概率统计B

Probability and Statistics

张思容
zhangsirong@buaa.edu.cn

数学与系统科学学院，北京航空航天大学
School of Mathematics and System Sciences, BUAA

September 15, 2014

1. 课程介绍
   - 介绍
2. 概率模型
   - 样本空间
   - 概率律
   - 从有限模型到公理化模型***
3. 古典概率和计数方法
   - 样本空间与集合论
   - 计数与组合

相互介绍

教师：

- 张思容: Ph.D. 几何分析，医学图像分析；
- 联系方式：134-3920-1025. zhangsirong@buaa.edu.cn
- 办公时间：欢迎dropby，地点：学院路校区图书馆西配楼501。
- 概率统计课程答疑： 星期一课上预约, 地点待定(教师休息室?) 中午12-1pm 或15：30pm-16:30pm
- 课程网站：个人主页 ppt和解答 http://smss.buaa.edu.cn/szdw/jxls/zsr/index.htm
- 欢迎大家学期中提建议和问题，不要考试后!

学生介绍：1313：（11，12，13，14，41，61，62） 公共邮箱 buaajtxy2013@163.com
联系电话?

课程介绍: 参考书目

预备要求：微积分.

参考书目:

- (教材) 概率统计及随机过程。张福渊等，北家航空航天大学出版社。ISBN: 7810770047
- (推荐) 概率导论. (MIT教材) 人民邮电出版社 ISBN 9787115215444。
  MIT 公开课程 6.041，Probabilistic Systems Analysis and Applied Probability。
- （参考) 概率论基础教程，S.M.ROSS，人民邮电出版社，ISBN 978-7-115-22110-0.
- (考研) 概率论与数理统计。盛骤等 (浙江大学教材). 高等教育出版社。 ISBN 9787040238969
- (习题）北航概率统计习题集；浙江大学概率统计习题解答；

课程介绍：内容学习

概率统计：最有用的数学课程，没有之一。

- 例子: 赌博, 股票， 市场调查。
- 我们生活在随机世界里: 随机现象建模+统计决策。
- 困难：理论（数学分析，实变函数，测度论...), 应用 （统计思想，数据挖掘,AI...）计算：（大数据，软件SAS,...)

课程学习

- 目标：学会概率建模的思想和方法；掌握经典例子，练习技巧(解题考试).
- 作业：每周上课交一次。下周一返回。
- 成绩： 平时成绩 10+期末考试 90=100
  小测验若干(quizs).

问题? Q&A(Questions and Answers)

概率与真实世界

什么是概率？ 《概率沉思录》 E.T.Jaynes

- 赌博与彩票等：概率是等可能的 → 古典概率模型 Laplace
- 红楼梦是否是曹雪芹写的？ 明天会下雨吗？ 概率是主观判断(经验) → 贝叶斯学派 Bayes
- 人口出生率： 男孩 vs 女孩 105 : 100 概率是数据发生的频率 → 频率学派 Fisher

概率模型：在不同的应用中选择不同的模型。
"抓住老鼠就是好猫!"

Remark (数学建模)

确定现象：数 → 函数
随机现象：集合 → 随机变量

样本空间

Definition

样本空间: 表示一个试验的所有可能结果的集合。记为S或Ω。
集合+所有结果: (选择合适样本空间是*ART*!)

- 样本空间的例子：
  投硬币：(Head, Tail)
  掷骰子：(1,2,3,4,5,6)
  射击打靶：(整个平面，或者3D)
- 事件： 样本空间中关心的结果(子集合). 用A, B等表示。
  $A=\{$ 骰子是偶数 $\}$,$B=\{$打靶中9环 $\}$。
- 事件与子集；集合运算。空集即不可能事件；全集是必然事件；
  注意：不是所有子集都是事件!

概率律(law)

概率律P:给每个事件指定一个概率大小。

1. $P(A) \geq 0$
2. $P(\Omega) = 1$
3. $A \cap B = \emptyset$,
   $P(A \cup B) = P(A) + P(B)$ 可加性

EXAMPLE

投硬币： *(Head, Tail)* $A = \{Head\}, B = \{Tail\}$, 指定$P(A) = P(B) = \frac{1}{2}$.
类似对掷骰子可以指定$P(\{1\}) = \frac{1}{6} \dots$
*(*有限空间*)*古典概率模型： $P(A) = \frac{\#A}{\#\Omega}$

问题： 空间无穷怎么办？ 古怪子集怎么办？

无穷样本空间：可数**

- 古典模型: 集合有n个元素，每个元素可看成一个子集,称为基本事件(样本点);
- 可数模型(离散模型)：集合有可数个元素。每个元素可看成一个子集,可看成事件;

EXAMPLE (射击命中)

设一个人打靶命中率为p，重复射击直到打中靶停止。我们关心的要射击多少次才能打中?

解答.

样本空间：$\Omega = \{1, 2, 3, \ldots, n, \ldots, \infty\}$
概率
律：$P(x = 1) = p, P(X = 2) = p(1-p), \ldots, P(X = n) = p(1-p)^{n-1}$.
验证 $P(X = 1) + P(X = 2) + \cdots = 1$ □

一般的：可数模型的概率律可对应一个收敛的正项级数(和为1)。

无穷样本空间: 连续空间(不可数)**

连续模型：集合有不可数个元素。每个元素可看成一个子集,但一点作为事件无意义! 因为一点的概率必须为零!
古典概率推广：几何概率 定义 $P(A)=\frac{L(A)}{L(\Omega)}$,其中$L(A)$表示其区域的面积或长度或体积。

EXAMPLE (约会问题)

两个人约定在8点钟见面，都可能迟到至多一个小时。如果先到的等候15分钟，问两个人能见面的概率。

解答.

设到达时间分别为x, y, 样本空间：$\Omega=\{(x,y)|8\le x\le 9, 8\le y\le 9\}$
事件：两个人见面 $A=\{(x,y)||x-y|\le 15/60\}$
概率律：$P(A)=\frac{L(A)}{L(\Omega)}$, 计算有$P(A)=1-P(\bar{A})=1-\frac{3}{4}\cdot\frac{3}{4}=\frac{7}{16}$ □

一般的： 连续模型的概率律对应一个可积函数。

概率律的公理化**

Definition (概率: 1930 Kolmogorov)

P是定义在样本空间Ω上所有事件$\mathfrak{F}$的一个函数；$P:\mathfrak{F}\to[0,1]$, 满足

1. 非负性： $P(A)\geq 0$;
2. 规范性： $P(\Omega)=1$;
3. 可数(列)可加性： 如果A_i是互不相容事件，则$P(\bigcup_{i=1}^{\infty}A_i)=\sum_{i=1}^{\infty}P(A_i)$;

称P为Ω的一个概率(律)，记$P(A)$为事件A的概率。

基本属性

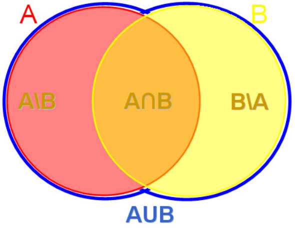
- $P(\emptyset)=0$；有限可加性；
- 单调性：$A\subset B$，则$P(A)\leq P(B)$;
- ***连续性：有单调增序列A_i,单调减序列B_i, 则 $P(\bigcup_{n=1}^{\infty}A_n)=\lim_{n\to\infty}P(A_n)$,$P(\bigcap_{n=1}^{\infty}B_n)=\lim_{n\to\infty}P(B_n)$
- ***数学上称概率为测度；见实变函数论。

思考题

将52张扑克扣在桌上一张张翻开；一直到出现一张"A"为止。再翻一张牌，问下一张牌是黑桃A的概率和方块2的概率谁大?
一样大! $p = \frac{1}{52}$.
解释参考 '概率论基础教程'（S.M.ROSS，ISBN 978-7-115-22110-0) 例子 5j.p34,p65.

事件与集合

- 集合Ω: 元素 $\omega \in \Omega$
  记子集$A = \{\omega | \omega \in A\}$.
- 集合运算: $\bar{A} = \Omega - A$, 乘法$A \cap B = AB$,"加法" $A \cup B$, 减法$A - B = A\bar{B}$
  对应的事件意义?
- 不相容事件: $A \cap B = \emptyset$.
  对立事件: $\bar{A}$
- 集合运算的规律：交换律，结合律;
  Venn韦恩图
- 有用公式:
  $A \cup B = A + \bar{A}B, A = AB + A\bar{B}$
  De-Morgan公式:
  $\overline{\bigcup_{j=1}^{j=n} A_i} = \bigcap_{j=1}^{j=n} \bar{A}_i$.



配对问题

EXAMPLE (配对问题)

任意n个同学交了n本作业，随机每人发回一本作业，试求有$k=0,1,2,...,n$个同学得到自己作业的概率？ 简单情形：$n=2$, $n=3$,

解答.

设$n=3$, 记三个人A,B,C,作业a,b,c, Ab表示A拿到b的作业。
样本空间：$\Omega=\{AaBbCc, AaBcCb, AbBaCc, AbBcCa, AcBbCa, AcBaCb\}$
事件: $X=k$个同学配对成功; $k=0$: $X=\{AbBcCa, AcBaCb\}$
$k=1$: $X=\{AaBcCb, AbBaCc, AcBbCa\}$
$k=2$: $X=\emptyset$; $k=3$: $X=\{AaBbCc\}$
古典概率计算有 $P(\{X=0\})=\frac{1}{3}, P(\{X=1\})=\frac{1}{2}$,
$P(\{X=2\})=0, P(\{X=3\})=\frac{1}{6}$, $\square$

n变大时，情况迅速变复杂! 见后续讲解。

计数方法 counting

Theorem (乘法原理)

完成一件事要k步，每一步分别有$n_1, n_2, n_3, \ldots, n_k$方法，则完成这件事的方法共有$n = n_1 \cdot n_2 \cdots n_k$.

基本结果:

- (m次试验): 从n个元素中有放回的每次取一个；取出m个元素,排成一列；共有 n^m种可能；不同排列是等可能的；
- (m元排列) 从n个元素中无放回的每次取一个；取出m个元素,排成一列；共有 $A_n^m = \frac{n!}{(n-m)!}$种可能；不同排列是等可能的；
- (m元组合) 从n个元素中无放回的每次取一个；取出m个元素,放在一组；共有 $C_n^m = \frac{n!}{m!(n-m)!}$种可能；不同组合是等可能的；

例子：(集合子集的个数) 集合有n个元素，则所有子集的个数为2^n。 利用：组合数即二项式系数, 由二项式展开$(x+y)^n = \sum C_i^n x^i y^{n-i}$可得。

例子

EXAMPLE (生日问题)

任意n个人中有（至少两个人）有相同生日的概率。

解答.

假设每个人等可能出生于365天中任一天。

n个人生日的样本空间Ω大小：365^n

$\bar{A}$ =[没有两个人生日相同], $\bar{A}$的集合大小为 $n!C_{365}^n$,

$P(A) = 1 - P(\bar{A}) = 1 - \frac{n!C_{365}^n}{365^n}$

n	20	30	40	50	60	70	80
p_n	0.411	0.706	0.891	0.970	0.994	0.999	0.9999

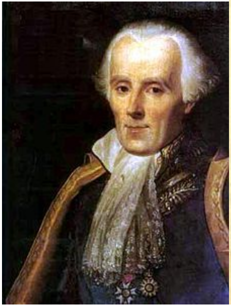
$\square$

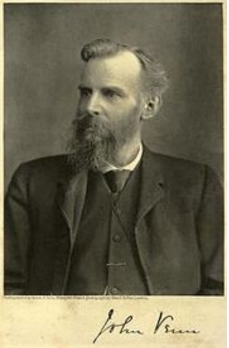
作业

北航教材:
P31 习题一.
1(2),(3),(5);
4
5
6
8

从古典到频率学派

P.Laplace VS J.Venn



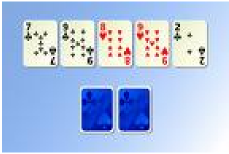


概率的真实大小?

- 投硬币：coin flipping 正面的概率?
- 拉普拉斯：1/2
- De Morgan 德摩根：1061/2048
- 蒲丰Buffon： 2048/4040
- John Kerrich: 5067/10000
- K Pearson: 12012/24000, Feller: 4979/10000;

频率学派：概率由大量数据的频率决定。

Texas Hold’em



德州扑克Texas Hold’em,是世界上最流行的扑克游戏。每个玩家最后用五张扑克牌比大小。 得克萨斯扑克的大小规则如下:

1. 皇家同花顺 Royal straight flush (比如黑桃10,J,Q,K,A)
2. 同花顺 Straight flush (比如黑桃3,4,5,6,7)
3. 四条,炸弹 Four of a kind (比如四条9和一个其它任何牌)
4. 满堂彩 Full house (三条加一对) 5.清一色 Flush (比如梅花2,5,6,8,J)
5. 一条龙 Straight (比如4,5,6,7,8不同花色混杂)
6. 三条 Three of a Kind (比如三个8和其它任意两张单牌)
7. 两对 Two Pair
8. 一对 Pair

思考：这个规则合理吗？各种情形的概率是多少？