# 离散概率

南京大学离散数学

# 提要

- 直觉概率分析：三门问题
- 直觉的形式化：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

# 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。

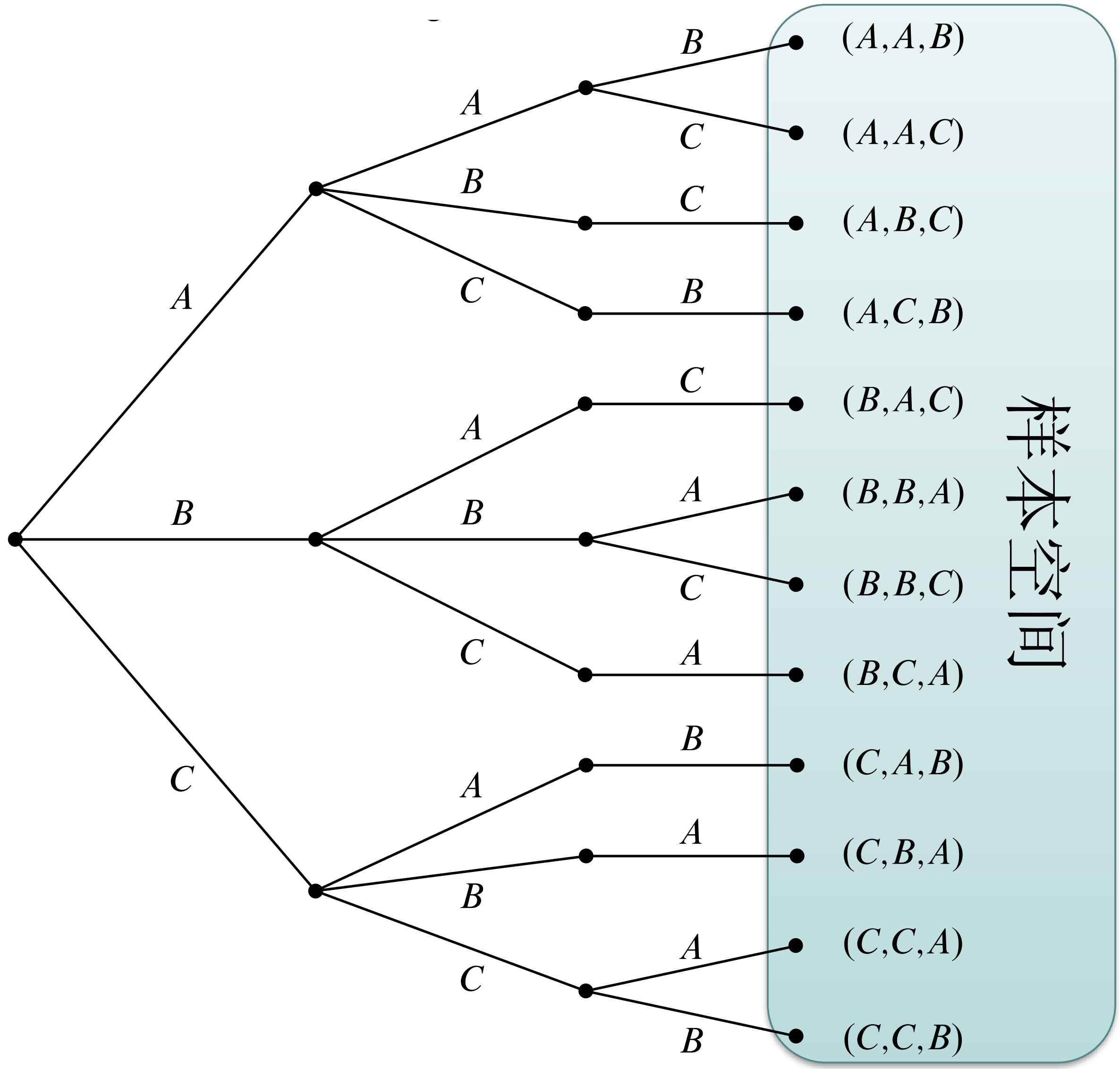
# 直觉的概率分析

- 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



# 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:

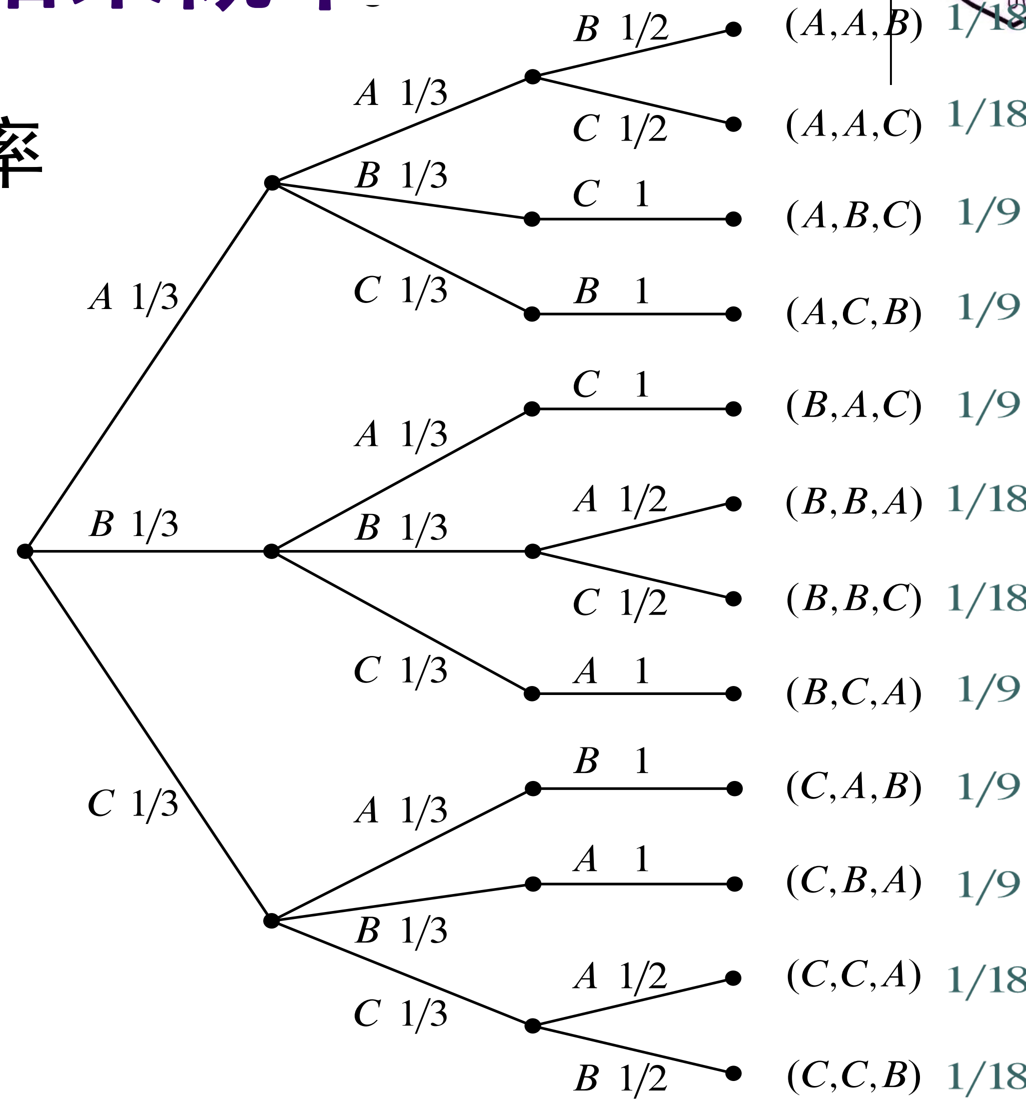
      {(A, A, B), (A, A, C), (B, B, A), (B, B, C), (C, C, A), (C, C, B)}
    - 改变选择才赢的情况：

      {(A, B, C), (A, C, B), (B, A, C), (B, C, A), (C, A, B), (C, B, A)}

6对6，似乎换不换都一样？ ✗

# 第三步：确定结果概率

- 给每个边确定概率




- 计算各结果概率

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

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

# 第四步：计算事件概率

Pr[改变选择而赢]

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

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

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

# 概率空间：基于集合论给概率以数学定义

- 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 $\omega$ 称为一个**结果**。

- 定义：满足下列条件的函数 $\Pr: \mathcal{S} \rightarrow \mathbb{R}$ 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$\mathcal{S}$是一个含 $n$ 个元素的样本空间. 均匀分布 (***uniform distribution***) 赋给 $\mathcal{S}$ 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* ⋂ *F* = {BB}.
  - $\Pr[F] = \frac{3}{4}, \Pr[E \cap F] = \frac{1}{4}$
  - $\Pr[E \mid F] = \frac{\Pr[E \cap F]}{\Pr[F]} = \frac{1}{3}$

# 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

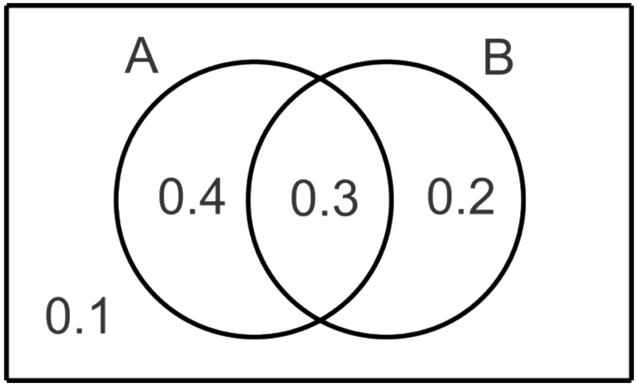
$$\Pr[\, F \mid E \,] \Pr[E] = \Pr[F \cap E]$$
$$= \Pr[E \cap \mathrm{F}] = \Pr[\, E \mid F \,] \Pr[F]$$

- 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 $\Pr[D \mid E]$, $\Pr[\overline{D} \mid \overline{E}]$ 。

# 贝叶斯定理的应用（续）

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

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

$$= \frac{0.99\times 0.00001}{0.99\times 0.00001+0.005\times 0.99999}$$

$$\approx 0.002$$

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）

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

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

$$\approx 0.9999999$$

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

呈阴性，高枕无忧！

# 举例

- 朋友来看我，乘坐交通工具的概率和这些工具可能晚点的概率分别是
  - 乘坐概率：自驾(0.3), 公交 (0.1), 高铁(0.4), 飞机(0.2)
  - 晚点概率：自驾(0.3), 公交 (0.15), 高铁(0.05), 飞机(0.5)
  - 朋友迟到了, 何种原因最有可能导致这种现象？
- 解：A自驾，B公交，C高铁，D飞机，E迟到。
  - $p(A)=0.3$, $p(B)=0.1$, $p(C)=0.4$, $p(D)=0.2$;
  - $p(E|A)=0.3$, $p(E|B)=0.15$, $p(E|C)=0.05$, $p(E|D)=0.5$;
  - 求$p(A|E)$, $p(B|E)$, $p(C|E)$, $p(D|E)$中的最大者。

# 在众多线索中探究

$$p(A|E)=\frac{p(E|A)p(A)}{p(E|A)p(A)+p(E|B)p(B)+p(E|C)p(C)+p(E|D)p(D)}$$

$p$(A|E)=90/225=2/5, $p$(B|E)=15/225=1/15, $p$(C|E)=20/225=4/45,

$p$(D|E)=100/225=4/9

误事的很可能是飞机！

# 贝叶斯Spam过滤器

- 如何确定一个电子邮件是Spam?
  - 假设我们有一个垃圾邮件的集合*B* 和一个不是垃圾的邮件集合*G*。利用贝叶斯定理来预测一个新的电子邮件是Spam 的概率。
- 考察一个特定的单词*w*, 它在*B* 和*G*中出现的次数分别为$n_B(w)$和$n_G(w)$.
- 设*S*是邮件为Spam的事件, *E*是邮件内容含单词w的事件. 需要计算 $p(S|E)$,

需要估算$p(E|S)$ 和 $p(E|\bar{S})$

# 贝叶斯Spam过滤器

估算

$p(E \mid S) = \mathbf{p(w) = n_B(w)/|B|}$

$p(E \mid \overline{S}) = \mathbf{q(w) = n_G(w)/|G|}$

$$p(S|E) = \frac{p(E|S)p(S)}{p(E|S)p(S) + p(E|\overline{S})p(\overline{S})}$$

假设 $\boldsymbol{p(S)} = ½$

垃圾邮件的频率

$$p(S|E) = \frac{p(E|S)}{p(E|S) + p(E|\overline{S})}$$

$$r(w) = \frac{p(w)}{p(w) + q(w)}$$

**若大于某个经验值，则被认为是Spam**

# 贝叶斯Spam过滤器

**举例: “Rolex” 在 2000 封垃圾邮件的250个当中出现，而在1000封非垃圾邮件中只有5封包含这个单词。估计一条含有“Rolex”的消息是Spam的概率. 假设收到的消息是Spam和不是Spam是等可能的。假设把一条消息作为Spam而拒绝的阈值为0.9，那么我们应该拒绝这条消息吗？**

**解: $p(Rolex)$ = 250/2000 =0.125, $q(Rolex)$ = 5/1000 = 0.005.**

$$r(Rolex) = \frac{p(Rolex)}{p(Rolex)+q(Rolex)} = \frac{0.125}{0.125+.005} = \frac{0.125}{0.125+.005} \approx 0.962$$

**将含有“Rolex”的消息分类为Spam，并拒绝这种消息。**

# 课堂练习

例1（男女比例谜题）：在某个国家，人们只想要男孩，每个家庭都会一直要孩子直到得到一个男孩。如果生的是女孩，他们就会再生一个；如果生的是男孩，就不再生了。假设生男孩和生女孩的概率相等，那么经过足够长时间，这个国家的男女比例将会如何？

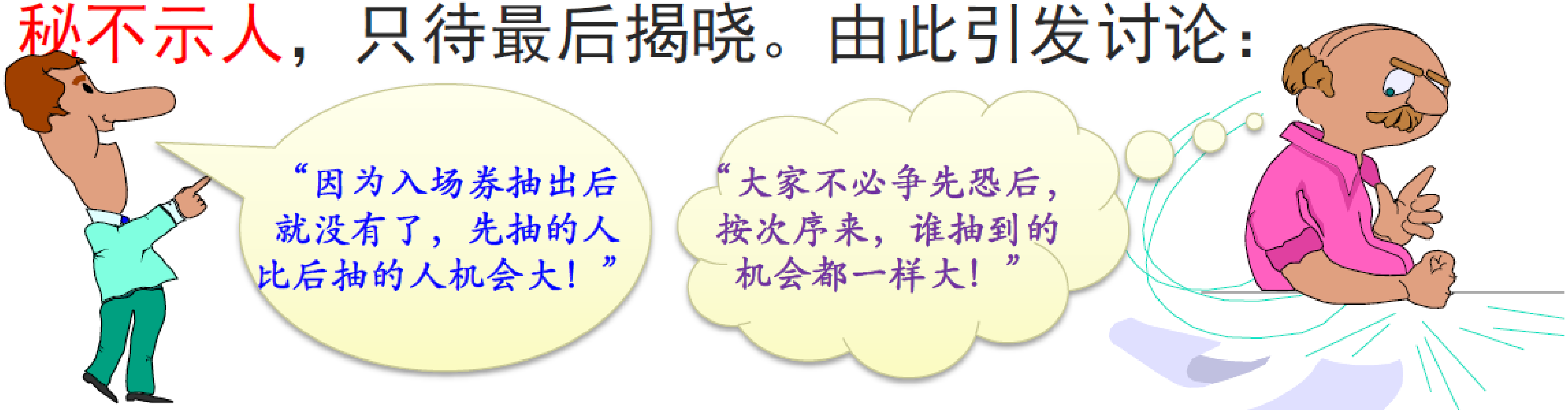
# 课堂练习

- 例2：设某种动物由出生算起活到20年以上的概率为0.8，活到25年以上的概率为0.4；问现年20岁的这种动物，它能活到25岁以上的概率是多少？
    - 解：设事件$B=\{$可以活到20年以上$\}$；事件$A=\{$可以活到25年以上$\}$；由题设，$P(A)=0.4$，$P(B)=0.8$，$A\subset B$且事件$B$已经发生，故根据条件概率公式，

      有：$P(A|B)=\dfrac{P(AB)}{P(B)}=\dfrac{P(A)}{P(B)}=0.5$

# 课堂练习

- 例3（抽签问题）：一场精彩的足球赛将要举行，5个球迷好容易才弄到一张入场券。大家都想去，只好用抽签的方法来解决，5张同样的卡片，只有一张上写有“入场券”，其余的什么也没写。将它们放在一起洗匀，让5个人依次抽取。每人抽签之后秘不示人，只待最后揭晓。由此引发讨论：



# 课堂练习

**解**（**抽签问题概率分析**）：用$A_i$表示事件“第$i$个人抽到入场券（$i=1,2,\cdots,5$）”

- **第一人抽签时：** 显然有：$P(A_1)=1/5$，$P(\overline{A_1})=4/5$；
- **第二人抽签时：** 由于第1人抽中与否的信息未向外泄露，若要抽中，必然蕴含条件“第1人未抽中”，故由乘法原理，事件$A_2=\overline{A_1}A_2$，故由条件概率得：$P(A_2)=P(\overline{A_1})P(A_2|\overline{A_1})=\frac{4}{5}\cdot\frac{1}{4}=1/5$；
- **第三人抽签时：** 第1、2人信息均未知，故若要抽中，必然蕴含前两任均未抽中，即$A_3=\overline{A_1A_2}A_3$，由条件概率公式：$P(A_3)=P(\overline{A_1})P(\overline{A_2}|\overline{A_1})P(A_3|\overline{A_1A_2})=\frac{4}{5}\cdot\frac{3}{4}\cdot\frac{1}{3}=1/5$；
- **以此类推：** 每人抽中的概率均为1/5，故不泄露信息的抽签模型中抽中的概率与顺序无关，等同于有放回的摸球模型

# 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

- 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：$X: \mathcal{S} \rightarrow \mathbb{R}$
  - 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

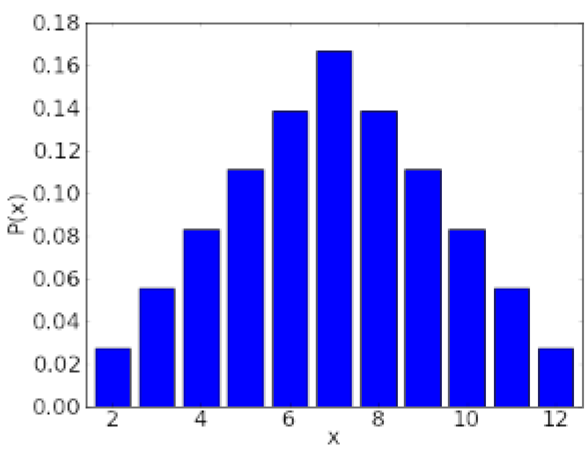
$$\Pr[X = 3] = \frac{1}{8},\ \ \Pr[X = 2] = \frac{3}{8},\ \Pr[X = 1] = \frac{3}{8}$$

# 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 $\omega$ 处的偏差(deviation)

# 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。

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

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

# 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

**Theorem 18.4.5** (Law of Total Expectation). *Let $R$ be a random variable on a sample space $S$, and suppose that $A_1, A_2, \ldots,$ is a partition of $S$. Then*

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

*Proof.*

$$\begin{aligned}
\text{Ex}[R] &= \sum_{r \in \text{range}(R)} r \cdot \Pr[R = r] && \text{(by 18.3)} \\
&= \sum_r r \cdot \sum_i \Pr\left[R = r \mid A_i\right] \Pr[A_i] && \text{(Law of Total Probability)} \\
&= \sum_r \sum_i r \cdot \Pr\left[R = r \mid A_i\right] \Pr[A_i] && \text{(distribute constant } r\text{)} \\
&= \sum_i \sum_r r \cdot \Pr\left[R = r \mid A_i\right] \Pr[A_i] && \text{(exchange order of summation)} \\
&= \sum_i \Pr[A_i] \sum_r r \cdot \Pr\left[R = r \mid A_i\right] && \text{(factor constant } \Pr[A_i]\text{)} \\
&= \sum_i \Pr[A_i] \, \text{Ex}[R \mid A_i]. && \text{(Def 18.4.4 of cond. expectation)}
\end{aligned}$$

# 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：Expected Value in the Hatcheck Problem

- 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。

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

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

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

# 独立随机变量

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

# 独立随机变量

对于样本空间 *S* 上**独立的**随机变量 *X* 和 *Y* 有

$E(XY) = E(X)E(Y)$

$$
\begin{aligned}
E(XY) &= \sum_{r \in XY(S)} r \cdot p(XY = r) \\
&= \sum_{r_1 \in X(S), r_2 \in Y(S)} r_1 r_2 \cdot p(X = r_1 \text{ and } Y = r_2) \\
&= \sum_{r_1 \in X(S)} \sum_{r_2 \in Y(S)} r_1 r_2 \cdot p(X = r_1 \text{ and } Y = r_2) \\
&= \sum_{r_1 \in X(S)} \sum_{r_2 \in Y(S)} r_1 r_2 \cdot p(X = r_1) \cdot p(Y = r_2) \\
&= \sum_{r_1 \in X(S)} \Big(r_1 \cdot p(X = r_1) \cdot \sum_{r_2 \in Y(S)} r_2 \cdot p(Y = r_2)\Big) \\
&= \sum_{r_1 \in X(S)} r_1 \cdot p(X = r_1) \cdot E(Y) \\
&= E(Y)\Big(\sum_{r_1 \in X(S)} r_1 \cdot p(X = r_1)\Big) \\
&= E(Y)E(X)
\end{aligned}
$$

# 方差

- 样本空间 *S* 上的随机变量 *X* 的**方差**(variance), 记作 *V(X)*，定义为

$$V(X)=\sum_{s\in S}(X(s)-E(X))^2p(s)$$

  - 方差是变量 *X* 在 *s* 处的偏差的平方的加权平均。
  - $\sqrt{V(X)}$ 称为 *X* 的**标准差**(standard deviation)，记为 $\sigma(X)$ .

# 方差

- 定理：样本空间 $S$ 上的随机变量 $X$ 的方差

$$V(X) = E(X^2) - E(X)^2$$

- 证明：

$$\begin{aligned} V(X) &= \sum_{s\in S}(X(s) - E(X))^2 p(s) \\ &= \sum_{s\in S} X(s)^2 p(s) - 2E(X)\sum_{s\in S} X(s)p(s) + E(X)^2 \sum_{s\in S} p(s) \\ &= E(X^2) - 2E(X)E(X) + E(X)^2 \\ &= E(X^2) - E(X)^2. \end{aligned}$$

# 方差

- 例：扔一个骰子点数的方差

$$V(X) = E(X^2) - E(X)^2$$
$$= \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) - \left(\frac{7}{2}\right)^2$$
$$= \frac{35}{12}$$

# Bienaymé’s formula

- 对于样本空间 $S$ 上**独立的**随机变量 $X$ 和 $Y$ 有

$$V(X+Y)=V(X)+V(Y)$$

并可推广至n个两两相互独立的随机变量

$$V(X_1+X_2+\cdots+X_n)$$
$$=V(X_1)+V(X_2)+\cdots+V(X_n)$$

# 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。

$$V(X)=V(X_1+X_2)=V(X_1)+V(X_2)=\frac{35}{12}+\frac{35}{12}$$

# 切比雪夫不等式 Chebyshev's Inequality

- 对于样本空间 $S$ 上随机变量 $X$,和任意的正实数 $r$ 有

$$p(|X(s) - E(X)| \geq r) \leq \frac{V(X)}{r^2}$$

# 利用切比雪夫不等式进行概率估算

- 例：已知正常男性成人每毫升血液中白细胞数目的平均值是7300，标准差是700。试利用切比雪夫不等式估算每毫升血液含白细胞数在5200-9400之间的概率。
- 解：设X表示每毫升血液中白细胞个数，则

$$E(X) = 7300, \qquad V(X) = \sigma^2 = 700^2$$

而 $p(5200 \le X \le 9400) = p(|X - 7300| \le 2100)$

$$= 1 - p(|X - 7300| \ge 2100)$$

又 $p(|X - 7300| \ge 2100) \le \frac{700^2}{2100^2} = \frac{1}{9}$

故 $p(5200 \le X \le 9400) \ge \frac{8}{9}$

# 作业

- 见课程QQ群