离散概率

离散数学

马晓星

南京大学 · 计算机科学与技术系

提要

- 直觉概率分析：三门问题
- 概率的公理化：概率空间
- 条件概率与贝叶斯定理
- 随机变量及其期望与方差

本投影片许多内容源自Eric Lehman, F Thomson Leighton and Albert R Meyer, *Mathematics for Computer Science.* https://courses.csail.mit.edu/6.042/spring18/mcs.pdf

三门问题（Monty Hall Problem）

三门问题（Monty Hall Problem）

- 假设你正在参加一个有奖游戏。
  - 你被要求在三扇门中选择一扇，其中一扇后面有一辆车，其余两扇后面则是山羊；
  - 你选择了一道门；
  - 然后知道门后面有什么的主持人，开启了另一扇后面有山羊的门。
  - 他然后问你：“你想改变主意而选择剩下来的这个门吗？”
- 问题是：**改变选择对你来说有利吗？**

进一步明确

- 你在三扇门中挑选一扇。你并不知道门内有什么。
- 主持人知道每扇门后面有什么。
- 主持人必须开启剩下的其中一扇门，并且必须提供你换门的机会。
- 主持人永远都会挑一扇有山羊的门。
  - 如果你挑了一扇有山羊的门，主持人必须挑另一扇有山羊的门。
  - 如果参赛者挑了一扇有汽车的门，主持人随机（概率均匀分布）在另外两扇门中挑一扇有山羊的门。
- 你会被问是否保持原来选择，还是选择剩下的那道门。

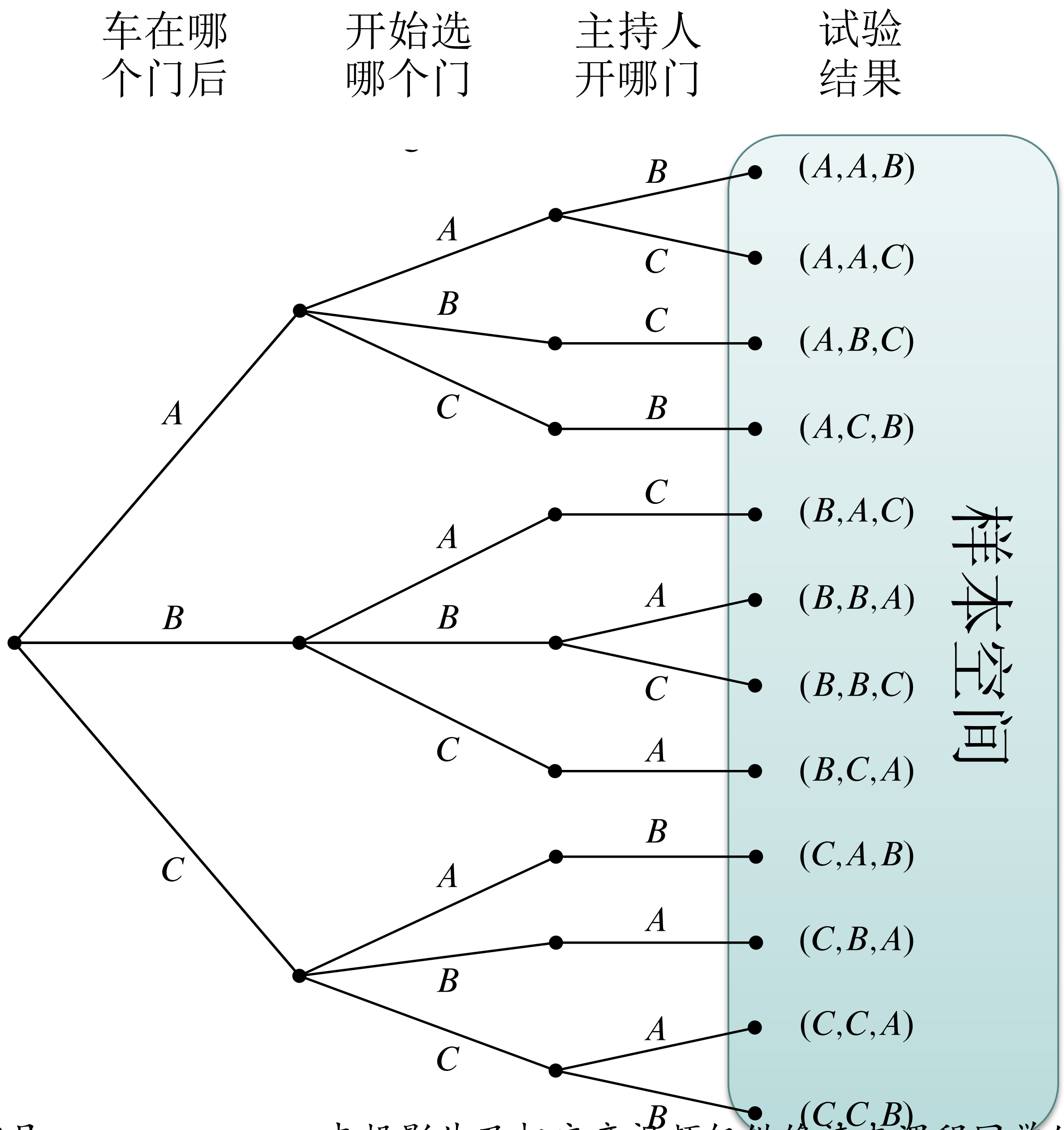
直觉的概率分析

- 四步法
  1. 选定样本空间 (Find the sample space)
  2. 定义相关事件 (Define events of interests)
  3. 确定结果概率 (Determine outcome probabilities)
  4. 计算事件概率 (Compute event probabilities)

第一步： 选定样本空间

- 试验：从一组可能的结果中得出一个结果的过程
  - 试验的某个特定“结果”通常是由若干随机因素的某种选择而导致的。这里
    - 因素一：车在哪个门后？
    - 因素二：你开始选的哪个门？
    - 因素三：主持人打开哪个门？

- 样本空间：所有可能结果的集合


车在哪个门后
开始选哪个门
主持人开哪门
试验结果
A
B
C
A
B
C
A
B
C
A
B
C
B
C
C
B
C
A
C
A
B
A
A
B
(A,A,B)
(A,A,C)
(A,B,C)
(A,C,B)
(B,A,C)
(B,B,A)
(B,B,C)
(B,C,A)
(C,A,B)
(C,B,A)
(C,C,A)
(C,C,B)
样本空间

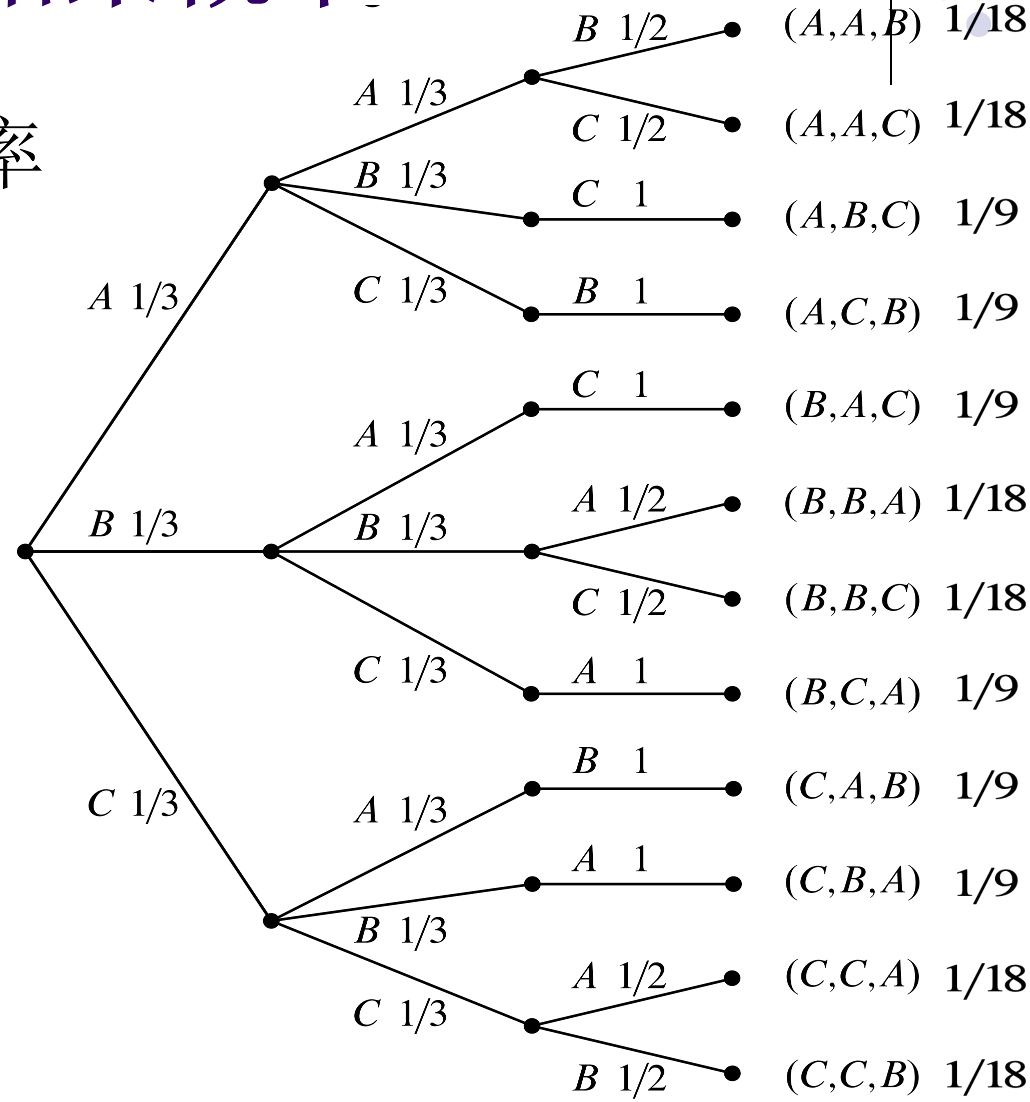
第二步：定义相关事件

- **事件**：样本空间的一个子集
  - 例如：
    - 车在C门后：{(C, A, B), (C, B, A), (C, C, A), (C, C, B) }
    - 第一次就选中有车的门:
      {(A, A, B), (A, A, C), (B, B, A), (B, B, C), (C, C, A), (C, C, B)}
    - 改变选择才赢的情况：
      {(A, B, C), (A, C, B), (B, A, C), (B, C, A), (C, A, B), (C, B, A)}

6对6，似乎换 换都一样？

第三步：确定结果概率

- 给每个边确定概率



- 计算各结果概率

$$\Pr[(A,A,B)]$$

$$=\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{1}{2}=\frac{1}{18}$$

第四步：计算事件概率

$$\Pr[\text{改变选择而赢}]$$

$$= \Pr[(A,B,C)] + \Pr[(A,C,B)] + \Pr[(B,A,C)] + \Pr[(B,C,A)] + \Pr[(C,A,B)] + \Pr[(C,B,A)]$$

$$= \frac{1}{9} + \frac{1}{9} + \frac{1}{9} + \frac{1}{9} + \frac{1}{9} + \frac{1}{9}$$

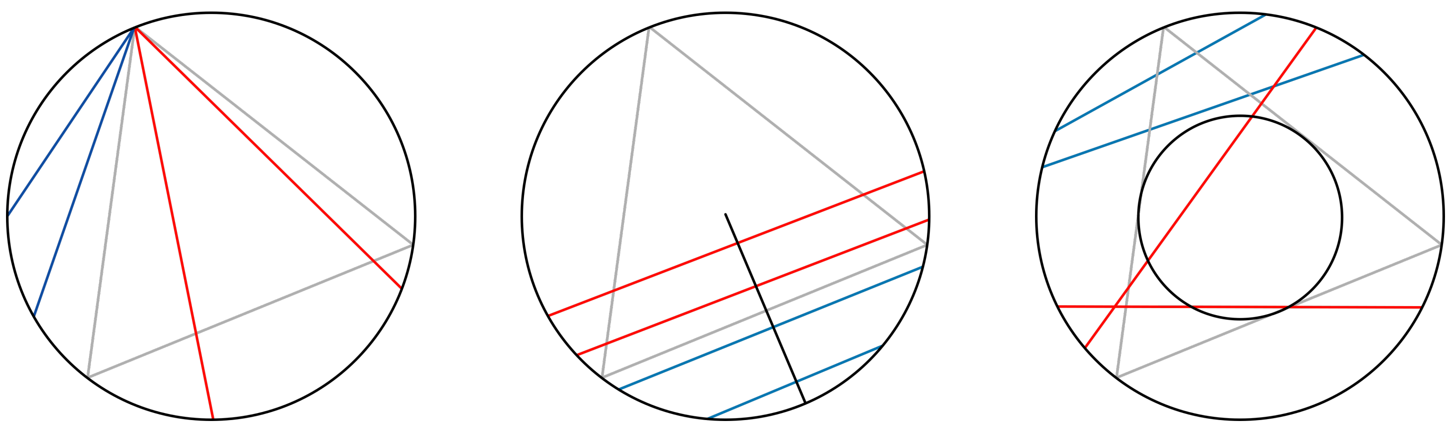
$$= \frac{2}{3}$$

概率的定义

- 古典概率(Laplace)
  - 对于结果具有相同可能性的有限样本空间S, 及其一个子集E, 称事件E的概率为
$$\Pr(E) ::= \frac{|E|}{|S|}.$$
- 频率主义的概率(Frequentism)
  - 令n_E为n次试验中事件E发生的次数. 定义事件E的概率为
$$\Pr(E) ::= \lim_{n\to\infty} \frac{n_E}{n}.$$

概率: 直觉不可靠

- Bertrand Paradox (1888)
  - 考虑一个内接于圆的等边三角形。若随机选圆上的弦，则此弦的长度比三角形的边更长的概率是多少?



方法一: 随机端点 1/3

方法二: 随机半径 1/2

方法三: 随机中点 1/4

概率空间、条件概率、贝叶斯定理

概率空间：
基于集合论的概率定义

- 定义：可数**样本空间** $\mathcal{S}$ 乃一个可数集合。
  - $\mathcal{S}$ 的每一个元素 ω 称为一个**结果**。

- 定义：满足下列条件的函数 $\Pr: \mathcal{S} \to \mathbb{R}$ 称为样本空间 $\mathcal{S}$ 上的一个**概率函数：**
  - $\forall_{\omega\in\mathcal{S}} \Pr[\omega] \geq 0$，且
  - $\Sigma_{\omega\in\mathcal{S}} \Pr[\omega] = 1.$

- 定义：$\mathcal{S}$ 的一个子集 $E \subseteq \mathcal{S}$ 称为一个**事件**。
  - **事件 E 的概率** $\Pr[E] ::= \sum_{\omega\in E} \Pr[\omega]$

基于集合论的概率计算

- 定理 1：设 E 是样本空间 $\mathcal{S}$ 中的一个事件，事件 $\bar{E}$（事件 E 的**补事件**）的概率为：

$$\Pr[\bar{E}] = 1 - \Pr[E]$$

- 定理 2：设 E_1 和 E_2 是样本空间 $\mathcal{S}$ 中的事件，那么

$$\Pr[E_1 \cup E_2] = \Pr[E_1] + \Pr[E_2] - \Pr[E_1 \cap E_2]$$

基于集合论的概率计算

- 例：从不超过**100**的正整数中随机选一个,它能被**2**或**5**整除的概率?

- 解：设E_1是选出一个被2整除的事件，E_2是选出一个被5 整除的事件。则$E_1 \cap E_2$是选出一个被10整除的事件。

$$\Pr[E_1 \cup E_2] = \Pr[E_1] + \Pr[E_2] - \Pr[E_1 \cap E_2]$$

$$= \frac{50}{100} + \frac{20}{100} - \frac{10}{100} = \frac{3}{5}$$

均匀分布

- 定义：假设S是一个含 n 个元素的样本空间. **均匀分布** (*uniform distribution*) 赋给 S 中每个结果$1/n$ 的概率.
  - 举例：对于均匀的硬币 $\Pr[H] = \Pr[T] = \frac{1}{2}$
  - 举例：公平的骰子 $\Pr[X] = \frac{1}{6}$， $X = 1 \cdots 6$
- 均匀分布下事件的概率可通过对其中的元素计数求得

条件概率

- 定义：设E和F是事件,且$\Pr[F] > 0$. E在**给定 F 条件下的概率**, 记作$\Pr[E \mid F]$, 定义为

$$\Pr[E \mid F] ::= \frac{\Pr[E \cap F]}{\Pr[F]}$$

条件概率

- 例：在至少有一个男孩的条件下,有两个孩子的家庭正好 均是男孩的条件概率?假设BB, BG, GB, 和GG是等可能的。

- 解：令E是家庭有两个男孩的事件,F是家庭至少有一 个男孩的事件。我们有$E = \{BB\}$, $F = \{BB, BG, GB\}$, $E \cap F = \{BB\}$.
  - $\Pr[F] = \frac{3}{4}$, $\Pr[E \cap F] = \frac{1}{4}$
  - $\Pr[E \mid F] = \frac{\Pr[E \cap F]}{\Pr[F]} = \frac{1}{3}$

三门问题的条件概率分析

- (错误分析) 假设你选A门. 主持人开B门, 门后是羊. 根据前面的分析, 你选A门且B门后是羊有三种情况:

  (A, A, B), (A, A, C), (C, A, B)

  于是 Pr[改变选择而赢 | 选A门且B门后是羊]

  =1/9 / (1/18+1/18+1/9) = 1/2 ???

- 错在何处? 条件选错了. 不是“选A门且B门后是羊”而应是“选A门且主持人开B门”. 就是说, 多包含了(A, A, C).

  Pr[改变选择而赢 | 选A门且主持人开B门]

  = 1/9 / (1/18+1/9) = 2/3

贝叶斯定理

- 设 E 和 F 是样本空间 $\mathcal{S}$ 中的事件, $\Pr[E] \neq 0$, $\Pr[F] \neq 0$, 则

$$\Pr[F \mid E] = \frac{\Pr[E|F]\Pr[F]}{\Pr[E]}$$

$$= \frac{\Pr[E|F]\Pr[F]}{\Pr[E|F]\Pr[F] + \Pr[E|\bar{F}]\Pr[\bar{F}]}$$

贝叶斯定理的推导

- 由条件概率定义

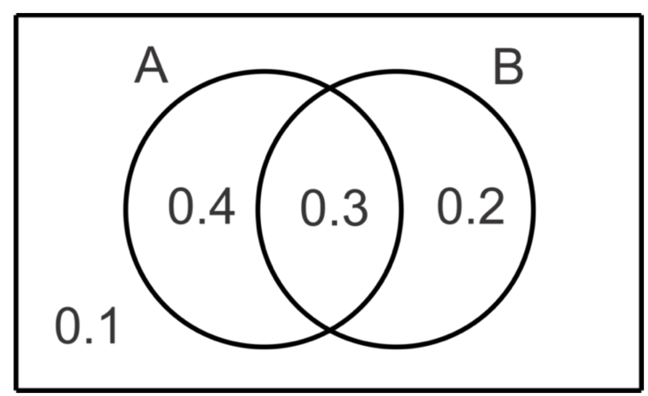
$$\Pr[\,F \mid E\,]\Pr[E] = \Pr[F \cap E]$$
$$= \Pr[E \cap \mathrm{F}] = \Pr[\,E \mid F\,]\Pr[F]$$

- 又

$$\Pr[E] = \Pr[(E \cap F) \cup (E \cup \bar{F})]$$
$$= \Pr[(E \cap F)] + \Pr[(E \cap \bar{F})]$$
$$= \Pr[E \mid F]\Pr[F] + \Pr[E \mid \bar{F}]\Pr[\bar{F}]$$

贝叶斯定理

- 一些常用说法
  - $\Pr[A]$ 是 A 的**先验概率**。之所以称为“先验”是因为它不考虑任何 B 方面的因素。
  - $\Pr[A \mid B]$ 是已知 B 发生后 A 的条件概率或**后验概率**。
  - $\Pr[B \mid A]$ 是已知 A 发生后 B 的条件概率或**后验概率**。
  - $\Pr[B]$ 是 B 的**先验概率**，也作**标准化常量**（normalizing constant）。



贝叶斯定理的应用

- 假设有一种罕见的疾病，100,000人只有1人会得这种病。如果某人得了此病，检测准确率高达99%；如果某人没有得此病，检测准确率为99.5%.
  - 疾病检测呈阳性，得此病的概率多大?
  - 疾病检测呈阴性，没有得此病的概率多大?

解：设D是此人得此病的事件，E是疾病检测呈阳性的事件。需要计算 $\Pr[D \mid E], \Pr[\overline{D} \mid \bar{E}]$。

贝叶斯定理的应用（续）

- $\Pr[D] = \dfrac{1}{100000} = 0.00001,\ \ \Pr[\bar{D}] = 1 - \Pr[D] = 0.99999$
- $\Pr[E \mid D] = 0.99,\ \ \Pr[\bar{E} \mid D] = 0.01,$
  $\Pr[E \mid \bar{D}] = 0.005,\ \Pr[\bar{E} \mid \bar{D}] = 0.995$

$$
\begin{aligned}
\Pr[\,D \mid E\,] &= \frac{\Pr[E|D]\Pr[D]}{\Pr[E|D]\Pr[D]+\Pr[E|\bar{D}]\Pr[\bar{D}]} \\
&= \frac{0.99\times 0.00001}{0.99\times 0.00001+0.005\times 0.99999} \\
&\approx 0.002
\end{aligned}
$$

为何结果如此小？

呈阳性，也不必太担心！

贝叶斯定理的应用（续）

$$\Pr[\,\bar{D}\mid\bar{E}\,]=\frac{\Pr[\bar{E}|\bar{D}]\,\Pr[\bar{D}]}{\Pr[\bar{E}|\bar{D}]\,\Pr[\bar{D}]+\Pr[\bar{E}|D]\Pr[D]}$$

$$=\frac{0.995\times0.99999}{0.995\times0.99999+0.01\times0.00001}$$

$$\approx 0.9999999$$

$$\Pr[\,D\mid\bar{E}\,]=1-\Pr[\,\bar{D}\mid\bar{E}\,]=0.0000001$$

呈阴性，高枕无忧！

事件独立性

- 定义：事件 E 和 F 是相互独立的当且仅当

$$\Pr[E \cap F] = \Pr[E] \cdot \Pr[F]$$

例：一个有两个孩子的家庭有四种情形 (BB, GG, BG,GB), 假设是等可能的。事件 E是两个孩子的家庭有两个男孩, 事件 F 是两个孩子的家庭至少有一个男 孩。事件 E 和 F 是否独立?

解：$\Pr[E] = \frac{1}{4}, \Pr[F] = \frac{3}{4}, \Pr[E \cap F] = \frac{1}{4}$

$$\Pr[E] \cdot \Pr[F] = \frac{3}{16} \neq \frac{1}{4} = \Pr[E \cap F]$$

故E 和 F 不是相互独立的。

随机变量、期望、方差

随机变量

- 一个**随机变量** X 是一个定义域为某样本空间 $\mathcal{S}$ 的函数 。
  - 其伴域(codomain)可为任意非空集合，但通常取实数集$\mathbb{R}$。即：$X: \mathcal{S} \to \mathbb{R}$
  - 一个随机变量是一个函数。它既不是一个变量, 也不是随机的。

随机变量（续）

- 举例: 假设一个硬币被掷 3 次. 令 $X(t)$ 是头像在结果 t 中出现的次数。那么随机变量 $X(t)$ 取值如下:
  - $X(HHH) = 3, \ \ X(TTT) = 0,$
  - $X(HHT) = X(HTH) = X(THH) = \ 2,$
  - $X(TTH) = X(THT) = X(HTT) \ = \ 1.$

  - 8 种结果的每一个出现的概率为1/8. 因此, $X(t)$的(概率)分布

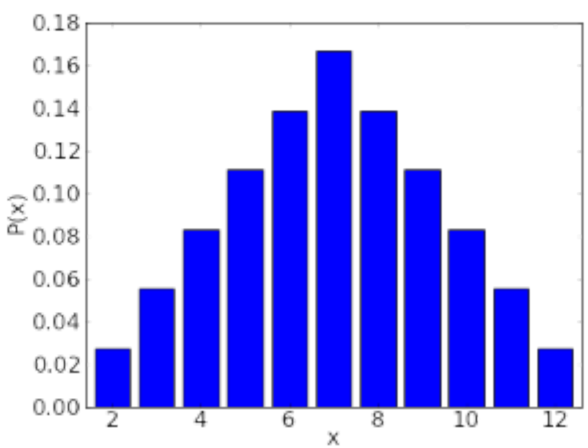
$$\Pr[X = 3] = \frac{1}{8}, \ \ \Pr[X \ = \ 2] = \frac{3}{8}, \ \Pr[X \ = \ 1] = \frac{3}{8}$$

随机变量的分布

- 定义: X 是样本空间 $\mathcal{S}$ 上的随机变量, X的分布是形如
$$(r, \Pr[X = r])$$
的二元组集合, 其中 $r \in X(\mathcal{S})$, $\Pr[X = r]$是 X 取值为 r 的概率。

随机变量分布特征的刻画

- 如何刻画随机变量取值分布的**整体特征**？
  - “平均”取值？
    - 当以概率加权之
  - “离散”程度？
    - 当以平均取值为基准，考虑偏差程度



期望值

- 定义：对于定义在样本空间 $\mathcal{S}$ 上的一个随机变量 X, 其**期望值**为
  - 以概率加权的随机变量平均取值

$$\mathrm{Ex}[X] ::= \sum_{\omega \in \mathcal{S}} X(\omega)\Pr[\omega]$$

$X(\omega) - \mathrm{Ex}[X]$ 称为 X 在 ω 处的**偏差**(deviation)

期望值(Expected value)常简称为期望(expectation), 也称为均值(mean).

期望值的直接计算

- 例：求扔一个骰子所得点数的期望值。

$$\mathrm{Ex}[X]=\frac{1}{6}\cdot 1+\frac{1}{6}\cdot 2+\frac{1}{6}\cdot 3+\frac{1}{6}\cdot 4+\frac{1}{6}\cdot 5+\frac{1}{6}\cdot 6=\frac{21}{6}=\frac{7}{2}$$

- 例：扔三个硬币，求头面朝上硬币个数的 期望值。

$$\begin{aligned}\mathrm{Ex}[X]=\frac{1}{8}[&X(HHH)+X(HHT)+X(HTH)+X(HTT)+\\&X(THH)+X(THT)+X(TTH)+X(TTT)]\\=\frac{1}{8}&(3+2+2+2+1+1+1+0)=\frac{3}{2}\end{aligned}$$

期望值的直接计算

- 例：求扔一个骰子所得点数的**倒数**的期望值。

$$\mathrm{Ex}\left[\frac{1}{X}\right]=\frac{1}{6}\cdot 1+\frac{1}{6}\cdot\frac{1}{2}+\frac{1}{6}\cdot\frac{1}{3}+\frac{1}{6}\cdot\frac{1}{4}+\frac{1}{6}\cdot\frac{1}{5}+\frac{1}{6}\cdot\frac{1}{6}=\frac{49}{120}$$

$$\frac{1}{\mathrm{Ex}[X]}\neq\mathrm{Ex}\left[\frac{1}{X}\right]$$

例：求扔两个骰子所得点数之和的期望值

$$\Pr[X=2]=\Pr[X=12]=\frac{1}{36}$$
$$\Pr[X=3]=\Pr[X=11]=\frac{1}{18}$$
$$\Pr[X=4]=\Pr[X=10]=\frac{1}{12}$$
$$\Pr[X=5]=\Pr[X=9]=\frac{1}{9}$$
$$\Pr[X=6]=\Pr[X=8]=\frac{5}{36}$$
$$\Pr[X=7]=\frac{1}{6}$$

$$\begin{aligned}\mathrm{Ex}[X] &= 2\cdot\frac{1}{36}+3\cdot\frac{1}{18}+4\cdot\frac{1}{12}+5\cdot\frac{1}{9}+6\cdot\frac{5}{36}+7\cdot\frac{1}{6}\\ &\quad+8\cdot\frac{5}{36}+9\cdot\frac{1}{9}+10\cdot\frac{1}{12}+11\cdot\frac{1}{18}+12\cdot\frac{1}{36}\\ &= 7.\end{aligned}$$

期望值的等价定义

- 定理：对于任意随机变量R

$$\mathrm{Ex}[R] = \sum_{x \in \mathrm{range}(R)} x \cdot \Pr[R = x]$$

Proof. Suppose R is defined on a sample space $\mathcal{S}$. Then,

$$\begin{aligned}
\mathrm{Ex}[R] &::= \sum_{\omega \in \mathcal{S}} R(\omega) \Pr[\omega] \\
&= \sum_{x \in \mathrm{range}(R)} \sum_{\omega \in [R=x]} R(\omega) \Pr[\omega] \\
&= \sum_{x \in \mathrm{range}(R)} \sum_{\omega \in [R=x]} x \Pr[\omega] && \text{(def of the event } [R = x]) \\
&= \sum_{x \in \mathrm{range}(R)} x \left(\sum_{\omega \in [R=x]} \Pr[\omega] \right) && \text{(factoring } x \text{ from the inner sum)} \\
&= \sum_{x \in \mathrm{range}(R)} x \cdot \Pr[R = x]. && \text{(def of } \Pr[R = x])
\end{aligned}$$

条件期望

- 给定一个随机变量R，R在已知事件A条件下的期望值是R在A中结果上的取值的概率加权平均值：

$$\mathrm{Ex}[R \mid A] = \sum_{r \in \mathrm{range}(R)} r \cdot \Pr[R = r \mid A]$$

例：已知一个公平骰子投出的点数不小于4点，此条件下投出的点数的期望值是多少？

$$\mathrm{Ex}[R \mid R \geq 4] = \sum_{i=1}^{6} i \cdot \Pr[R = i \mid R \geq 4]$$

Law of Total Expectation

定理: (全期望公式) 令R为样本空间$\mathcal{S}$上的随机变量, 且$\mathcal{S}$可以不重复不遗漏地划分为$A_1, A_2, \cdots$, 则

$$\mathrm{Ex}[R] = \sum_i \mathrm{Ex}[R \mid A_i] \Pr[A_i].$$

证明:

$$\begin{aligned}
\mathrm{Ex}[R] &= \sum_{r \in \mathrm{range}(R)} r \cdot \Pr[R = r] \\
&= \sum_r r \cdot \sum_i \Pr\left[R = r \mid A_i\right] \Pr[A_i] \\
&= \sum_r \sum_i r \cdot \Pr\left[R = r \mid A_i\right] \Pr[A_i] \\
&= \sum_i \sum_r r \cdot \Pr\left[R = r \mid A_i\right] \Pr[A_i] \\
&= \sum_i \Pr[A_i] \sum_r r \cdot \Pr\left[R = r \mid A_i\right] \\
&= \sum_i \Pr[A_i] \, \mathrm{Ex}[R \mid A_i].
\end{aligned}$$

Mean Time to Failure

- 假设某计算机系统在**每个小时**的使用过程中崩溃的概率为p, 如果现在它还没奔溃, 那么我们**期望**它在崩溃之前能工作多久?
- $\mathrm{Ex}[C]$: C是该系统在崩溃之前能工作的小时数. 令
  A: 在头一个小时系统崩溃的事件;
  $\bar{A}$: 上述事件的补事件, 则
  $$\mathrm{Ex}[C] = \mathrm{Ex}[C|A]\Pr[A] + \mathrm{Ex}[C|\bar{A}]\Pr[\bar{A}]$$
  $$\mathrm{Ex}[C|A] = 1$$
  $$\mathrm{Ex}[C|\bar{A}] = 1 + \mathrm{Ex}[C]$$
- $\mathrm{Ex}[C] = \frac{1}{p}$

期望的线性特性

- 定理：对于样本空间 $\mathcal{S}$ 上的一组任意的随机变量 X_i，（$i = 1,2,\ldots,n$）和任意实数 a, b，有
  - $\mathrm{Ex}[X_1 + X_2 + \cdots + X_n] = \mathrm{Ex}[X_1] + \mathrm{Ex}[X_2] + \cdots + \mathrm{Ex}[X_n]$
  - $\mathrm{Ex}[aX + b] = a\mathrm{Ex}[X] + b$

- 由上述定理可知，扔两个骰子所得点数之和的期望值等于第一个骰子点数期望值与第二个骰子点数期望值之和，即 7/2 + 7/2 = 7 .

例：寄存帽子问题

- 负责寄存帽子的服务生把帽子搞乱了，只能随机发还。问他可以期望还对几个？
  - 令$X_i = 1$ 若第 i 个客人拿到他的帽子；否则 $= 1$。

$$X = X_1 + X_2 + \cdots + X_n$$

$$\mathrm{Ex}[X_i] = 1 \cdot \Pr[X_i = 1] + 0 \cdot \Pr[X_i = 0] = \frac{1}{n}$$

$$\mathrm{Ex}[X] = \mathrm{Ex}[X_1] + \mathrm{Ex}[X_2] + \cdots + \mathrm{Ex}[X_n] = n \cdot \frac{1}{\mathrm{n}} = 1$$

用期望来计算平均计算复杂度

- 例：线性搜索算法的平均复杂度
  - 给定一个实数 x 及一个长度为 n 的列表，线性搜索算法依次检查列表中的元素，看是否为 x，直至找到或者所有元素都查完了而未找到 x。

ALGORITHM 2 The Linear Search Algorithm.

procedure *linear search*(x: integer, $a_1, a_2, \ldots, a_n$: distinct integers)
$i := 1$
while ($i \le n$ and $x \ne a_i$)
　　$i := i + 1$
if $i \le n$ **then** *location* $:= i$
else *location* $:= 0$
return *location*{*location* is the subscript of the term that equals x, or is 0 if x is not found}

只考虑比较操作的次数。
若 x 处于第 i 个位置，则需要 $2i$+1 次比较；若 x 不在列表中，则需要 $2n$+2 次。

用期望来计算平均计算复杂度

- 例：线性搜索算法的平均复杂度
  - 给定一个实数 x 及一个长度为 n 的列表，线性搜索算法依次检查列表中的元素，看是否为 x，直至找到或者所有元素都查完了而未找到 x。

$$
\begin{aligned}
E &= \frac{3p}{n} + \frac{5p}{n} + \cdots + \frac{(2n+1)p}{n} + (2n+2)q \\
&= \frac{p}{n}(3 + 5 + \cdots + (2n+1)) + (2n+2)q \\
&= \frac{p}{n}((n+1)^2 - 1) + (2n+2)q \\
&= p(n+2) + (2n+2)q.
\end{aligned}
$$

p 为 x 在列表中的概率
$q = 1\text{-}p$

独立随机变量

- 样本空间 $\mathcal{S}$ 上的随机变量 X 和 Y 若满足 $\Pr[X = r_1 且 Y = r_2] = \Pr[X = r_1] \cdot \Pr[Y = r_2]$，则称它们相互**独立**。
  - 例：扔两个骰子，第一个骰子点数与第二个骰子点数二者是否独立？
  - 例：扔两个骰子，第一个骰子点数与两个骰子点数之和二者是否独立？

- 对于样本空间 $\mathcal{S}$ 上**独立的**随机变量 X 和 Y 有 $\mathrm{Ex}[XY] = \mathrm{Ex}[X]\mathrm{Ex}[Y]$

方差

- 样本空间 $\mathcal{S}$ 上的随机变量X的**方差(variance)**

$$\mathrm{Var}[X] ::= \mathrm{Ex}[(X - \mathrm{Ex}[X])^2]$$

$$\mathrm{Var}[X] = \mathrm{Ex}[(X - \mathrm{Ex}[X])^2] = \sum_{\omega \in \mathcal{S}} (X(\omega) - \mathrm{Ex}[X])^2 \Pr[\omega]$$

- 方差是随机变量X在ω处的偏差的平方的加权平均
- $\sqrt{\mathrm{Var}[X]}$ 称为X的标准差 (standard deviation)
  记为 σ_X (或者$\sigma(X)$)

方差

- 定理：样本空间$\mathcal{S}$ 上的随机变量X的方差
$$\mathrm{Var}[X] = Ex[X^2] - \mathrm{Ex}^2[X]$$

- 例：扔一个骰子所得点数的方差

$$\mathrm{Var}[X] = Ex[X^2] - \mathrm{Ex}^2[X]$$
$$= \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) - \left(\frac{7}{2}\right)^2$$
$$= \frac{35}{12}$$

Bienaymé's formula

- 对于样本空间 $\mathcal{S}$ 上独立的随机变量 X 和 Y 有

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

并可推广至 n 个两两相互独立的随机变量

$$\mathrm{Var}[X_1+X_2+\cdots+X_n]$$
$$=\mathrm{Var}[X_1]+\mathrm{Var}[X_2]+\cdots+\mathrm{Var}[X_n]$$

Bienaymé's formula

- 例：求扔两个骰子点数之和的方差
  - 第一个骰子点数与第二个骰子点数两个随机变量相互独立；故可使用Bienaymé公式

$$\mathrm{Var}[X_1 + X_2] = \mathrm{Var}[X_1] + \mathrm{Var}[X_2]$$

$$= \frac{35}{12} + \frac{35}{12}$$

小结

- 基于古典概率的分析
- 离散概率空间的定义
- 条件概率与贝叶斯定理
- 随机变量, 期望, 与方差