{\lambda_1 + \lambda_2}
\end{aligned}$$

مثال ۳-۱۸. فرض کنید $\{Y_n;\ n = 1, 2, \cdots\}$ یک زنجیر مارکف با فضای حالت گسسته بوده و j یک حالت ثابت باشد. اگر تعداد مراحل متوالی تا دیدار مجدد j را با $S_1, S_2, \cdots$ نشان دهیم، یک فرایند تجدید با استفاده از زنجیر معرفی شده تعریف کنید.

حل. اگر حالت اولیه j باشد آن‌گاه $S_1, S_2 - S_1, S_3 - S_2, \cdots$ مستقل و هم‌توزیع هستند و بنابراین با تعریف $X_1 = S_1$ و $X_i = S_i - S_{i-1}$، $i = 1, 2, \cdots$، $\{X_i;\ i = 1, 2, \cdots\}$ یک فرایند تجدید است.

مثال ۳-۱۹. اگر $\{N_t;\ t \geq ۰\}$ یک فرایند پواسون با نرخ λ باشد توزیع $S_n = \sum_{i=۱}^{n} X_i$ را بدست آورید.
حل.

$$\begin{aligned} F_{S_n}(t) &= \mathrm{P}(S_n \leq t) = \mathrm{P}(N_t \geq n) \\ &= \sum_{k=n}^{\infty} \mathrm{P}(N_t = k) \\ &= \sum_{k=n}^{\infty} \frac{e^{-\lambda t}(\lambda t)^k}{k!} \\ &= ۱ - \sum_{k=۰}^{n-۱} \frac{e^{-\lambda t}(\lambda t)^k}{k!} \qquad t > ۰ \quad n = ۱, ۲, \cdots \end{aligned}$$

در نتیجه بدست می‌آوریم

$$\begin{aligned} f_{S_n}(t) &= F'_{S_n}(t) \\ &= \lambda e^{-\lambda t}\left(۱ + \frac{\lambda t}{۱!} + \frac{(\lambda t)^۲}{۲!} + \cdots + \frac{(\lambda t)^{n-۱}}{(n-۱)!}\right) \\ &- e^{-\lambda t}\left(\lambda + \lambda\frac{\lambda t}{۱!} + \lambda\frac{(\lambda t)^۲}{۳!} + \cdots + \lambda\frac{(\lambda t)^{n-۲}}{(n-۱)!}\right) \\ &= \frac{\lambda^n}{(n-۱)!} t^{n-۱} e^{-\lambda t} \qquad t > ۰ \\ &\Longrightarrow S_n \sim Gamma(n, \lambda) \end{aligned}$$

۳.۳ توزیع یکنواخت و فرایند پواسون

در این بخش نشان می‌دهیم که فرایند پواسون دارای خاصیت آماره‌های ترتیبی است یعنی به شرط اینکه در بازه زمانی $[۰, t]$، n رخداد اتفاق افتاده باشد، محل‌های زمانی رخدادها به صورت (یا مشابه) آماره‌های ترتیبی یک نمونه تصادفی n تایی از توزیع یکنواخت $(۰, t)$ توزیع شده‌اند. این نتیجه یک ابزار قدرتمند برای محاسبات مربوط به توابعی از فرایند پواسون بوجود می‌آورد.

قضیه ۳-۲۰. اگر $\{N_t;\ t \geq ۰\}$ یک فرایند پواسون با نرخ λ باشد و $U_i \sim U(۰, t)$، $i = ۱, ۲, \cdots, n$، آن‌گاه برای هر $t > ۰$ و $n \geq ۱$، داریم

$$(S_۱, S_۲, \cdots, S_n | N_t = n) =^d (U_{(۱)}, U_{(۲)}, \cdots, U_{(n)}).$$

همچنین به شرط اینکه $n+1$–امین رخداد در زمان t اتفاق افتاده باشد، زمان رخدادهای اول تا n–ام به شکل (یا شبیه) آماره های ترتیبی یک نمونه n تایی از توزیع یکنواخت $(0,t)$ است یعنی

$$\begin{aligned} f_{S_1,S_2,\cdots,S_n|S_{n+1}=t}(t_1,t_2,\cdots,t_n) &= \frac{n!}{t^n} \qquad 0<t_1<t_2<\cdots<t_n<t \\ &= f_{U_{(1)},U_{(2)},\cdots,U_{(n)}}(t_1,t_2,\cdots,t_n) \end{aligned}$$

مثال ۳–۲۱. طول عمر قطعه‌ای از یک سیستم صنعتی متغیرتصادفی با توزیع نمایی با پارامتر λ است. اگر قطعه خراب شد آن را فورا با یک قطعه جدید جایگزین می‌کنیم و وقتی قطعه جدید نیز از کارافتاد آن را نیز با یک قطعه جدید دیگر جایگزین می‌کنیم و همین‌طور این روند را ادامه می‌دهیم. فرض کنید هزینه جایگزینی هر قطعه c ریال و نرخ ثابت نزول بانکی پول برابر با α باشد یعنی ریالی که در زمان t صرف شود در زمان صفر (یعنی زمان شروع) دارای ارزش $ce^{-\alpha t}$ است.

الف. مقدار مورد انتظار میزان هزینه جایگزینی تا زمان t را بدست آورید.

ب. مقدار مورد انتظار میزان کل هزینه‌های جایگزینی را محاسبه کنید.

حل. فرض کنید X_i، $i=1,2,\cdots$، طول عمر قطعات متوالی جایگزین شده و $S_n=\sum_{i=1}^n X_i$ زمان از کار افتادن n–امین قطعه مورد استفاده باشد. در این صورت اگر N_t را به صورت تعداد قطعات جایگزین شده تا زمان t تعریف کنیم آنگاه $\{N_t;\ t\geq 0\}$ یک فرایند پواسون با پارامتر λ است.

با توجه به این تعاریف و اطلاعات مساله، هزینه جایگزینی تا زمان t که آن را با W_t نشان می‌دهیم به شکل زیر محاسبه خواهد شد

$$W_t=\sum_{k=1}^{N_t} e^{-\alpha S_n} \qquad t\geq 0.$$

و بنابراین داریم

$$\begin{aligned} \mathrm{E}(W_t) &= \mathrm{E}(\mathrm{E}(W_t|N_t)) \\ &= \sum_{n=0}^{\infty} \mathrm{E}(W_t|N_t=n)\mathrm{P}(N_t=n) \\ &= \sum_{n=0}^{\infty} \mathrm{E}\left(\sum_{k=1}^{N_t} e^{-\alpha S_k}|N_t=n\right)\mathrm{P}(N_t=n) \\ &= \sum_{n=1}^{\infty} \mathrm{E}\left(\sum_{k=1}^{n} e^{-\alpha S_k}|N_t=n\right)\mathrm{P}(N_t=n) \\ &= \sum_{n=1}^{\infty} \mathrm{E}\left(\sum_{k=1}^{n} e^{-\alpha U_{(k)}}\right)\mathrm{P}(N_t=n) \end{aligned}$$

با توجه به اینکه $U_{(1)}, U_{(2)}, \cdots, U_{(n)}$ آماره‌های ترتیبی یک نمونه تصادفی n تایی از $U \sim U(0, t]$ است بدست می‌آوریم

$$\begin{aligned}
\mathrm{E}(W_t) &= \sum_{n=1}^{\infty} \mathrm{E}\left(\sum_{k=1}^{n} e^{-\alpha U_k}\right) \mathrm{P}(N_t = n) \\
&= \sum_{n=1}^{\infty} n \mathrm{E}(e^{-\alpha U_1}) \mathrm{P}(N_t = n) \\
&= \mathrm{E}(e^{-\alpha U_1}) \sum_{n=1}^{\infty} n \mathrm{P}(N_t = n) \\
&= \mathrm{E}(e^{-\alpha U_1}) \mathrm{E}(N_t) \\
&= \left(\frac{1}{t} \int_0^t e^{-\alpha u}\right) \lambda t = \frac{\lambda}{\alpha}(1 - e^{-\alpha t})
\end{aligned}$$

در قسمت دوم سوال، میزان کل هزینه جایگزینی را با C نمایش داده و داریم $C = \sum_{k=1}^{\infty} ce^{-\alpha S_n}$. در نتیجه با توجه به اینکه $S_n \sim Gamma(n, \lambda)$ بدست می‌آوریم

$$\begin{aligned}
\mathrm{E}(C) &= \mathrm{E}(= \sum_{k=1}^{\infty} ce^{-\alpha S_n}) \\
&= c \sum_{k=1}^{\infty} \mathrm{E}(e^{-\alpha S_n}) \\
&= c \sum_{k=1}^{\infty} \left(\frac{\lambda}{\alpha + \lambda}\right)^n \\
&= c \frac{\lambda}{\alpha + \lambda} \left(1 - \frac{\lambda}{\alpha + \lambda}\right)^{-1} = \frac{c\lambda}{\alpha}
\end{aligned}$$