دانشکده‌ی علوم ریاضی

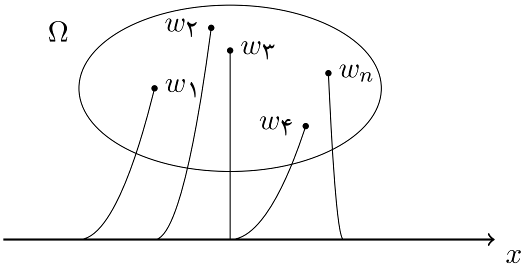
احتمال و کاربردآن ۲۰ اسفند ۹۲

جلسه‌ی ۱۰: متغیرتصادفی

مدرّس: دکتر شهرام خزائی *نگارنده: شراره عزّت‌نژاد و شیوا بناسازنوری*

۱ مقدمه

یک بادکنک پر از هوا را در نظر بگیرید. اگر مجموع مولکول‌های گاز داخل بادکنک را به عنوان فضای نمونه‌ای درنظر بگیریم، می‌توانیم کمیت‌های مختلفی از قبیل جرم، سرعت، انرژی‌جنبشی و اندازه‌حرکت به آنها نسبت دهیم. اما آیا مقدار این کمیت‌ها به‌ازای تمام مولکول‌ها ثابت است؟ به‌وضوح مقدار این کمیت‌ها برای هر مولکول متفاوت از مولکول دیگر است. یا به‌طور مثال دانش‌آموزان یک کلاس را به‌عنوان فضای‌نمونه‌ای در نظر بگیرید. می‌توان کمیت‌های قد، وزن و سن را به آنها نسبت داد که از فردی به فرد دیگر متفاوت است. در واقع مقدار عددی این کمیت‌ها به عضوی که از فضای‌نمونه‌ای انتخاب می‌کنیم، بستگی دارد. به هرکدام از این کمیت‌ها متغیرتصادفی‌گسسته می‌گوییم که به هریک از برآمدهای فضای‌نمونه‌ای، یک عدد حقیقی نسبت می‌دهد.



۲ متغیرتصادفی گسسته

تعریف ۱ *متغیرتصادفی‌گسسته تابعی است مانند $X : \Omega \to R$ که به هر عضو از فضای‌نمونه‌ای گسسته Ω یک عدد حقیقی نسبت می‌دهد.*

نکته ۱ *معمولاً متغیرتصادفی را با حروف بزرگ نمایش می‌دهند. به‌طورمثال در مورد مولکول‌های گاز، متغیرتصادفی سرعت را با V نشان‌می‌دهیم. در مورد مثال دانش‌آموزان کلاس، متغیرهای‌تصادفی قد (برحسب اینچ) و وزن (برحسب کیلوگرم) را می‌توان با H و W نشان‌داد. همچنین می‌توان متغیرهای‌تصادفی جدیدی*

مانند قد برحسب سانتی‌متر و نسبت قد به وزن را برحسب متغیرهای‌تصادفی قبل به‌صورت $۲/۵۴H$ و $\frac{H}{W}$ تعریف‌کرد.

می‌توانیم با استفاده از متغیرهای‌تصادفی، پیشامدهایی بدین‌گونه تعریف کنیم، که $I \subseteq R$:

$$\{X \in I\} = \{\omega \in \Omega | X(\omega) \in I\}$$

به‌عنوان مثال داریم:

$$\{X = x\} = \{\omega \in \Omega | X(\omega) = x\}$$

$$\{a \leq X \leq b\} = \{\omega \in \Omega | a \leq X(\omega) \leq b\}$$

با این نمادگذاری، عباراتی مانند $\Pr\{X \in I\}$ و $\Pr\{X = x\}$ و $\Pr\{a \leq X \leq b\}$ معنا پیدا می‌کنند. حال که مفاهیم تاحدودی روشن‌شده است، به بررسی چند مثال می‌پردازیم.

مثال ۱ *فرض‌کنید یک سکه‌ی سالم را آن‌قدر پرتاب کنیم تا شیر بیاید. اگر در این آزمایش X تعداد پرتاب‌ها باشد، آنگاه X یک متغیر تصادفی‌گسسته است. اگر برآمد شیر آمدن سکه را با H و برآمد خط‌بودن را با T نشان‌دهیم، آنگاه تابع X به صورت زیر خواهدبود:*

$$\Omega = \{H, TH, TTH, ..., \underbrace{TT...T}_{i-۱}H, ...\}$$

$$X(H) = ۱, X(TH) = ۲, ..., X(\underbrace{TT...T}_{i-۱}H) = i$$

احتمال‌های متناظر با این مقادیر به‌صورت زیر مقایسه می‌شوند:

$$\Pr\{X = ۱\} = \frac{۱}{۲}$$

$$\Pr\{X = ۲\} = (\frac{۱}{۲})(\frac{۱}{۲}) = \frac{۱}{۴}$$

$$\Pr\{X = i\} = (\frac{۱}{۲})^i$$

مثال ۲ *آزمایشی را در نظر بگیرید که در آن ۲ سکه‌ی سالم را پرتاب می‌کنیم. اگر در این آزمایش X را تعداد شیرهای رو آمده در نظر بگیریم، مقادیر X به‌ازای برآمدهای آزمایش و احتمال هر یک به صورت زیر خواهد بود:*

$$\Omega = \{TT, TH, HT, HH\}$$

$$X(TT) = ۰, X(TH) = ۱, X(HT) = ۱, X(HH) = ۲$$

به وضوح X فقط می‌تواند مقادیر ۰، ۱ و ۲ را به خود اختصاص دهد.

$$\Pr\{X = ۰\} = \frac{۱}{۴}$$

$$\Pr\{X = ۱\} = \frac{۱}{۲}$$

$$\Pr\{X = ۲\} = \frac{۱}{۴}$$

مثال ۳ *آزمایشی را در نظر بگیرید که در آن دو تاس سالم را می‌ریزیم. متغیر تصادفی X را مجموع برآمدها در نظر می‌گیریم. مشخص‌است که X می‌تواند مقادیری از ۲ تا ۱۲ داشته‌باشد؛ یعنی: $۲ \leq X \leq ۱۲$. تابع احتمال برای این آزمایش به‌صورت زیر است:*

$$\Pr\{X=k\} = \frac{۶-|k-۷|}{۳۶}\ , \quad k = ۱,۲,۳,\ldots,۱۲$$

k	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲
$\Pr\{X=k\}$	$\frac{۱}{۳۶}$	$\frac{۲}{۳۶}$	$\frac{۳}{۳۶}$	$\frac{۴}{۳۶}$	$\frac{۵}{۳۶}$	$\frac{۶}{۳۶}$	$\frac{۵}{۳۶}$	$\frac{۴}{۳۶}$	$\frac{۳}{۳۶}$	$\frac{۲}{۳۶}$	$\frac{۱}{۳۶}$

مثال ۴ *در آزمایش قبل اگر X را ماکسیمم برآمد دو تاس در نظر بگیریم، تابع احتمال به‌صورت زیر است:*

$$\Pr\{X=k\} = \frac{۲k-۱}{۳۶}\ ,\ ۱ \leq k \leq ۶$$

k	۱	۲	۳	۴	۵	۶
$\Pr\{X=k\}$	$\frac{۱}{۳۶}$	$\frac{۳}{۳۶}$	$\frac{۵}{۳۶}$	$\frac{۷}{۳۶}$	$\frac{۹}{۳۶}$	$\frac{۱۱}{۳۶}$

مثال ۵ *از یک‌دسته کارت، سه کارت انتخاب‌می‌کنیم. اگر متغیرتصادفی S را تعداد پیک‌ها درنظر بگیریم، تابع احتمال به‌صورت زیر خواهدبود:*

اگر s را نمایانگر یک خال پیک و n را نمایانگر بقیه کارت‌ها درنظر بگیریم، فضای نمونه‌ای به‌شکل زیر است:

$$\Omega = \{sss, ssn, sns, nss, snn, nss, nns, nnn\}$$

حالت‌اول: انتخاب کارت‌ها با جایگذاری

$$\Pr\{S=۰\} = (\frac{۳}{۴})^{۳}$$

$$\Pr\{S=۱\} = ۳(\frac{۳}{۴})^{۲}(\frac{۱}{۴})$$

$$\Pr\{S=۱\} = ۳(\frac{۳}{۴})(\frac{۱}{۴})^{۲}$$

$$\Pr\{S=۳\} = (\frac{۱}{۴})^{۳}$$

حالت‌دوم: انتخاب کارت‌ها بدون جایگذاری

$$\Pr\{S=k\} = \frac{\binom{۱۳}{k}\binom{۳۹}{۳-k}}{\binom{۵۲}{۳}} \qquad k = ۰,۱,۲,۳$$

۳ تابع جرم احتمال

باتوجه به مثال‌های قبل، می‌توان به هر متغیر تصادفی‌گسسته X یک تابع حقیقی‌مقدار $p_X : R \to R$ تخصیص‌دهیم که به‌صورت $p_X(x) = \Pr\{X = x\}$ تعریف می‌شود. به‌عنوان نمونه در مثال ۴ این تابع

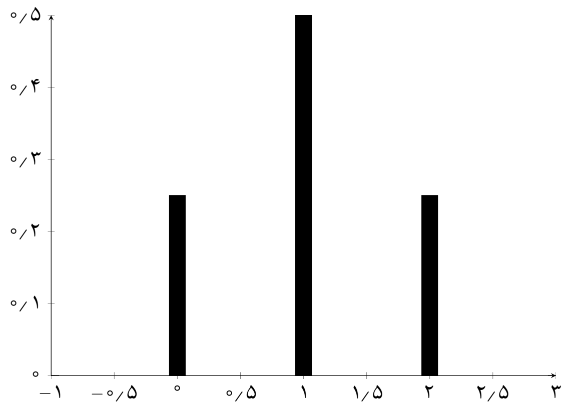
$$p_X(x) = \begin{cases} \frac{۲x-۱}{۳۶} & x \in \{۱, ۲, \ldots, ۶\} \\ ۰ & x \notin \{۱, ۲, \ldots, ۶\} \end{cases}$$

بود. این تابع را تابع جرم احتمال[1] می‌نامند و به‌صورت زیر تعریف می‌شود:

تعریف ۲ *p_X تابع جرم احتمال متغیر تصادفی X ، که مجموعه‌ی مقادیر ممکن آن $\{x_۱, x_۲, ...\}$ می‌باشد، تابعی است از R به R با ویژگی‌های زیر:*
الف) $p_X(x) = ۰$ *اگر* $x \notin \{x_۱, x_۲, ...\}$
ب) $p_X(x_i) = \Pr\{X = x_i\}$ *و* $p_X(x_i) \geq ۰$
پ) $\sum_{i=۱}^{\infty} p_X(x_i) = ۱$

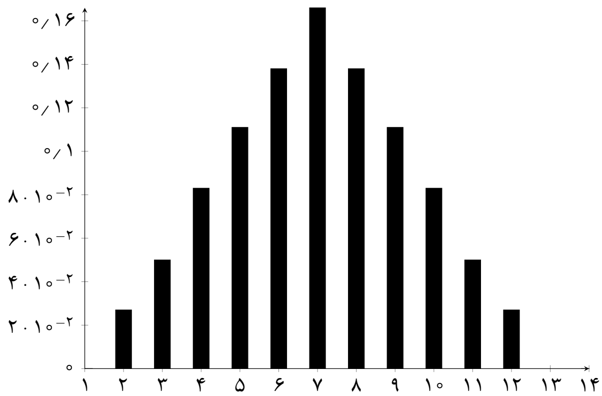
می‌توان برای تابع جرم احتمال هر متغیر تصادفی نمودار رسم کرد.

مثال ۶ *نمودار توابع چگالی احتمال مثال‌های ۲ و ۳ و ۴ در زیر آمده است:*

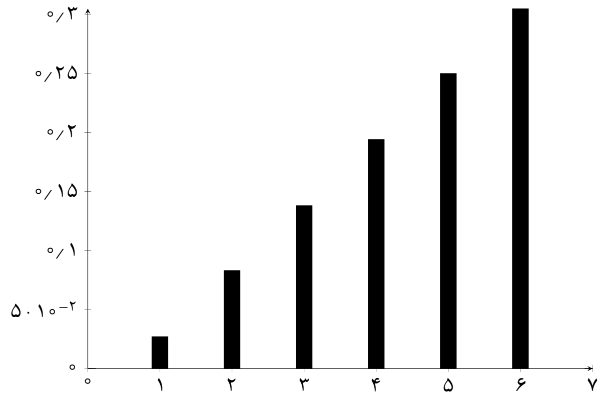


تابع جرم احتمال پرتاب ۲ سکه

[1] probability mass function



تابع جرم احتمال مجموع پرتاب ۲ تاس



تابع جرم احتمال ماکسیمم برآمد ۲ تاس

۴ تابع توزیع تجمعی احتمال

می‌توان تابع حقیقی‌مقدار دیگری بر روی متغیرهای تصادفی تعریف کرد به طوریکه که $F_X(t) = \Pr\{X \le t\}$. در واقع این تابع مشخص‌کننده‌ی مجموع توابع جرم احتمال به‌ازای $x_i \le t$ است.

تعریف ۳ *اگر X یک متغیر تصادفی باشد، آنگاه تابع F_X، که روی بازه‌ی $(-\infty, +\infty)$ به‌صورت $F_X(t) = \Pr\{X \le t\}$ تعریف‌می‌شود را تابع توزیع تجمعی*[۲] *متغیرتصادفی X می‌نامند.*

نکته ۲ *وقتی ابهامی درمورد متغیر تصادفی نداشته‌باشیم، به‌جای $p_X(\cdot)$ و $F_X(\cdot)$ از $p(\cdot)$ و $F(\cdot)$ استفاده می‌کنیم.*

قضیه ۱ *تابع توزیع تجمعی دارای خواص زیر است:*

۱. *F غیر نزولی است.*

[۲] Comulative Distribution Function (CDF)

۲. F در ∞ و $-\infty$ به‌ترتیب به ۱ و ۰ میل می‌کند؛ یعنی:

$$\lim_{t\to\infty} F(t) = ۱$$

$$\lim_{t\to-\infty} F(t) = ۰$$

۳. F از سمت‌راست پیوسته است.

برهان.

۱. برای اثبات غیرنزولی بودن، دو پیشامد $\{X \le n\}$ و $\{X \le m\}$ را درنظر بگیرید بطوری‌که $m < n$ باشد. دراینصورت وقوع پیشامد $\{X \le m\}$ ، وقوع پیشامد $\{X \le n\}$ را نتیجه می‌دهد. پس $\{X \le m\} \subseteq \{X \le n\}$ و درنتیجه $\Pr\{X \le m\} \le \Pr\{X \le n\}$. این نامساوی نتیجه می‌دهد $F(m) \le F(n)$ که این همان شرط غیرنزولی بودن یک تابع است.

۲. فرض کنید $t_۱, t_۲, \ldots$ یک دنباله صعودی از اعداد حقیقی همگرا به ∞ باشد. از خاصیت پیوستگی تابع احتمال (که در جلسات قبل مطرح شده است) می‌توانیم نتیجه بگیریم که پیشامدهای $\{X \le t_n\}$ تشکیل یک دنباله‌ی صعودی از پیشامدها را می‌دهند که به پیشامد $\bigcup_{n=۱}^{\infty} \{X \le t_n\} = \{X < \infty\}$ همگراست. یعنی، $\lim_{n\to\infty} \Pr\{X \le t_n\} = \{X < \infty\}$. بنابراین

$$\lim_{n\to\infty} \Pr\{X \le t_n\} = \Pr(\bigcup_{n=۱}^{\infty} \{X \le t_n\}) = \Pr\{X < \infty\} = ۱,$$

یا به‌عبارت دیگر

$$\lim_{t\to\infty} F(t) = ۱$$

برای $-\infty$ هم می‌توان به طریق مشابه اثبات کرد که آن را به‌عهده‌ی خواننده می‌گذاریم.

۳. برای اثبات پیوستگی از راست، دنباله‌ای نزولی از اعداد حقیقی مانند a_n را درنظر بگیرید که به a همگرا باشد. بنابراین پیشامدهای $\{X \le a_n\}$ تشکیل یک دنباله‌ی نزولی می‌دهند که به پیشامد $\bigcap_{n=۱}^{\infty} \{X \le a_n\} =$ $\{X \le a\}$ همگراست. پس باتوجه به خاصیت پیوستگی تابع حتمال،

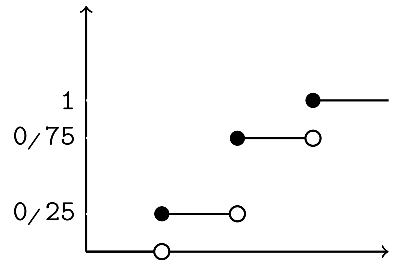
$$\lim_{n\to\infty} \Pr\{X \le a_n\} = \Pr\{\bigcap_{n=۱}^{\infty} \{X \le a_n\}\} = \Pr\{X \le a\}$$

بنابراین

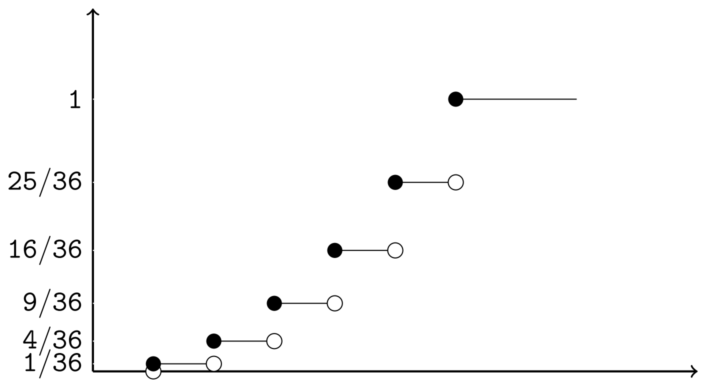
$$\lim_{n\to\infty} F(a_n) = F(a)$$

و این به این معناست که برای هر مقدار $a \in R$ ، $F(a^+) = F(a)$ که این به معنای پیوستگی از راست تابع F است.

■



شکل ۱: تابع توزیع تجمعی پرتاب دو سکه



شکل ۲: تابع توزیع تجمعی ماکسیمم برآمد دو تاس

مثال ۷ *نمودار توابع توزیع‌تجمعی متغیرهای‌تصادفی مثال‌های ۲ و ۴ به‌صورت زیر است:*

نکته ۳ *باتوجه به تعریف F، می‌توان $\Pr\{X < t\}$ را محاسبه کرد. دنباله‌ی پیشامدهای $\{X \leq t-(\frac{۱}{n})\}$ را که یک دنباله‌ی صعودی است و به $\{X < t\}$ = $\bigcup_{n=۱}^{\infty} \{X \leq t - \frac{۱}{n}\}$ همگرا است، درنظر بگیرید. بنابر خاصیت پیوستگی تابع احتمال،*

$$\lim_{n\to\infty} \Pr\{X \leq t - \frac{۱}{n}\} = \Pr\{\bigcup_{n=۱}^{\infty} \{X \leq t - \frac{۱}{n}\}\} = \Pr\{X < t\}$$

یعنی،

$$\Pr\{X < t\} = \lim_{n\to\infty} F(t - \frac{۱}{n})$$

و این بدین معناست که

$$\Pr\{X < t\} = F(a^-)$$

حال با استفاده از تابع توزیع تجمعی و نکته بالا می‌توان احتمال بسیاری از پیشامدهای مطلوب را به‌صورت زیر بدست آورد:

$$\Pr\{X = a\} = F(a) - F(a^-)$$

$$\Pr\{X \leq a\} = F(a)$$

$$\Pr\{X > a\} = ۱ - F(a)$$

$$\Pr\{a < X \leq b\} = F(b) - F(a)$$

$$\Pr\{X \geq a\} = ۱ - F(a^-)$$

$$\Pr\{a \leq X < b\} = F(b^-) - F(a^-)$$

$$\Pr\{a < X < b\} = F(b^-) - F(a)$$

مثال ۸ *فرض‌کنید درون کیسه‌ای، N نوع تمبر (و از هر نوع به تعداد نامتناهی) وجود دارد و شما n بار از کیسه تمبر برداشته‌اید. احتمال اینکه از تمام N نوع، درمیان تمبرهایی که بیرون‌کشیده‌اید وجود داشته باشد، چه‌قدر است؟*

پاسخ *متغیرتصادفی T را به‌شکل زیر تعریف‌می‌کنیم.*
T: تعداد تمبرهایی که باید جمع‌آوری شوند تا از تمام N نوع، داشته باشیم.

$$p_T(n) = \Pr\{T = n\} = \begin{cases} ۰ & n < N \\ \frac{n!}{N^n} & n = N \\ we\ should\ find & n > N \end{cases}$$

هنگامی‌که تعداد تمبرهای جمع‌آوری‌شده از تعداد انواع موجود کمتر باشد، این احتمال صفر است. درحالتی‌که به تعداد انواع، تمبر بیرون کشیده‌باشیم، این احتمال برابر $\frac{n!}{N^n}$ می‌باشد زیرا حالت‌مطلوب زمانی رخ می‌دهد که n تمبری که برداشته‌ایم، متمایز باشند که $n!$ حالت دارد؛ همچنین هرکدام از تمبرهایی که برداشته‌ایم N حالت

دارند، پس کل حالات با N^n برابر می‌شود. اما برای حالت‌سوم نمی‌توان به این سادگی احتمال را محاسبه کرد. اولین گام برای محاسبه‌ی این احتمال، تعریف یک پیشامد به‌شکل زیر است:
A_j: پیشامد اینکه، تمبر jاُم، دربین n تمبر جمع‌آوری‌شده نباشد. ($1 \leq j \leq N$).
داریم:

$$\begin{aligned}
\Pr\{\bigcup_{j=1}^{N} A_j\} &= \Pr\{T > n\} \\
&= \sum_{j=1}^{N} \Pr\{A_j\} - \sum_{j_1<j_2}\sum \Pr\{A_{j_1} A_{j_2}\} \\
&\quad + \ldots + (-1)^{k+1} \sum_{j_1<\ldots<j_k} \ldots \sum \Pr\{A_{j_1} \ldots A_{j_k}\} + \ldots \\
&\quad + (-1)^{N+1} \Pr\{A_1 \ldots A_N\}
\end{aligned}$$

همچنین داریم:

$$\Pr\{A_j\} = (\frac{N-1}{N})^n$$
$$\Pr\{A_{j_1} A_{j_2}\} = (\frac{N-2}{N})^n$$
$$\Pr\{A_{j_1} \ldots A_{j_k}\} = (\frac{N-k}{N})^n$$

بنابراین:

$$\Rightarrow \Pr\{T > n\} = \sum_{i=1}^{N-1} \binom{N}{i} (\frac{N-i}{N})^n (-1)^{i+1}$$

که از جایگذاری عبارت بالا در تساوی زیر احتمال موردنظر بدست‌خواهدآمد.

$$\Pr\{T = n\} = \Pr\{T > n\} - \Pr\{T > n-1\}$$

۵ متغیرهای‌تصادفی درحالت کلی

حال که با مفهوم متغیرتصادفی‌گسسته آشنا شدید، قصد داریم تعریف کلی از متغیرهای تصادفی بیان‌کنیم که به دلیل پیچیدگی، از بازگو کردن آن در ابتدای بخش سرباز زدیم. ابتدا به این نکته توجه کنید.

نکته ۴ *متغیرتصادفی، یک تابع حقیقی‌مقدار است. هر تابع حقیقی‌مقدار روی یک فضای‌نمونه‌ای شمارا، یک متغیرتصادفی است. درمورد فضای‌نمونه‌ای ناشمارا، این عبارت لزوماً درست نیست. برای اینکه تابع حقیقی‌مقدار روی فضای‌نمونه‌ای، متغیرتصادفی باشد، باید معکوس آن تابع روی هر بازه‌ی حقیقی از R یک پیشامد از فضای‌نمونه‌ای باشد.*

تعریف ۴ *فرض‌کنید Ω فضای‌نمونه‌ای یک آزمایش است. تابع حقیقی‌مقدار $X : \Omega \to R$ را یک متغیرتصادفی آن آزمایش می‌گوییم هرگاه برای هر بازه‌ی $I \subseteq R$ ، مجموعه‌ی $\{\omega : X(\omega) \in I\}$ یک پیشامد باشد.*