آمار و احتمال مهندسی

جلسه اول و دوم

مدرس: امیر نجفی

زمینه‌های شکل‌گیری علم آمار و احتمال

زمینه‌های شکل‌گیری علم آمار و احتمال:

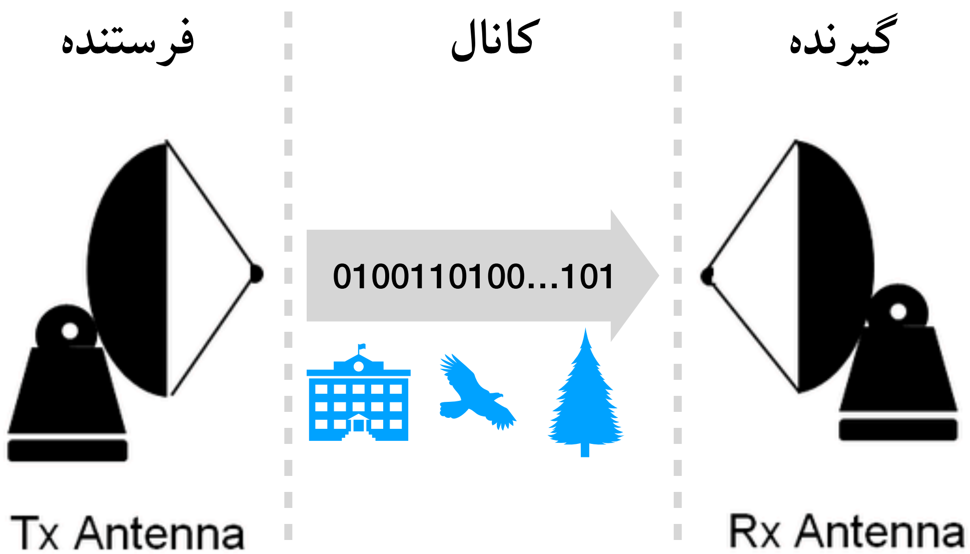
پدیده‌ها و آزمون‌های تکراری

- در صنعت، علوم مختلف نظری و حتی زندگی روزمره عادی با انبوهی از پدیده‌های تکرارشونده روبرو هستیم.

- مثال مخابرات: ارسال بیت‌های پیاپی هزاران میلیارد بیت از یک فرستنده به یک گیرنده رادیویی

- هر بیت ممکن است بسته به شرایط کانال مخابراتی پیچیده‌ای که میان گیرنده و فرستنده وجود دارد، **با یا بدون خطا** به دست گیرنده برسد.

- محاسبه شرایط بروز خطا در یک ارسال خاص، و مهمتر از آن ارائه راهکار برای جلوگیری از بروز آن با علم و فناوری امروز غیرممکن است.

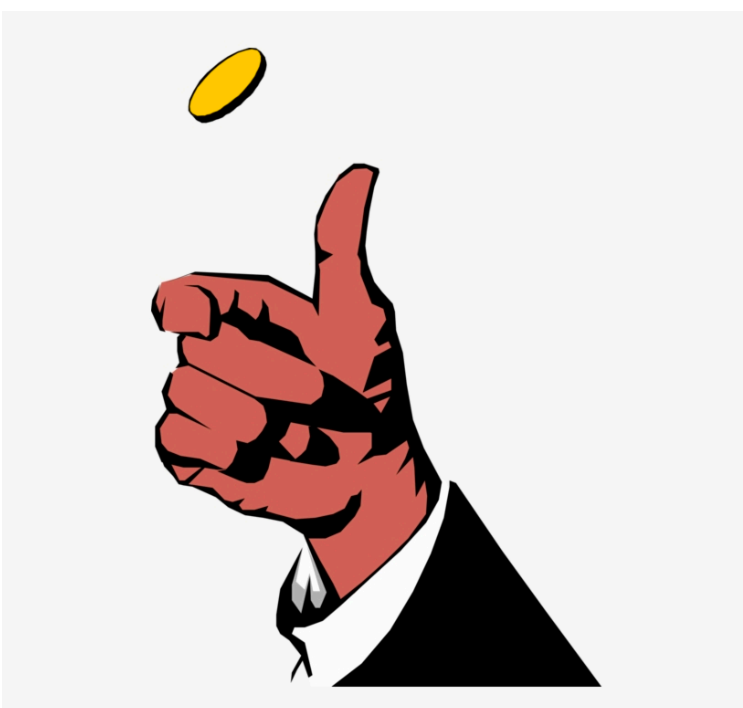


- در مثال فوق دو جنبه مشترک و بسیار مهم وجود دارد:
    - یک پدیده (یا در اصطلاح یک **رویداد**، **رخداد** یا **آزمون**) به **تعداد دفعات زیادی** اتفاق می‌افتد.
    - رخدادهای مختلف یک پدیده، یا اصطلاحاً آزمون‌های منتسب به آن دارای نوعی **استقلال** از یکدیگر هستند.

زمینه‌های شکل‌گیری علم آمار و احتمال:

پدیده‌ها و آزمون‌های تکراری

یا مثال بازی‌های شانسی: در یک پرتاب سکه، آزمون دیدن شیر/خط از معادلات پیچیده و لاینحل فیزیکی و همچنین ماهیت تصادفی پرتاب کننده (عامل انسانی) پیروی می‌کند.



- نیروی وارده چقدر باشد؟
- نسبت گشتاور به پیش‌رانش چقدر است؟
- مقاومت هوا، میزان رطوبت هوا، جریان‌های کوچک هوایی و ... چقدر است؟
- عامل انسانی در چه شرایط بدنی و روحی قرار دارد؟

زمینه‌های شکل‌گیری علم آمار و احتمال:

پدیده‌ها و آزمون‌های تکراری

- در تمامی این موارد با پدیده‌ها (یا در اصطلاح) آزمون‌هایی روبرو هستیم که تحلیل قطعی و فیزیکی آنان خارج از توانایی‌ها و علم روز دنیا است.

- اما از مدت‌ها پیش مشاهده گردیده که در صورت انجام مکرر هر یک از آزمون‌های فوق در **شرایط مطلوب**، کمیت حدی زیر وجود دارد (نتیجه ساده **قانون اعداد بزرگ**):

$$\mathbb{P}(A) = \lim_{N\to\infty} \frac{n_A}{N}$$

  - n_A نشان‌دهنده تعداد دفعاتی است که واقعه A رخ داده است.
  - N تعداد کل آزمون‌های انجام گرفته برای A است.

- به کمیت حدی فوق (در صورت وجود)، **احتمال** واقعه یا موفقیت آزمون A اطلاق می‌گردد که عددی در بازه $[0,1]$ است.

دلیل کارایی احتمال

- چرا علم آمار و احتمال مهم است؟ تعریف احتمال و استفاده از آن (حتی با فرض وجود داشتن)، چه سودی دارد؟ چرا شما لازم است آن را فرا بگیرید؟
  - در اکثریت قریب به اتفاق موارد اشاره شده، بررسی یک پدیده خاص (ارسال یک بیت خاص، نتیجه انداختن یک سکه و ...) مورد توجه نیست.
  - اتفاقاً در این موارد، اشخاص، سازمان‌ها و ... به دنبال **برآیند**، **میانگین** یا **نرخ** رخداد یک پدیده یا موفقیت یک آزمون هستند.
  - علم آمار و احتمال به ما در فهم، و اندازه‌گیری مورد بالا کمک خواهد کرد.

مقدمات ریاضیاتی لازم

مقدمات لازم

برای فراگیری علم آمار و احتمال، قدری مقدمات ریاضی لازم است:

- نظریه مجموعه‌ها (Set theory)
- نظریه اندازه (Measure theory)
- ترکیبیات و شمارش (Combinatorics)

مقدمات لازم

نظریه مجموعه‌ها (Set theory)

- آمار و احتمال علم بررسی و محاسبه احتمال رخدادهاست. مثلاً، در هر بار انداختن یک تاس شش وجهی:
  - احتمال آمدن عدد ۳ (رویداد A)،
  - احتمال ظاهر شدن عددی فرد (رویداد B)،
  - یا احتمال ظاهر شدن عددی بزرگ تر از ۴ (رویداد C)

  می‌توانند مد نظر باشند.

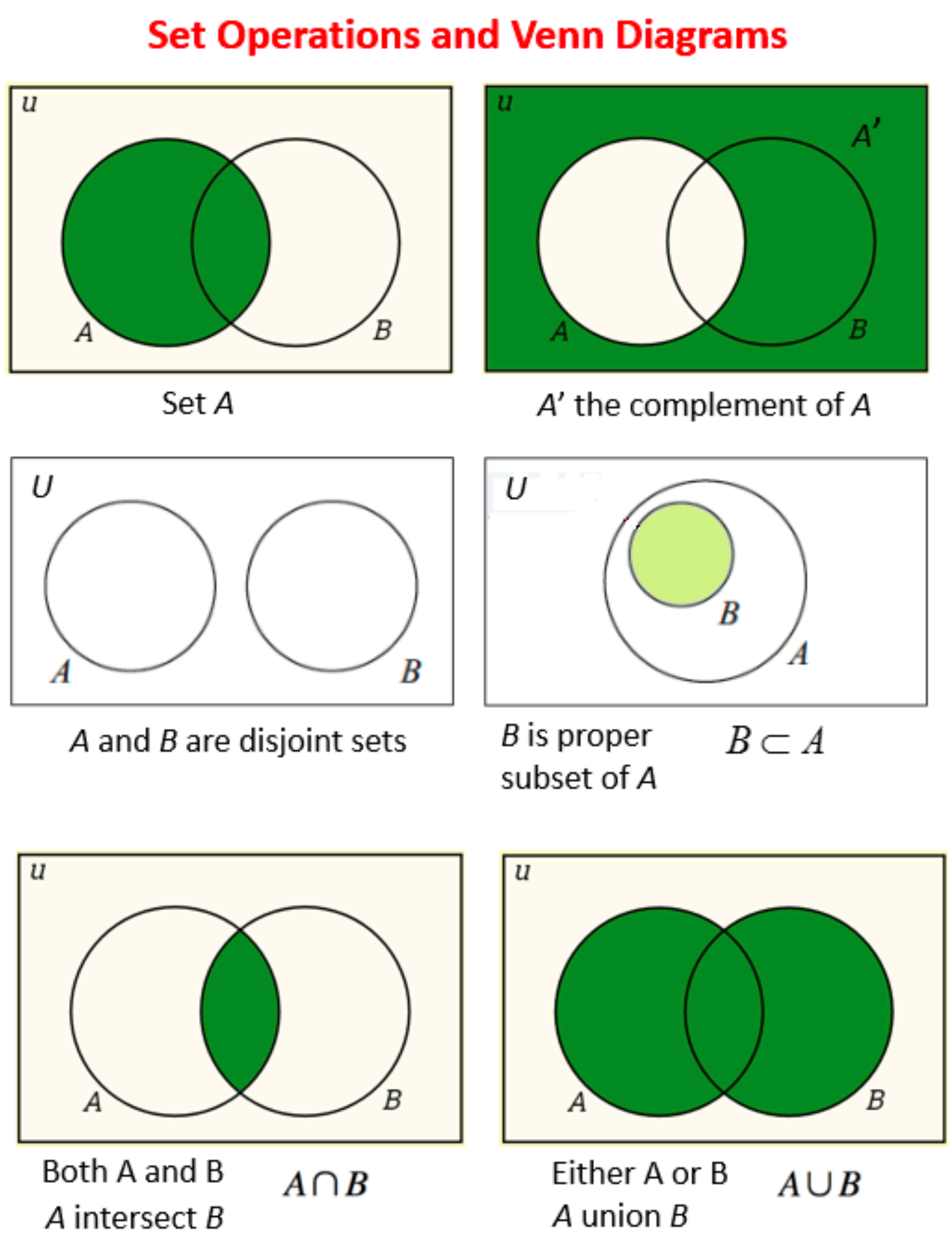
  - در اینجا، رویداد A بخشی (یا در اصطلاح **زیرمجموعه**) رویداد B است.
  - رویدادهای A و C **ناسازگار** هستند. یعنی در صورتیکه یکی به وقوع پیوسته باشد، دیگری نمی‌تواند روی دهد.
  - رویدادهای B و C یکسان نیستند. ولی قدری **اشتراک** دارند.

- به عبارت دیگر، برخی رویدادها (مانند ظهور عددی فرد یا همان رویداد B)، نتیجه اجتماع تعدادی رویداد ساده‌تر هستند؛ در اینجا، روی دادن یکی از اعداد ۱، ۳ یا ۵ در تاس شش وجهی.

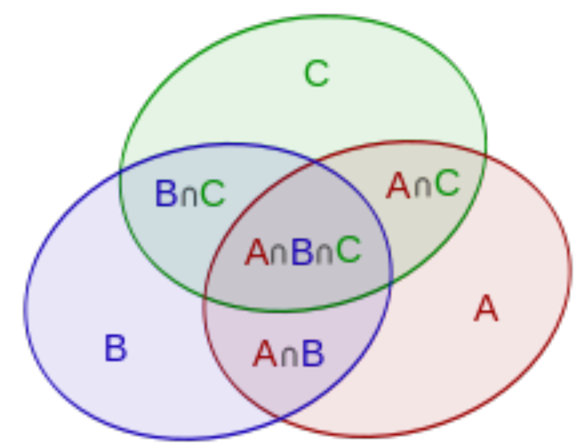
- برای مطالعه اینگونه روابط، نیاز به ابزاری تحلیلی داریم که بتواند مفاهیم فوق را به زبان ریاضی بیان کند. نظریه مجموعه‌ها قادر به انجام این کار است.

مقدمات لازم

نظریه مجموعه‌ها (Set theory)



اصل شمول و عدم شمول

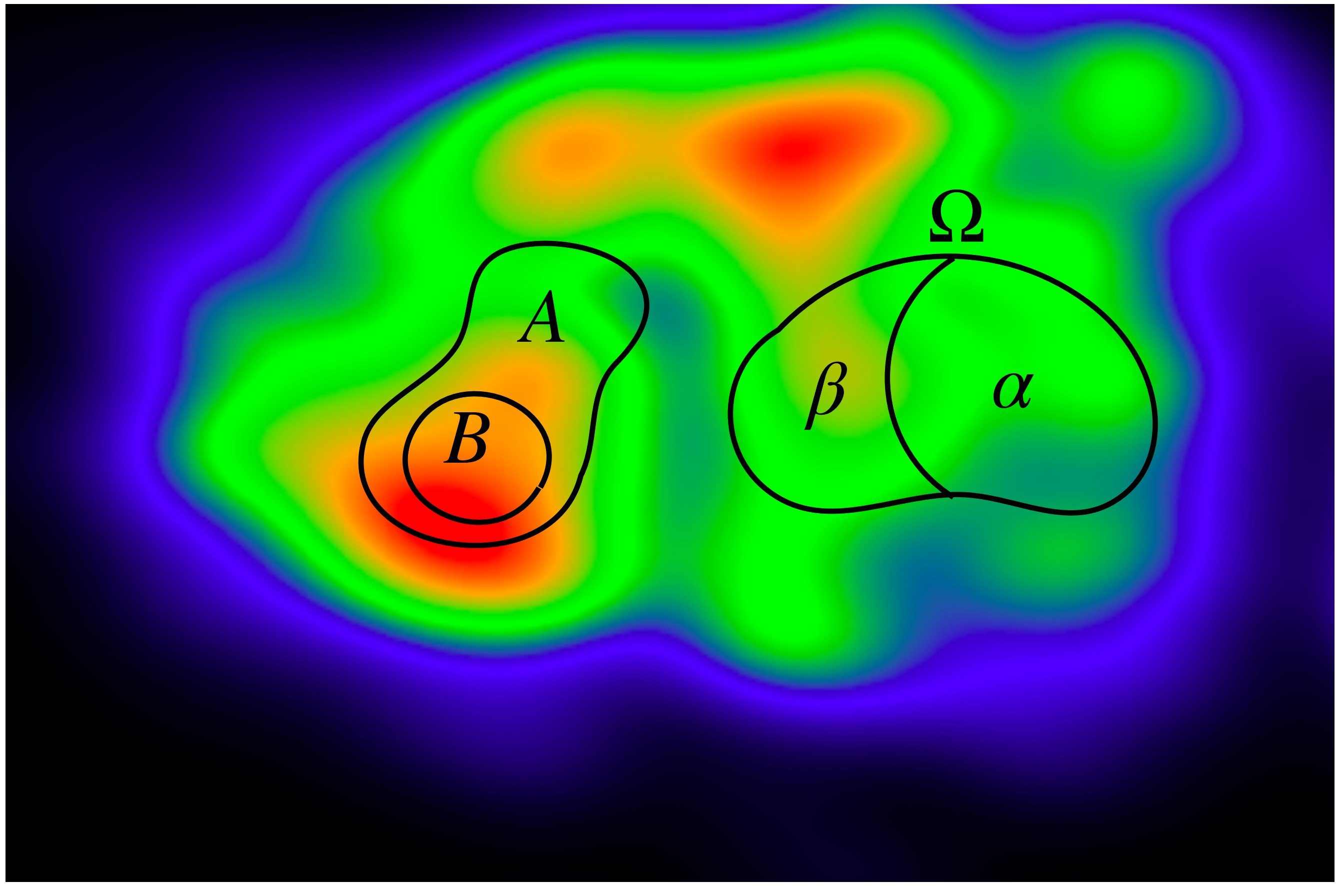


$$
|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|
$$

مقدمات لازم

نظریه اندازه (Measure theory)

- فرض کنید شکل زیر (نمای از بالا) توزیع جرم در یک ناحیه در عمق زمین را نشان دهد.



  - هر چه رنگ تصویر به قرمز نزدیکتر باشد، چگالی جرم بیشتر بوده و عناصر تشکیل‌دهنده آن ناحیه سنگین‌تر یا فشرده‌تر هستند.
  - هر چه رنگ به آبی نزدیک‌تر باشد، فشردگی ذرات کمتر و یا عناصر تشکیل‌دهنده سبک‌تر خواهند بود.

- چه روابط منطقی می‌تواند میان جرم متمرکز در نواحی خط‌کشی‌شده وجود داشته باشد؟
  - آیا می‌توانیم بنویسیم $m(B) > m(A)$ ؟
  - یا امکان دارد $m(\Omega) \neq m(\alpha) + m(\beta)$ ؟

مقدمات لازم

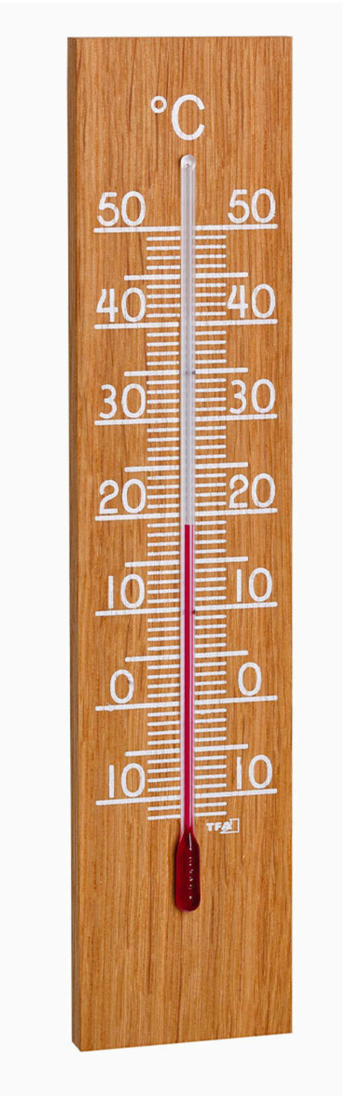
نظریه اندازه (Measure theory)

- روابط اشاره شده در اسلاید قبل نامعقول و غیرقابل قبول هستند. به عبارتی، انتظار ما از معنی و مفهوم جرم، اجازه این گونه روابط را نمی‌دهد.
- دلیل این امر، این حقیقت است که جرم فیزیکی نوعی **اندازه (measure)** است که برای فضا تعریف می‌شود.
- **به شکلی نادقیق**: اندازه‌ها، انواع خاصی از توابع هستند که بر روی تمام زیرمجموعه‌های ممکن از یک مجموعه تعریف می‌شوند و می‌بایست ویژگی‌های خاصی را دارا باشند.
- نظریه اندازه، علم بررسی و تعریف اندازه برای مجموعه‌های مختلف است.
- احتمال نیز نوع خاصی اندازه (measure) است که بر روی تمامی زیرمجموعه‌های ممکن از فضا (یا مجموعه) تمامی رخدادها تعریف می‌شود.

مقدمات لازم

نظریه اندازه (Measure theory) (مثال)

- فرض کنید که شما یک هواشناس هستید و قصد دارید دمای لابی دانشکده کامپیوتر را در یک روز خاص از سال (مثلاً ۱۸ اردیبهشت) بررسی کنید.



- برای این کار، در سال‌های متمادی دما را در روز ۱۸ام اردیبهشت با یک دماسنج بینهایت دقیق جیوه‌ای اندازه گرفته و ثبت می‌کنید.

- به دنباله‌ای مانند شکل زیر خواهید رسید. اعداد حقیقی با بینهایت رقم اعشار... و تقریباً با قطعیت می‌توان گفت هیچ دو عددی با یکدیگر برابر نخواهند بود...

$$\ldots, T_{1397}, T_{1398}, T_{1399}, T_{1400}, \ldots \in \mathbb{R}$$

- چگونه می‌توان به دمای دانشکده احتمال نسبت داد؟ آیا می‌توان در اینجا از تعریف حدی اسلایدهای پیشین استفاده نمود؟

- در اینگونه موارد به **نظریه اندازه** (**measure theory**) نیاز خواهیم داشت.

مقدمات لازم
شمارش و ترکیبیات (Combinatorics)

- به زبان ساده و ابتدایی، ترکیبیات علم شمارش تعداد حالات کوچک‌مقیاسی (microscpoic) است که یک یا چند ویژگی بزرگ‌مقیاس (macroscopic) را برآورده نمایند.
- مثال: اعدادی مانند ۱۲، ۳۴۵ یا ۵۶۳۴۶ مشتمل بر تک‌حالات کوچک مقیاس هستند. اما، برای مثال، مجموعه اعداد فرد پنج‌رقمی بیان‌گر یک مجموعه دارای دو ویژگی بزرگ‌مقیاس است، یعنی: داشتن ۵ رقم، و فرد بودن.
- در بسیاری از موارد، تمامی حالات کوچک‌مقیاس متناظر با یک حالت بزرگ‌مقیاس دلخواه هم‌احتمال (Equally likely) هستند. در این موارد، محاسبه احتمال به یک مسئله شمارش تبدیل خواهد شد.

اصول کولموگروف
(Kolmogorov’s principles)

اصول کولموگروف
(Kolmogorov’s Principles)

- تا پیش از قرن ۲۰ میلادی، تعاریف شهودی و پراکنده فراوانی از احتمال وجود داشت.

- امروزه، بنیان‌های احتمال نوین را بر پایه اصول احتمال کولموگروف (همچنین تحت عنوان Axioms of Probability) تعریف می‌کنیم.

From Wikipedia

Andrey Nikolaevich Kolmogorov
زاده: ۱۹۰۳
فوت: ۱۹۸۷
ریاضیدان شهیر اتحاد جماهیر شوروی

فضای احتمال (Probability Space)

- پیش از انجام هر گونه تحلیل احتمالاتی، ابتدا لازم است تا یک فضای احتمال را تعیین کرده و تا انتهای مسئله ثابت نگه داریم.

- فضای احتمال، به یک سه تایی مرتب مطابق با زیر اطلاق می‌گردد:

$$\left(\Omega\ ,\ \mathscr{F}\ ,\ P\right)$$

فضای احتمال (Probability Space)

فضای نمونه Ω (Sample Space)

- فضای نمونه Ω (Sample space) به مجموعه تمامی خروجی‌های ممکن پدیده تصادفی مورد مطالعه اطلاق می‌گردد.

- مثال:

  - پرتاب سکه: $\Omega = \{H, T\}$
  - پرتاب تاس: $\Omega = \{1,2,\ldots,6\}$
  - پرتاب سه بار متوالی یک سکه: $\Omega = \{H, T\}^3 = \{HHH, HHT, \ldots, TTT\}$
  - دماسنج با دقت بی‌نهایت: $\Omega = (\text{Kelvin's zero}, +\infty)$
  - مکان یک الکترون در فیزیک کوانتوم: سراسر فضای $\Omega = \mathbb{R}^3$

فضای احتمال (Probability Space)

فضای رخدادها $\mathscr{F}$ (Event Space)

- فضای رخدادهای $\mathscr{F}$ (Event space) به مجموعه تمامی رخدادهای ممکن گفته می‌شود. این مجموعه، می‌تواند مجموعه تمام زیرمجموعه‌های فضای نمونه Ω باشد.

- برای یک مجموعه محدود Ω (مانند پرتاب سکه و یا تاس)، داریم:

$$\mathscr{F} = 2^{\Omega} = \{S \mid S \subseteq \Omega\}$$

- مجموعه رخدادها، شامل $\varnothing$ و خود Ω نیز هست.

- در صورتیکه مجموعه Ω، نامحدود ولی شمارا باشد، کماکان می‌توان $\mathscr{F}$ را مجموعه توانی آن در نظر گرفت.

فضای احتمال (Probability Space)

تابع احتمال P (Probability Function)

- تابع احتمال P، یک تابع از فضای رخدادها به اعداد $[0,1]$ است.
یعنی:

$$P : \mathscr{F} \rightarrow [0,1]$$

- تابع احتمال P یک اندازه (measure) است و لذا نمی‌تواند هر تابع دلخواهی صرفاً با دامنه و برد فوق باشد.

- به قوانینی که P می‌بایست از آنان تبعیت کند، قوانین یا اصول کولموگروف اطلاق می‌گردد.

استقلال آماری

استقلال آماری

- در یک آزمایش احتمالاتی، دو واقعه E_1 و E_2 را مستقل از یکدیگر می‌نامیم، اگر و تنها اگر

$$P\left(E_1 \cap E_2\right) = P\left(E_1\right) \times P\left(E_2\right)$$

- در صورت پرتاب یک تاس شش وجهی و مدلسازی تصادفی آن، دو رویداد **فرد** و یا **زوج** آمدن نمی‌توانند از یکدیگر مستقل باشند.

- اما در صورت پرتاب دو تاس پشت سر هم، و با فرض مستقل بودن نتایج آنان، تمام رویدادهای مرتبط با پرتاب اول از تمامی رویدادهای متناظر با پرتاب دوم مستقل می‌شوند.

- وجود رابطه ضرب در احتمال وقایع **مستقل**، در ارتباط تنگاتنگ با **اصل ضرب** است (چرا؟).

چند مثال

مثال:

ترکیب دو یا چند عامل تصادفی با Ω محدود

- یک تاس تصادفی و کاملا متقارن را سه بار پشت سر هم می‌اندازیم. فرض کنید که پرتاب‌ها از یکدیگر مستقل هستند. احتمال اینکه جمع اعداد ظاهر شده برابر با ۸ شود چقدر است؟

- جواب:

  - سه بار پرتاب یک تاس تصادفی:

$$\mathscr{F} = 2^{\Omega}$$

$$\Omega = \{1,2,\ldots,6\}^3 = \left\{(1,1,1),\ldots,(6,6,6)\right\}$$

  - تاس متقارن و استقلال آماری پرتاب‌ها از یکدیگر:

به واسطه استقلال آماری پرتاب‌ها

$$P(i,j,k) = \frac{1}{6}\times\frac{1}{6}\times\frac{1}{6}, \quad \forall i,j,k \in \{1,2,\ldots,6\}$$

به واسطه تقارن تاس

$$P(1) = P(2) = \ldots = P(6) = \frac{1}{6}$$

مسئله روز تولد
(Birthday Party Problem)

- در یک جمع n نفره از افراد مستقل و تصادفی، احتمال اینکه هیچ دو نفری روز تولد یکسان نداشته باشند چقدر است؟

مسئله روز تولد
(Birthday Party Problem)

- جواب:

- احتمال اینکه روز تولد نفر دوم با اولی یکی نباشد: $P(E_2) = 1 - \frac{1}{365}$

  احتمال اینکه روز تولد نفر سوم با اولی و دومی یکی نباشد: $P(E_3) = 1 - \frac{2}{365}$

  ⋮

  احتمال اینکه روز تولد نفر nام با قبلی‌ها یکی نباشد: $P(E_n) = 1 - \frac{n-1}{365}$

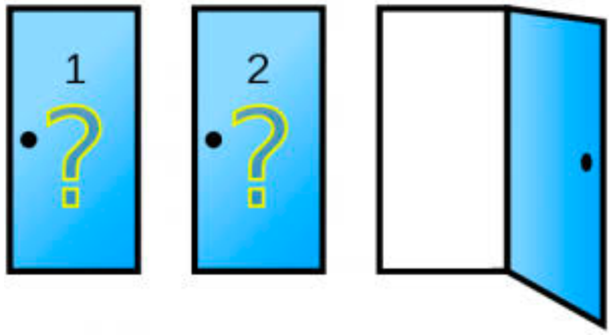
- به دلیل استقلال، روز تولد یک نفر تاثیری در احتمالات روزهای تولد سایر افراد ندارد.

- لذا، احتمال یکی نبودن هیچ روز تولدی:

$$P\left(E_2 \cap E_3 \cap \ldots E_n\right) = \prod_{i=1}^{n-1}\left(1 - \frac{i}{365}\right) \simeq \prod_{i=1}^{n-1} e^{-\frac{i}{365}} = \exp\left(-\frac{n(n-1)}{2 \times 365}\right)$$

مسئله Monty Hall

- فرض کنید که در یک مسابقه تلویزیونی شرکت کرده‌اید و جایزه‌ای پشت یکی از سه در زیر پنهان شده است.



- شما یک در را به صورت کاملاً تصادفی انتخاب می‌کنید...

- سپس، مجری برنامه یکی از دو در باقیمانده را برای شما پوچ می‌کند و تنها یک در باقی خواهد ماند که شما انتخاب نکرده‌اید.

- سوال: آیا انتخاب خود را تغییر می‌دهید؟

احتمال شرطی
(Conditional Probability)

احتمال شرطی (Conditional Probability)

- در بسیاری موارد، **اطلاعاتی** در مورد یک واقعه تصادفی به دست ما می‌رسد. در این شرایط، به دنبال یافتن احتمال‌های به‌روز شده هستیم.

- در پرتاب یک تاس ۶ وجهی متقارن، در صورتیکه پیش‌تر بدانیم واقعه **آمدن یک عدد فرد** $E_{\text{odd}}=$ حتماً اتفاق افتاده است، آنگاه:

    - احتمال رخ‌دادن عدد ۲ صفر خواهد شد،

    - احتمال واقعه رخ‌دادن عددی کوچکتر یا مساوی ۳ بیشتر خواهد شد،

    - و ...

احتمال شرطی (Conditional Probability)

- در آمار و احتمال، **احتمال شرطی** واقعه A به شرط B
  (A given B یا A conditioned on B)
  را به صورت احتمال به‌روز شده رویداد A تعریف می‌کنیم، به شرطی که بدانیم رویداد B حتماً اتفاق افتاده است.

- تعریف می‌کنیم:

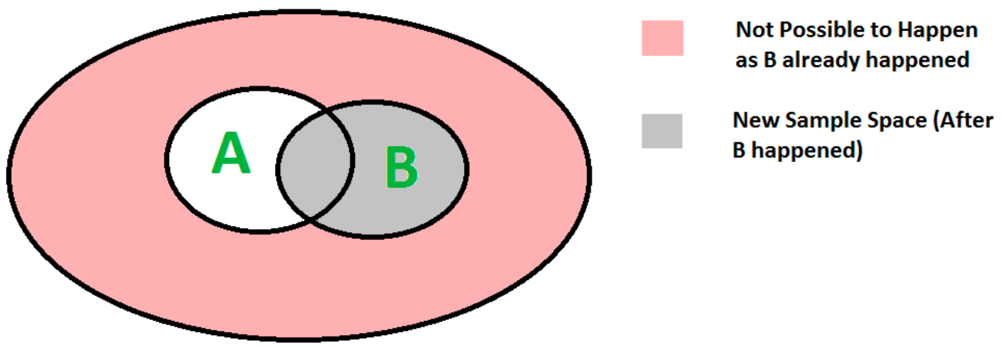
$$P\left(A \mid B\right) = \frac{P(A \cap B)}{P(B)}$$

- دلیل ظاهر شدن مخرج، این واقعیت است که پس دانستن رخ دادن B، گویی فضای نمونه‌ها Ω کوچک‌تر شده است.

  - تنها وقایعی که برابر یا زیرمجموعه B هستند باقی خواهد ماند.

احتمال شرطی (Conditional Probability)

- **نمایش گرافیکی:** بعد از رخداد B، فضای نمونه مؤثر (و به تبع آن، فضای وقایع مؤثر) کوچک‌تر خواهد شد.



- دقت کنید که $P\left(A\,|\,B\right)$ می‌تواند از $P\left(A\right)$ بزرگ‌تر، کوچک‌تر یا مساوی باشد. به این بستگی خواهد داشت که رویداد B، چه میزان اطلاعاتی در مورد رویداد A بدست می‌دهد.

- اگر و تنها اگر دو رویداد A و B مستقل باشند، خواهیم داشت:

$$P\left(A\,|\,B\right)=\frac{P\left(A\cap B\right)}{P\left(B\right)}=\frac{P\left(A\right)P\left(B\right)}{P\left(B\right)}=P\left(A\right)$$

احتمال شرطی

مثال

- فرض کنید احتمالات زیر مربوط به واقعه تصادفات جاده‌ای در شرایط مختلف جوی باشند.

	صاف	مه‌آلود
تردد امن	0.73	0.17
تصادف	0.01	0.09

- سوال ۱: احتمال تصادف چقدر است؟
- سوال ۲: احتمال تصادف به شرط هوای مه‌آلود چقدر است؟
- سوال ۳: آیا احتمال تصادف، از شرایط جوی جاده مستقل است؟

احتمال شرطی

مثال

- سوال ۱: احتمال تصادف چقدر است؟

$$P\left(\text{crash}\right) = \boxed{0.01 + 0.09}$$

- سوال ۲: احتمال تصادف به شرط هوای مه‌آلود چقدر است؟

برابر نیستند.
پس واقعه تصادف
مستقل از مه‌آلود
بودن نیست.

	صاف	مه‌آلود
تردد امن	0.73	0.17
تصادف	0.01	0.09

$$P\left(\text{crash} \cap \text{mist}\right) = 0.09$$

$$P\left(\text{mist}\right) = 0.17 + 0.09$$

$$P\left(\text{crash} \mid \text{mist}\right) = \boxed{\frac{0.09}{0.09 + 0.17}}$$

قانون زنجیره‌ای
(Chain Rule)

قانون زنجیره‌ای (حالت دوتایی و سه‌تایی)

- با توجه به تعریف احتمال شرطی خواهیم داشت:

$$\boxed{P(A,B) = P\left(A\,|\,B\right)P(B)}$$

- در صورتیکه سه واقعه A، B و C را مدنظر قرار داده باشیم، می‌توان رابطه بالا را به صورت تو در تو تکرار کرد:

$$P(A,B,C) = P\left(A\,|\,B,C\right)P(B,C)$$

$$\boxed{P(A,B,C) = P\left(A\,|\,B,C\right)P\left(B\,|\,C\right)P(C)}$$

- **نکته**: وقتی احتمالی را به بیش از یک رویداد مشروط می‌کنیم، منظور این است که فرض کنید تمامی آنان رخ داده باشند.

قانون زنجیره‌ای
(حالت کلی و موارد کاربردی)

- در حالت کلی، برای احتمال رخ دادن توأمان تمامی رویدادهای $A_1, A_2, \ldots, A_n$، خواهیم داشت:

$$P\left(A_1, A_2, \ldots, A_n\right) = P\left(A_1\right) \prod_{i=2}^{n} P\left(A_i \mid A_{i-1}, \ldots, A_1\right)$$

- در مواردی محاسبه احتمالات شرطی بسیار ساده است. مثلاً، یک مدلسازی آماری را از ابتدا بر حسب احتمال‌های شرطی بنا می‌کنیم.

- نکته اضافی: بسیاری از پدیده‌ها، خاصیت Markovian دارند. به عنوان مثال، رویداد A_i می‌تواند به شرط دانستن A_{i-1} از تمام رویدادهای $A_{i-2}, A_{i-3}, \ldots, A_1$ مستقل شود. در این صورت، محاسبه رابطه فوق بسیار ساده می‌شود.

قانون بیز (Bayes’ Rule)

قانون بیز (Bayes' Rule)

- قانون بیز، رابطه ساده ولی بسیار پرکاربردی است که کمک می‌کند تا ترتیب وقایع شرطی در احتمال را عوض کنیم.

- یعنی، با داشتن $P\left(A|B\right)$، و همینطور احتمال‌های حاشیه‌ای $P\left(B\right)$ و $P\left(A\right)$، بتوانیم $P\left(B|A\right)$ را محاسبه کنیم.

$$\boxed{P\left(B|A\right)} = \frac{P\left(B\cap A\right)}{P\left(A\right)} = \boxed{\frac{P\left(A|B\right)P\left(B\right)}{P\left(A\right)}}$$

قانون بیز (Bayes' Rule)

$$P\left(B|A\right)=\frac{P\left(A|B\right)}{P\left(A\right)}P\left(B\right)$$

در رابطه بالا:

- به احتمال $P\left(B\right)$ در اصطلاح **احتمال پیشین** (Prior Probability)،

و

- به احتمال $P\left(B|A\right)$ در اصطلاح **احتمال پسین** (Posterior Probability) می‌گوییم.

قانون بیز (Bayes’ Rule)

- دیدگاه Bayesian در آمار و احتمال بر این اصل مبتنی است که:
برای مدلسازی تصادفی تاثیر دو رویداد تصادفی بر یکدیگر (مثلا تاثیر فضای وقایع Ω_B بر Ω_A، با فضای وقایع کلی $\Omega_B \times \Omega_A$)، صرفاً لازم است تا موارد زیر را داشته باشیم (یا مدلسازی کنیم):

  - احتمال در فضای وقایع منتسب به Ω_B
  - احتمالات شرطی در فضای وقایع منتسب به Ω_A به شرط وقایع B.

- موارد کاربری دیدگاه Bayesian بسیار زیاد هستند:

  - Bayesian inference, Bayesian optimization, Bayesian network, Bayesian Hierarchy و ...

متغیر تصادفی

(Random Variable)

متغیر تصادفی

(مقدمه)

- پیش‌تر دیدیم که مدل‌سازی تصادفی یک واقعه، معادل با تعریف یک فضای احتمال به صورت زیر است:

$$\left(\Omega\ ,\ \mathscr{F}\ ,\ P\right)$$

متغیر تصادفی
(اندازه‌گیری در فضای نمونه)

- دستیابی به تمامی مقادیر لازم برای تشکیل فضای Ω و مشاهده یک پیشامد دل‌خواه $\omega \in \Omega$، بعضاً بسیار سخت و یا حتی غیرممکن است.

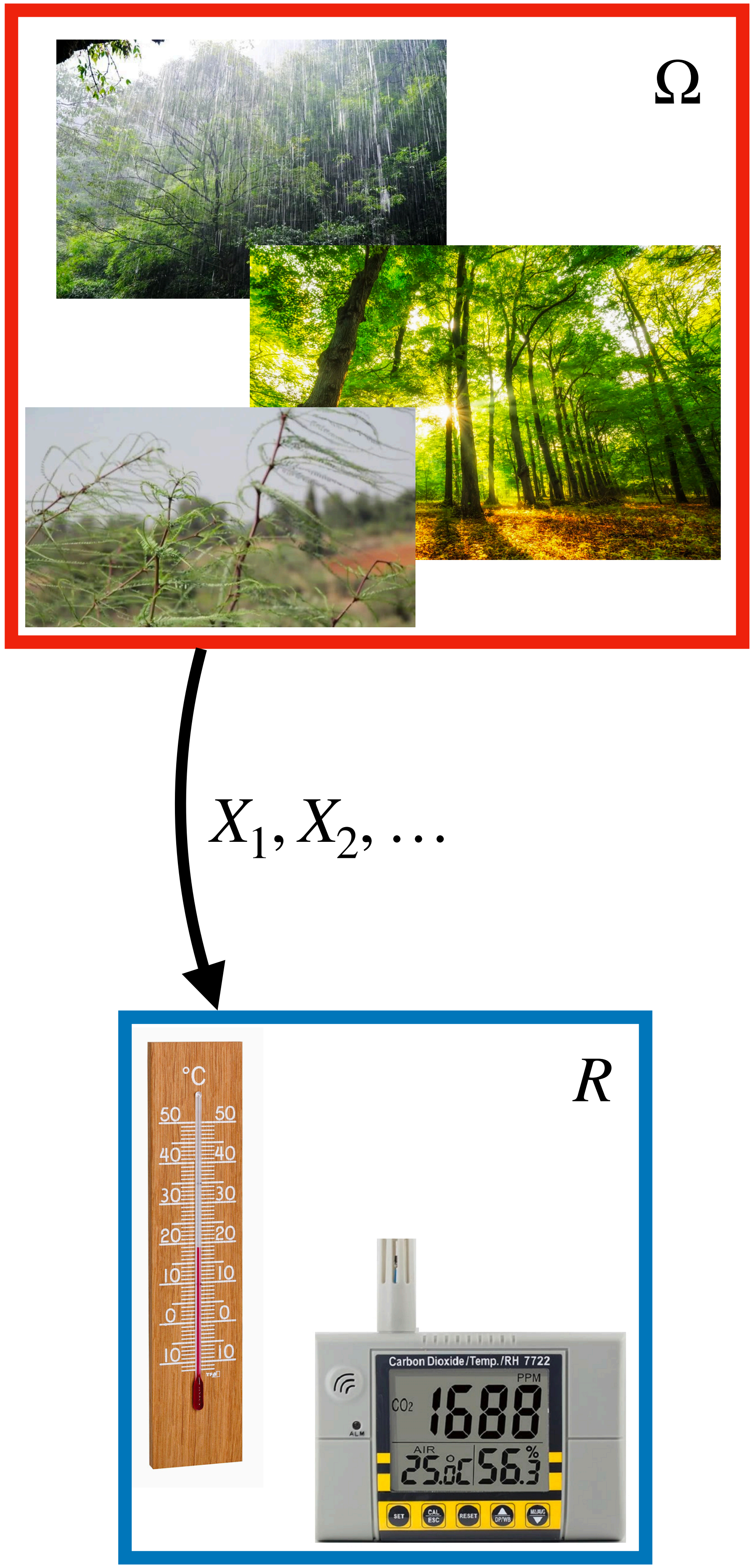
- لذا، به جای مشاهده (یا معادلاً اندازه‌گیری) خود ω، تابعی از آن را مانند $X(\omega)$ اندازه می‌گیریم:

$$X : \Omega \to R$$

- در اینجا، برد تابع X (یعنی R) یک مجموعه عددی دل‌خواه است. R می‌تواند مجموعه اعداد حقیقی $\mathbb{R}$، اعداد طبیعی $\mathbb{N}$ یا مجموعه محدود $\{0,1\}$ باشد.

متغیر تصادفی

(ادامه مثال)



در مدل‌سازی وضعیت آب و هوایی یک ناحیه از کره زمین به صورت تصادفی:

- امکان سنجش فضای Ω اصلی وجود ندارد. لذا، از اندازه‌گیری محدود و در دسترس استفاده خواهد شد:

  - **دماسنج**: نگاشت فضای Ω به یک ناحیه از فضای $\mathbb{R}$ از طریق یک حسگر (sensor) دمایی.

  - **رطوبت سنج**: اندازه‌گیری نادقیق تجمع ملکول‌های آب از طریق حسگر رطوبتی.

  - و ...

متغیر تصادفی به مثابه یک «جعبه سیاه»

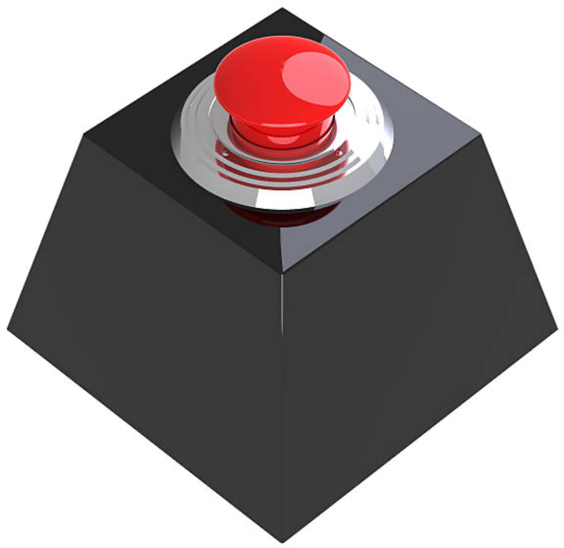
- برای فهم بهتر، می‌توانید یک متغیر تصادفی مانند X را به مثابه یک **جعبه سیاه** در نظر بگیرید که نمی‌دانید در داخل آن چه خبر است...

- هر بار که دکمه روی جعبه سیاه را فشار می‌دهید، یک آزمایش تصادفی درون آن انجام می‌گیرد (یک مقدار $\omega \in \Omega$ به صورت تصادفی مقدار می‌گیرد)، و یک مقدار تصادفی به X می‌دهد.

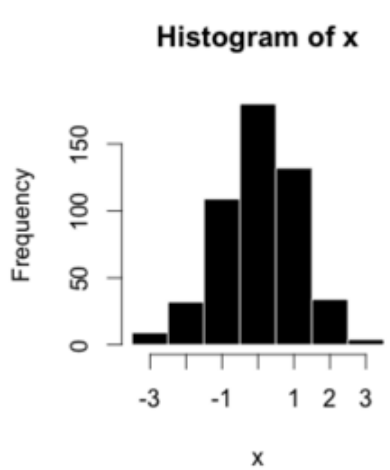
- به این مقدار یک **نمونه** از X گفته می‌شود. در حالت کلی، هر بار که این کار را می‌کنید مقادیر مختلفی بدست خواهید آورد. در واقع X یک مقدار مشخص ندارد، و در هر آزمایش یک مقدار خاص خواهد گرفت.

- اما «**توزیع**» یا distribution مقادیر آن مشخص است.

X



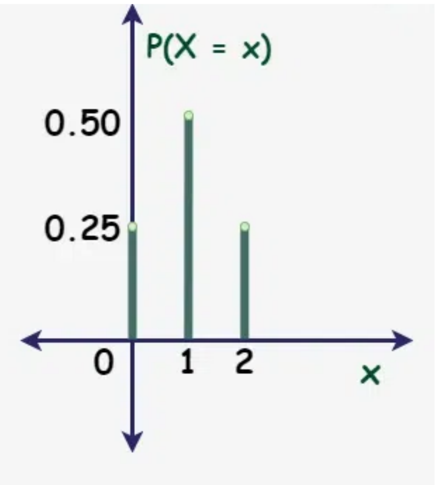
$1.2,\ 3.45,\ -2.1,\ 5,\ -0.4,\ \ldots$



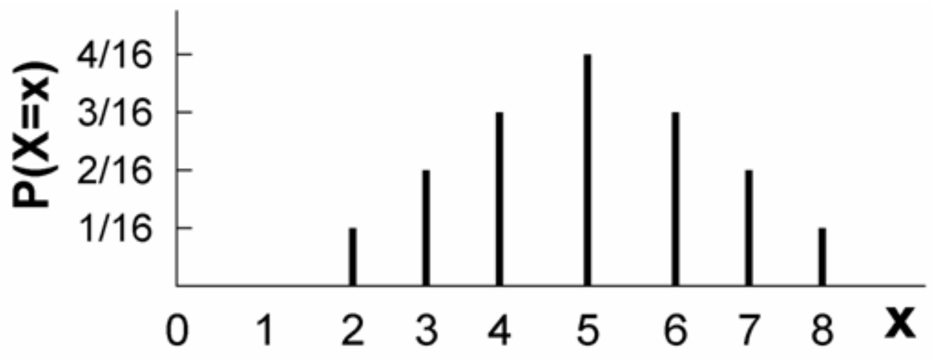
تابع توزیع جرم احتمال (Probability Mass Function – PMF)

- متغیرهای تصادفی می‌توانند پیوسته و یا گسسته باشند. متغیرهای تصادفی گسسته می‌توانند مقادیری از یک مجموعه محدود و یا حداکثر شمارا را داشته باشند.

- در صورتیکه یک متغیر تصادفی گسسته باشد، به طور کامل توسط PMF تعیین خواهد شد.



x	P(x)
2	1/16
3	2/16
4	3/16
5	4/16
6	3/16
7	2/16
8	1/16



تابع توزیع انباشته
(Cumulative Distribution Function)

تابع توزیع انباشته
Cumulative Distribution Function (CDF)

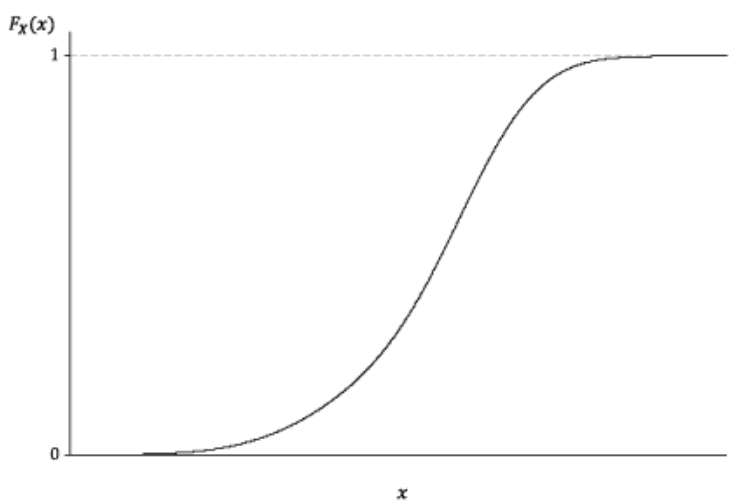
تابع توزیع انباشته (CDF) متغیر تصادفی X را به صورت زیر تعریف می‌کنیم:

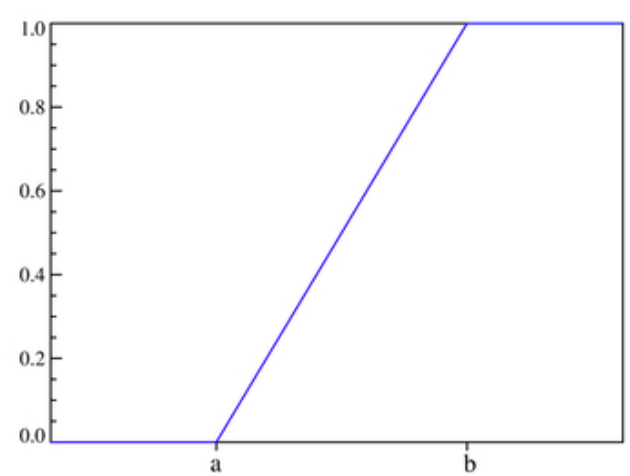
$$F_X(x) = P\{X \leq x\}$$

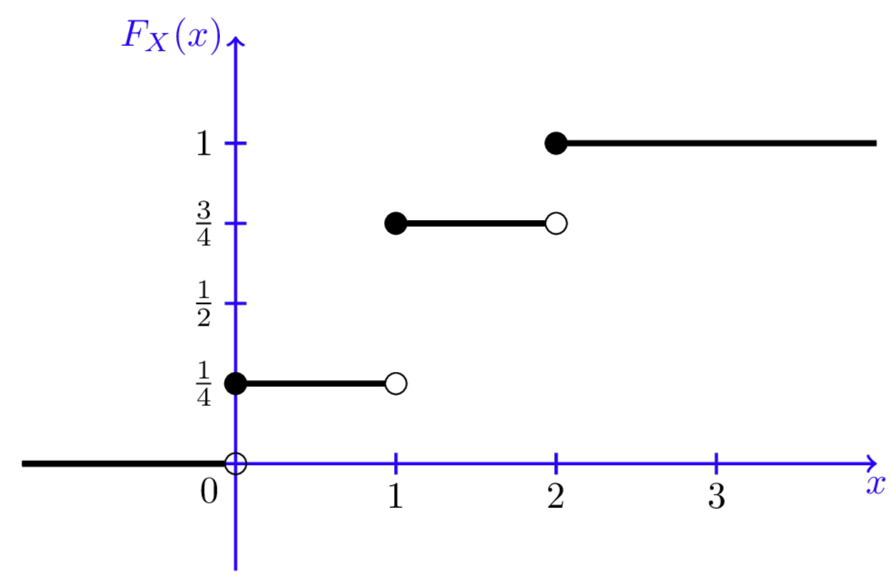
- لذا، F_X یک تابع از اعداد (مثلاً اعداد حقیقی) به بازه بسته $[0,1]$ است.
- **اشتباه رایج**: دقت کنید که در این رابطه، X بیان‌گر منتسب بودن F_X به یک متغیر تصادفی خاص است. ولی x یک عدد (مثلاً یک عدد حقیقی) و در اینجا صرفاً یک متغیر کمکی (در اصطلاح dummy variable) است.

تابع توزیع انباشته
(ویژگی‌ها)

- تابع F_X **صعودی** است، ولی لزوماً **اکیداً صعودی** نیست (چرا؟).







- تابع CDF برای هر متغیر تصادفی X مقادیر حدی زیر را خواهد داشت:

$$F_X(-\infty) = \lim_{x\to-\infty} F_X(x) = 0 \ ,$$
$$F_X(+\infty) = \lim_{x\to+\infty} F_X(x) = 1$$

تابع چگالی احتمال
(Probability Density Function)

تابع چگالی احتمال (PDF) (مفهوم)

- کارکردن با تابع توزیع انباشته (CDF) مقداری سخت است:
  - شهود بالایی نسبت به آن نداریم.
  - اطلاعات محلی (localized) مناسبی از نواحی مختلف فضا بدست نمی‌دهد.

- برای حل این مشکل، از تابع چگالی احتمال (PDF) استفاده خواهد شد:

$$f_X(x) = \lim_{h \to 0^+} \frac{P\left(x < X \leq x + h\right)}{h} = \frac{\mathrm{d}F_X(x)}{\mathrm{d}x}$$

تابع چگالی احتمال (PDF) (مفهوم)

- کارکردن با تابع توزیع انباشته (CDF) مقداری سخت است:
  - شهود بالایی نسبت به آن نداریم.
  - اطلاعات محلی (localized) مناسبی از نواحی مختلف فضا بدست نمی‌دهد.

- برای حل این مشکل، از تابع چگالی احتمال (PDF) استفاده خواهد شد:

مفهوم چگالی دارد

$$f_X(x) = \boxed{\lim_{h \to 0^+} \frac{P\left(x < X \leq x + h\right)}{h}} = \frac{\mathrm{d}F_X(x)}{\mathrm{d}x}$$

تابع چگالی احتمال (PDF)
(خواص و مقایسه با سایر توابع)

- تابع چگالی احتمال (PDF):

- $f_X(x)$ همواره نامنفی است ($f_X \geq 0$).

- $f_X(x)$ لزوماً بین ۰ و ۱ نیست. حتی ممکن است گاهی بی‌نهایت شود (دقت کنید که f_X به خودی خود معنای احتمال ندارد).

- شرط بهنجارش (Normalization criteria):

$$\int_{-\infty}^{+\infty} f_X(x)\,\mathrm{d}x = F_X(+\infty) - F_X(-\infty) = 1$$

تابع چگالی احتمال (PDF)
(خواص و مقایسه با سایر توابع)

- مقایسه توابع **CDF**، **PDF**، **PMF** و اندازه احتمال P (**Probability Measure or Law**):

- **اندازه احتمال P**: عضو سوم فضای احتمال $(\Omega, \mathscr{F}, P)$، که به هر رویداد در $\mathscr{F}$ عددی در بازه $[0,1]$ نسبت می‌دهد که می‌بایست قوانین کولموگروف را رعایت نماید.

- **تابع توزیع انباشته CDF**: در نقطه $x \in R$ به معنای $P\left(X \leq x\right)$ خواهد بود. CDF معنای احتمال دارد و بین $[0,1]$ خواهد بود. همچنین تابعی صعودی است که در $-\infty$ برابر با صفر، و در $+\infty$ برابر با یک است.

- **تابع چگالی احتمال PDF**: مشتق تابع CDF، که همواره نامنفی و عددی در $\mathbb{R}_{\geq 0}$ خواهد بود. PDF مفهوم احتمال ندارد، و صرفاً انتگرال آن معنای احتمال می‌دهد:

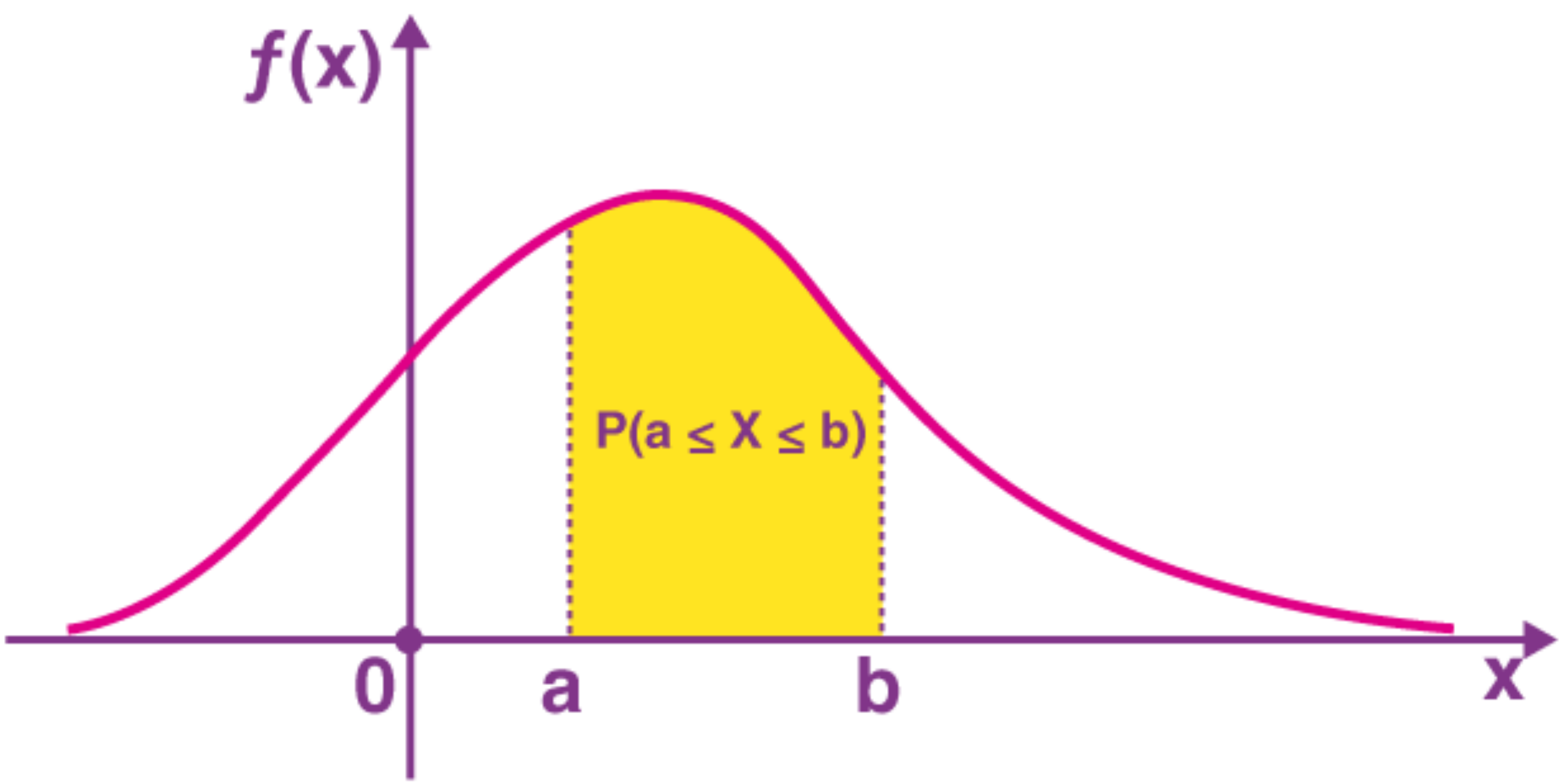
$$\boxed{P\left(a < X \leq b\right) = \int_a^b f_X(x)\,\mathrm{d}x = F_X\left(b\right) - F_X(a)}$$

- **تابع جرم احتمال PMF**: زمانیکه Ω مجموعه‌ای محدود باشد، برای $\Omega = \left\{\omega_1, \ldots, \omega_N\right\}$ به صورت $p\left(\omega_i\right) = P\left(\left\{\omega_i\right\}\right)$ تعریف می‌شود. لذا معنای احتمال دارد و همواره بین ۰ و ۱ است. همچنین، جمع PMF روی کل Ω برابر با ۱ می‌شود.

تابع چگالی احتمال (PDF)
(مفهوم گرافیکی)

- مفهوم گرافیکی PDF یک متغیر تصادفی پیوسته

$$\boxed{P\left(a < X \leq b\right) = \int_a^b f_X(x)\,\mathrm{d}x = F_X\left(b\right) - F_X(a)}$$



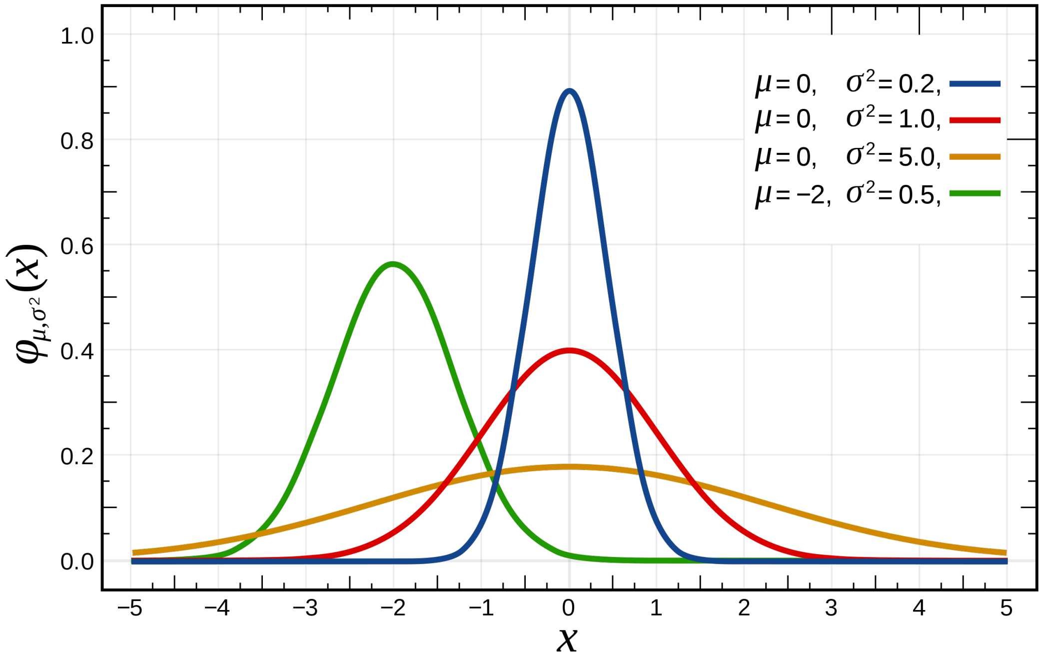
متغیرهای تصادفی معروف
(پیوسته)

متغیرهای تصادفی معروف

(توزیع گاوسی یا نرمال – Gaussian or Normal Dist.)

- یکی از پرکاربردترین توزیع آماری، توزیع احتمال گاوسی یا نرمال (Gaussian or Normal) است.
- توزیع گاوسی $\mathcal{N}\left(\mu, \sigma^2\right)$ با دو پارامتر آزاد مشخص می‌شود:
  - **میانگین** (mean) $\mu \in \mathbb{R}$، و **واریانس** (variance) $\sigma^2 \in \mathbb{R}_{\geq 0}$.
  - به خود مقدار $\sigma \in \mathbb{R}_{\geq 0}$ **انحراف معیار** (standard deviation) اطلاق می‌گردد.
- فرمول $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$:

$$\boxed{f_X(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}}$$

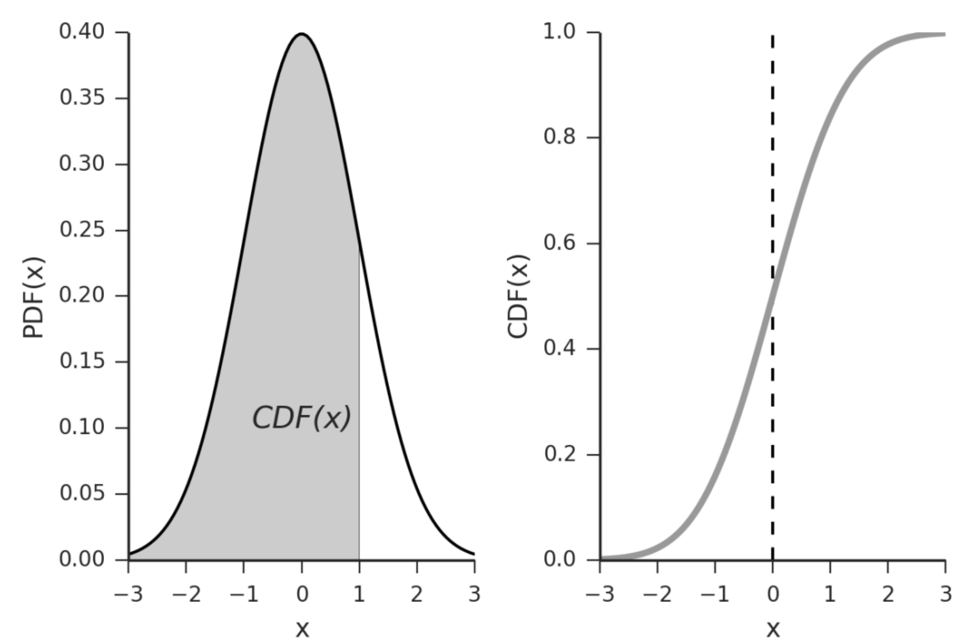


متغیرهای تصادفی معروف (خواص توزیع گاوسی)

- برخی ویژگی‌ها:
  - حول میانگین μ متقارن است. همچنین، بیشنیه چگالی احتمال نیز در همین نقطه رخ می‌دهد.
  - دم توزیع (tail) خیلی سریع افت می‌کند.
  - در طبیعت بسیار ظاهر می‌شود (**قضیه حد مرکزی**).
- تابع توزیع انباشته:

$$\boxed{\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} \mathrm{d}t}$$

$$\boxed{F_X(x) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{x-\mu}{\sqrt{2}\sigma} \right)}$$



متغیرهای تصادفی معروف

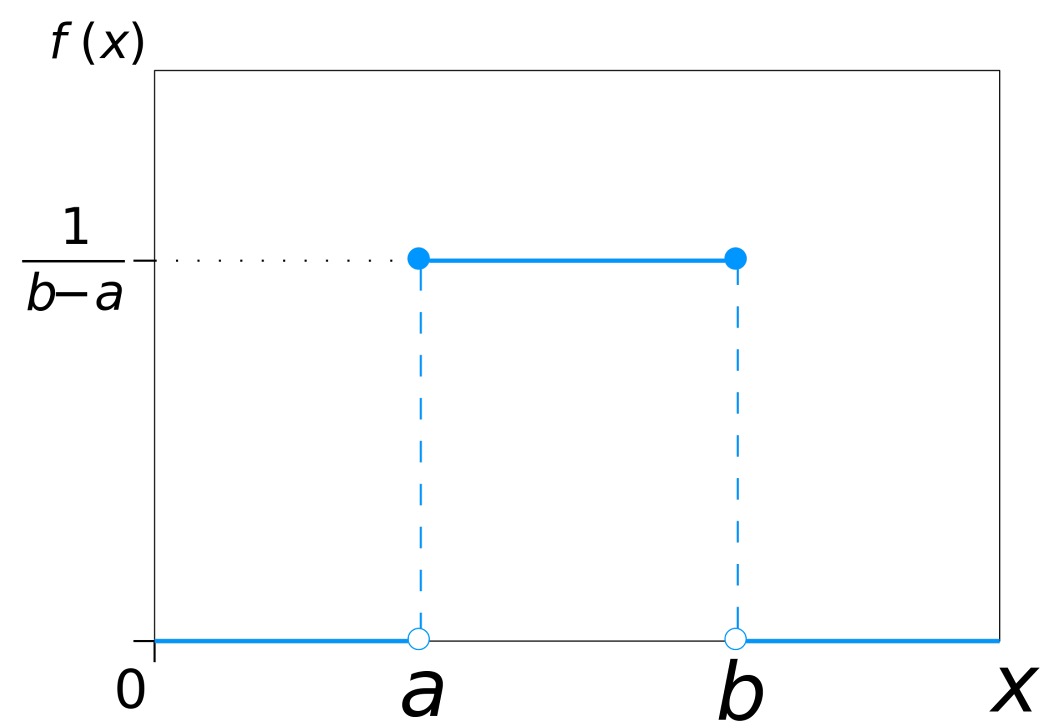
(توزیع یکنواخت - Uniform Dist.)

- متغیر تصادفی یکنواخت، دارای بیشترین آنتروپی (تصادفی بودن) بین تمامی متغیرهای تصادفی با $\Omega = [a, b]$ است.

- همانطور که مشخص است، این متغیر تصادفی دارای دو پارامتر آزاد a و b $(b > a)$ است.

- فرمول $X \sim \text{Uniform}(a, b)$:

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & a \le x \le b \\ 0 & \text{o.w.} \end{cases}$$



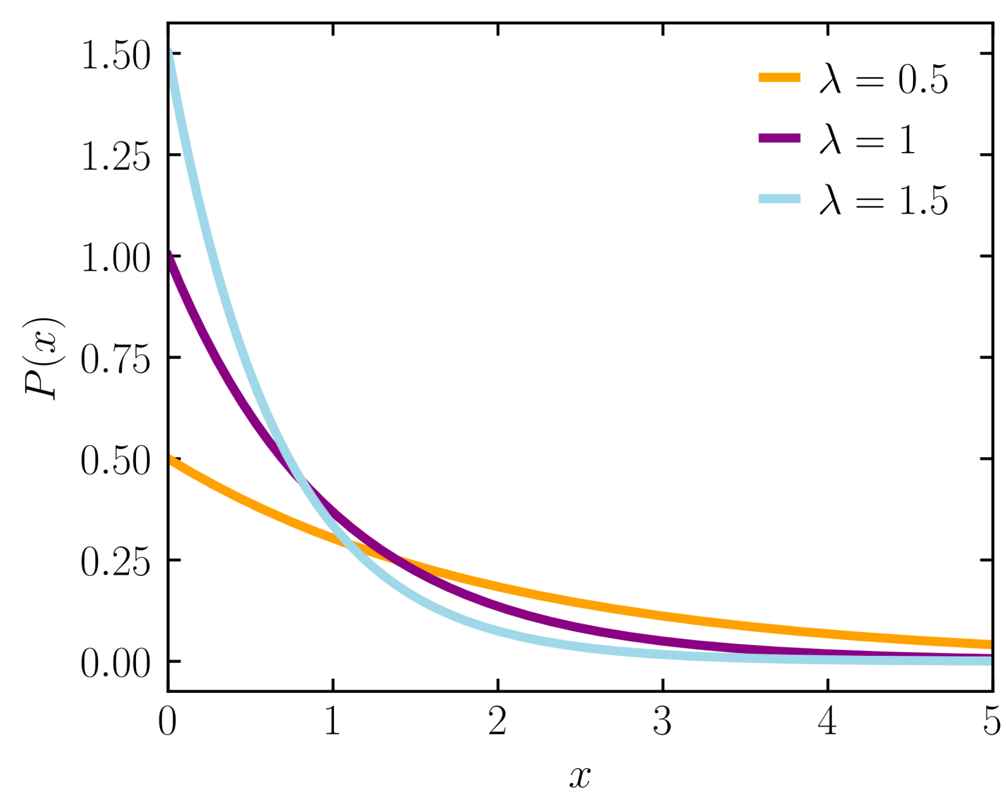
متغیرهای تصادفی معروف

(توزیع نمایی – .Exponential Dist)

- توزیع نمایی (Exponential) دارای یک پارامتر آزاد به صورت $\lambda \in \mathbb{R}_{>0}$ است.

- فرمول $X \sim \text{exponential}\,(\lambda)$:

$$\boxed{f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}}$$



متغیرهای تصادفی معروف
(خواص توزیع نمایی)

- این توزیع معمولاً برای مدل‌سازی زمان وقوع یک رخداد **بی‌حافظه** مورد استفاده قرار می‌گیرد.

- تابع توزیع انباشته:

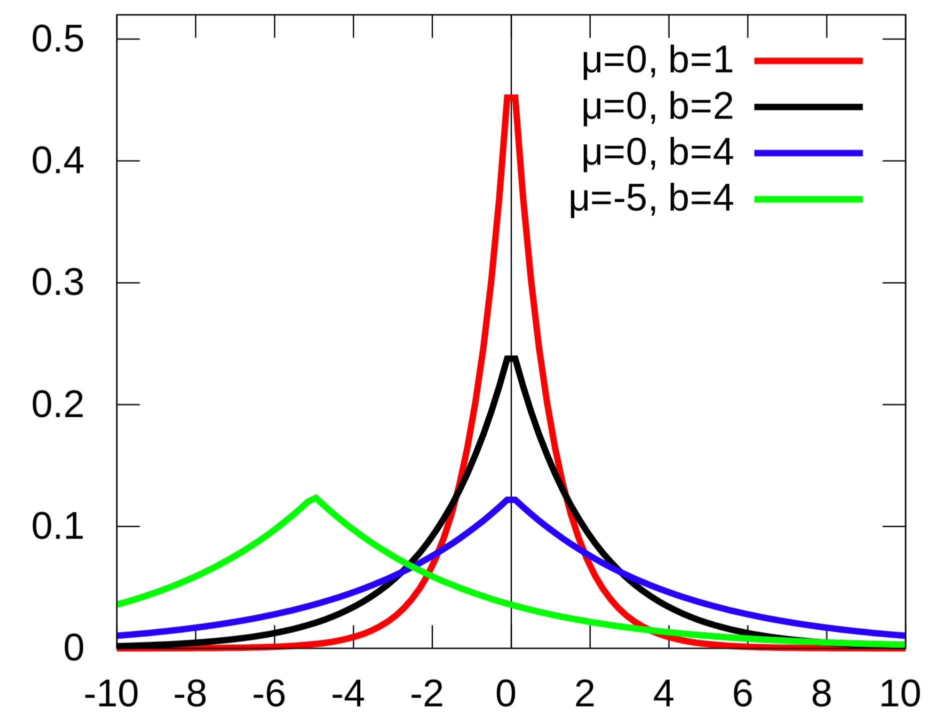
$$\boxed{F_X(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}}$$

متغیرهای تصادفی معروف
(توزیع لاپلاس – Laplace Dist.)

- توزیع لاپلاس (Laplace) دارای دو پارامتر آزاد به صورت $b \in \mathbb{R}_{>0}$ و میانگین $\mu \in \mathbb{R}$ است.

- فرمول $X \sim \text{Laplace}\left(\mu, b\right)$:

$$\boxed{f_X(x) = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}}}$$



متغیرهای تصادفی معروف

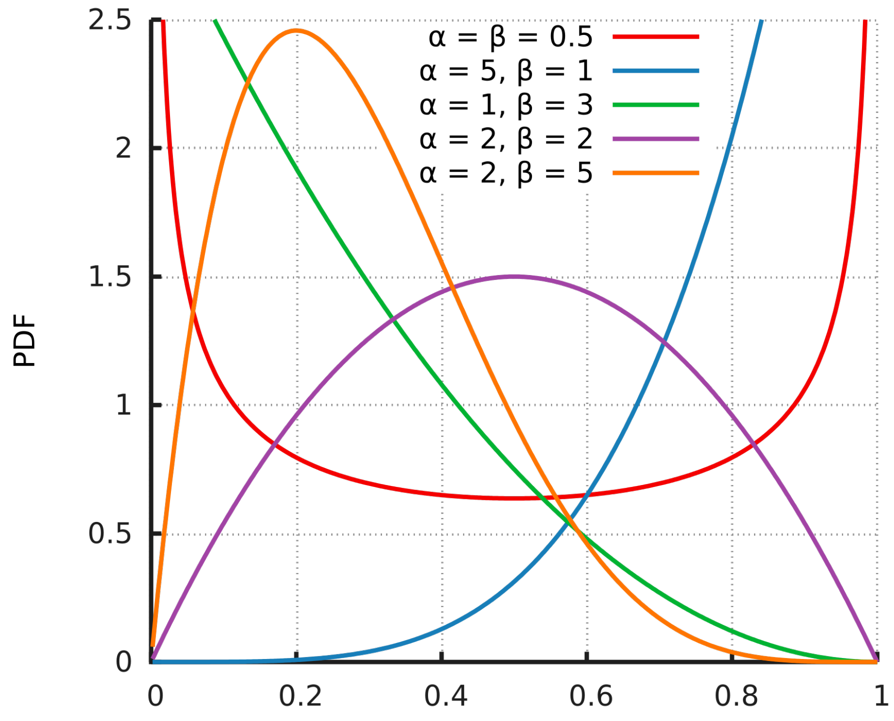
(توزیع بتا – Beta Dist.)

- متغیر تصادفی بتا بین ۰ و ۱ تعریف شده و معمولاً برای مدلسازی احتمالاتی **خود احتمال وقوع یک پدیده** به کار می‌رود.

- این متغیر تصادفی دارای دو پارامتر آزاد $\alpha, \beta \in \mathbb{R}_{>0}$ است. به این پارامترها، پارامترهای شکلی (shape parameters) نیز اطلاق می‌گردد.

- فرمول $X \sim \text{Beta}\left(\alpha, \beta\right)$ در بازه $\left[0,1\right]$ یا $\left(0,1\right)$ (بسته به پارامترها):

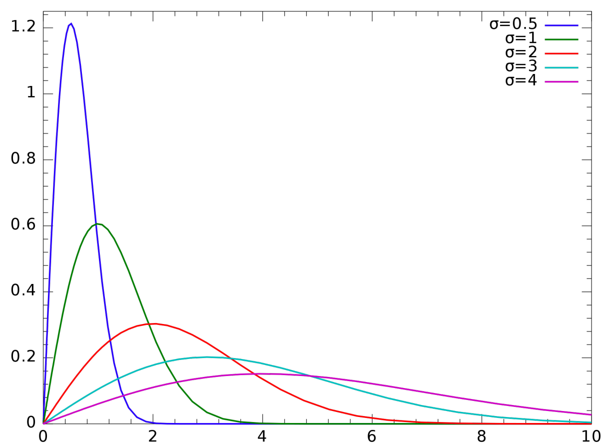
$$\boxed{f_X(x) = \frac{\Gamma\left(\alpha+\beta\right)}{\Gamma(\alpha)\Gamma\left(\beta\right)} x^{\alpha-1}\left(1-x\right)^{\beta-1}}$$



متغیرهای تصادفی معروف

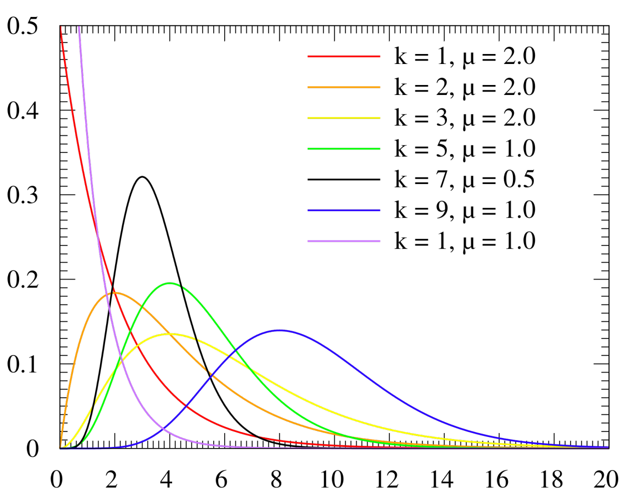
(برخی توزیع‌های کمتر کاربردی...)

- توزیع رایلی (Rayleigh)



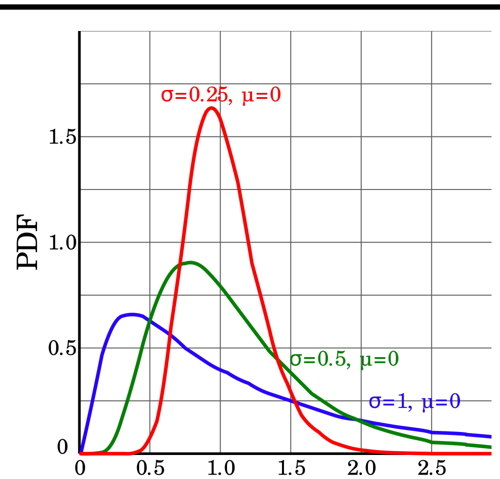
Parameters	scale: $\sigma > 0$
Support	$x \in [0, \infty)$
PDF	$\frac{x}{\sigma^2} e^{-x^2/(2\sigma^2)}$
CDF	$1 - e^{-x^2/(2\sigma^2)}$

- توزیع ارلانگ (Erlang)



Parameters	$k \in \{1, 2, 3, \ldots\}$, shape $\lambda \in (0, \infty)$, rate alt.: $\mu = 1/\lambda$, scale
Support	$x \in [0, \infty)$
PDF	$\frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!}$

- توزیع Log−normal



Parameters	$\mu \in (-\infty, +\infty)$, $\sigma > 0$
Support	$x \in (0, +\infty)$
PDF	$\frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right)$

- و ...

ادامه متغیرهای تصادفی معروف
(گسسته)

متغیرهای تصادفی معروف

(توزیع دوجمله‌ای ـ Binomial Dist.)

- توزیع دوجمله‌ای مربوط به تعداد موفقیت‌های تکرار nباره یک آزمون برنولی است.

- فرمول $k \sim \text{Binomial}\left(n, p\right)$ برای $k = 0,1,2,\ldots,n$:

$$\boxed{P\left(k\right) = \binom{n}{k} p^{k} \left(1 - p\right)^{n-k}}$$

متغیرهای تصادفی معروف

(توزیع چندجمله‌ای – Multinomial Dist.)

- توزیع چندجمله‌ای، تعمیم توزیع دوجمله‌ای برای تکرار nباره آزمون‌هایی است که بیش از دو خروجی ممکن داشته باشند.

- فرض کنید آزمونی با m خروجی ممکن، هر یک با احتمال‌های $p_1, p_2, \ldots, p_m$ (با جمع ۱)، را n بار به صورت مستقل انجام می‌دهیم. هدف، بدست آوردن احتمال وقوع k_i بار از پیشامد iام است ($n = k_1 + \ldots + k_m$).

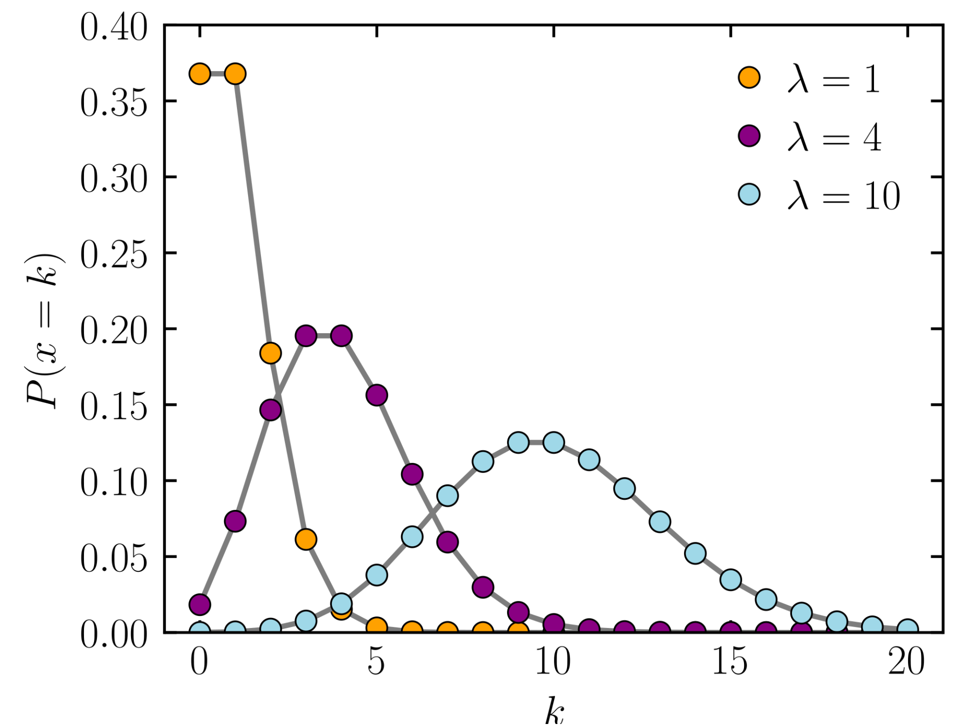
- فرمول $\left(k_1, \ldots, k_m\right) \sim \text{Multinomial}\left(n, p_1, \ldots, p_m\right)$:

$$\boxed{P\left(k_1, \ldots, k_m\right) = \binom{n}{k_1, \ldots, k_m} \prod_{i=1}^{m} p_i^{k_i}}$$

متغیرهای تصادفی معروف
(توزیع پواسون ـ .Poisson Dist)

- متغیر تصادفی پواسون کابرد فراوانی در مدلسازی تعداد پیشامدهای یک ماهیت تصادفی بی‌حافظه در طول زمان دارد.
- این توزیع بر روی اعداد حسابی $k = 0,1,2,\ldots$ تعریف شده و تنها دارای یک پارامتر آزاد $\lambda \in \mathbb{R}_{\geq 0}$ (نرخ) است.
- فرمول $k \sim \text{Pois}\,(\lambda)$:

$$\boxed{P\left(k\right) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0,1,\ldots}$$



متغیرهای تصادفی معروف

(خواص توزیع پواسون)

- توزیع پواسون نتیجه یکی از حالت‌های توزیع دوجمله‌ای (Binomial) است.
- در توزیع دوجمله‌ای $\text{Binomial}\left(n,p\right)$، فرض کنید $n \to \infty$ و $p \to 0$ به طوریکه حاصل‌ضرب آنها عددی ثابت بماند، یعنی $np = \lambda$.
- آنگاه، برای $k \ll n$ داریم:

$$\binom{n}{k} p^k \left(1-p\right)^{n-k} = \frac{n\left(n-1\right)\ldots\left(n-k+1\right)}{k!} p^k \left(1-p\right)^{n-k}$$

$$\simeq \frac{n^k p^k}{k!} \left(1-p\right)^n \simeq \frac{\left(np\right)^k}{k!} e^{-np}$$

متغیرهای تصادفی معروف
(خواص توزیع پواسون)

- توزیع پواسون نتیجه یکی از حالت‌های توزیع دوجمله‌ای (Binomial) است.

- در توزیع دوجمله‌ای $\text{Binomial}\left(n, p\right)$، فرض کنید $n \to \infty$ و $p \to 0$ به طوریکه حاصل‌ضرب آنها عددی ثابت بماند، یعنی $np = \lambda$.

- آنگاه، برای $k \ll n$ داریم:

$$\binom{n}{k} p^k \left(1-p\right)^{n-k} = \frac{n\left(n-1\right)\ldots\left(n-k+1\right)}{k!} p^k \left(1-p\right)^{n-k}$$

$$\simeq \frac{n^k p^k}{k!} \left(1-p\right)^n \simeq \boxed{\frac{\left(np\right)^k}{k!} e^{-np}}$$

در حد دقیقاً برابر با
این مقدار خواهد شد.

پایان جلسات اول و دوم