

Probability and Statistics

概率统计

南京大学

高 尉

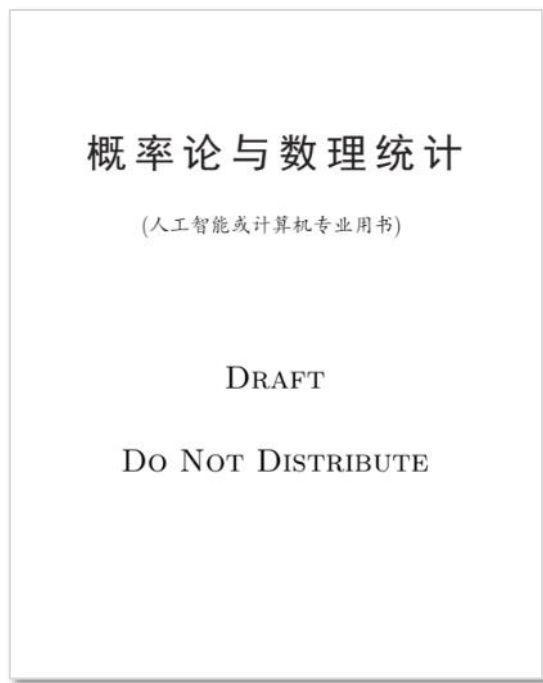


Course Information and Textbooks

- Instructor: 高尉
 - gaow@nju.edu.cn
 - gaow@lamda.nju.edu.cn
- Office: 国际学院A-305

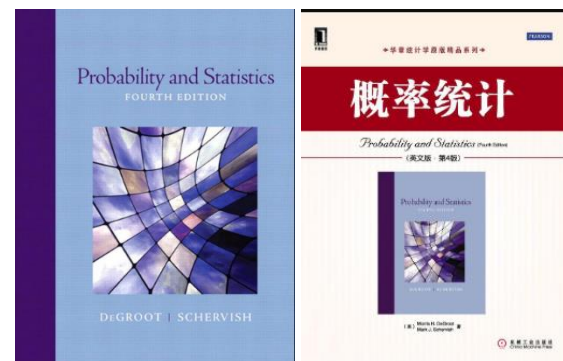
要求或建议

- 不能躺着上课
- 备好纸笔、做笔记
- 上课专心、课后少花时间
- 独立完成作业



概率论与数理统计

- 盛骤、谢式千等编
- 高等教育出版社



Probability and statistics

- M. H. DeGroot and M. J. Schervi
- 机械工业出版社

概率的萌芽：赌博（点数分配问题）

两人进行一场赌博，5局3胜，赌金为1000；假设当前比分为2 : 1，而比赛由某种原因不得不中止。

问题：最“公平合理”的奖金分配方式？

— 1495年意大利数学家/修道士帕西奥尼 (Luca Pacioli)

这个问题持续了150年左右



Luca Pacioli

概率的起源

1650左右的法国

- 赌博流行且时尚,不受法律限制
- 赌博变得更加复杂, 风险增大
- 有必要通过数学方法来计算胜率
- 德梅根(De Mere): 点数分配问题

古典概型 (等可能概型)

- 有限个基本事件
- 每个基本事件发生的可能性相同



Blaise Pascal



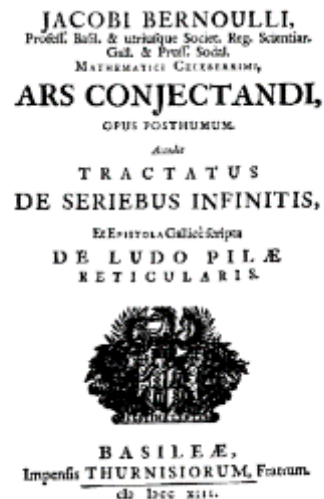
Pierre Fermat

概率论的形成和发展

《推想的艺术》 1713年

- 贝努利 (James Bernoulli)

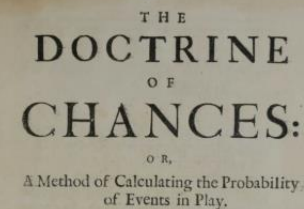
- 大数定律
- 频率稳定性理论化
- 特殊问题到一般理论



《机遇原理》 1718年

- 棣谟佛 (Abraham de Moivre)

- 概率乘法法则
- 正态分布律
- 中心极限定理的一个特例



By A. De Moivre. F. R. S.
L O N D O N:
Printed by W. Pearson, for the Author. MDCCXVIII.

概率论的形成和发展(续)

《偶然性的算术试验》 1706

-蒲丰(Comte de Buffon)

- 结合概率和几何, 开始几何概率的研究
- 采取概率的方法来求圆周率的尝试

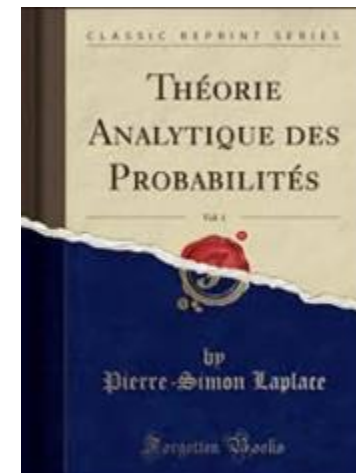


19世纪：概率论进一步发展-应用

Theorie Analytique des Probabilités

- Pierre-Simon Laplace

A mathematical theory of probability with an emphasis on scientific applications



高斯
Carl F. Gauss



麦克斯韦
James C. Maxwell



吉布斯
Josiah W. Gibbs

20世纪： 概率的公理化定义

1900年希尔伯特提出了著名的23个数学问题

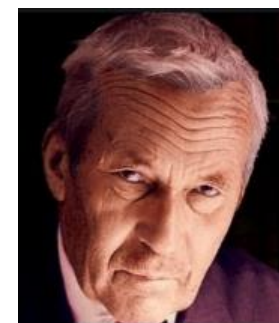
概率公理化



Grundbegriffe der Wahrscheinlichkeitsrechnung 1933

- 柯尔莫哥洛夫(Andrey Kolmogorov)

建立概率公理化理论体系，利用基本性质来定义概率，可媲美于欧几里得几何公理化



概率公理化 (三点)

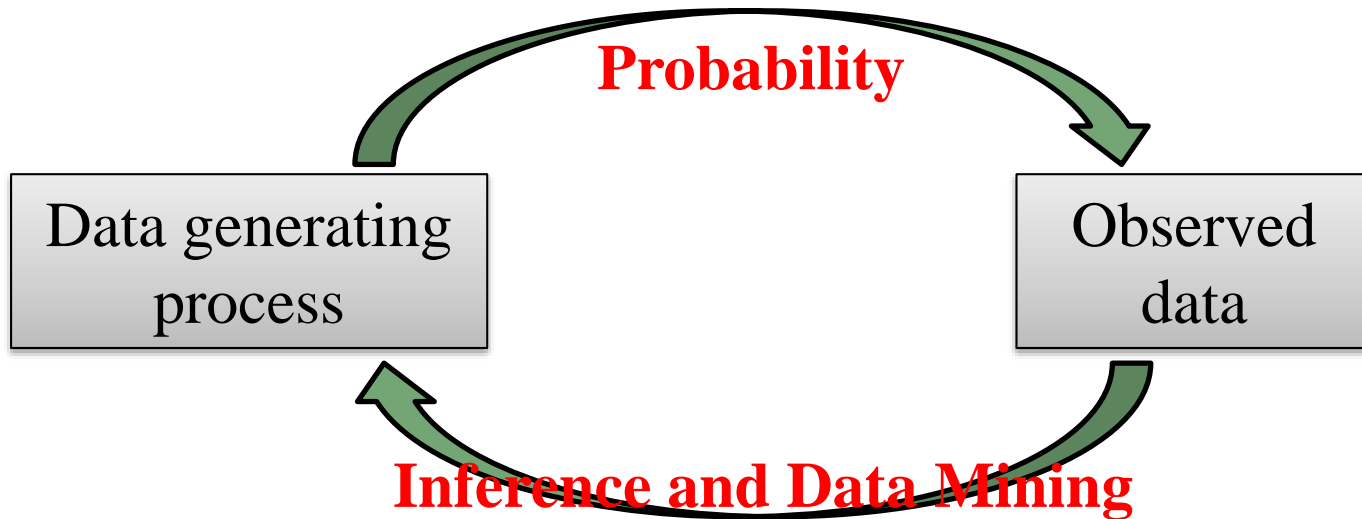
- 非负性
- 规范性
- 可列可加性

现代概率统计：测度论



概率与统计

- **概率**：研究事件的不确定性，在给定数据生成过程中观察、研究数据的性质，强调**公理体系**、**推理**
- **统计**：收集与分析数据，根据观察的数据反思其数据生成过程，强调“**归纳**”



- 概率: 随机变量、分布、大数定律、概率不等式等
- 统计: 推断、参数估计、假设检验、模型、算法 (与机器学习相关)

应用



顶级国际期刊与会议

- 概率统计及其相关的顶级国际期刊
 - Annals of Statistics
 - Journal of the American Statistical Association (JASA)
 - Annals of Probability
 - Journal of Machine Learning Research
- 国际会议
 - ICML: International Conference on Machine Learning
 - COLT: Annual Conference on Learning Theory
 - STOC: ACM Symposium on Theory of Computing
 - FOCS: IEEE Symposium on Foundations of Computer Science

Researcher work (ICML/COLT/STOC)

- DeltaGrad: Rapid retraining of machine learning models
- on the noisy gradient descent that generalizes as SGD
- club: a contrastive log-ratio upper bound of mutual information
- optimistic bounds for multi-output prediction
- on the generalization effects of linear transformations in data augmentation
- uniform convergence of rank-weighted learning
- online learning with imperfect hints
- the implicit and explicit regularization effects of dropout
- Single Point Transductive Prediction
- Sample Amplification: Increasing Dataset Size even when Learning is Impossible
- In Defense of Uniform Convergence: Generalization via Derandomization with an Application to Interpolating Predictors
- Deep k-NN for Noisy Labels
- Deep Streaming Label Learning
- Fine-Grained Analysis of Stability and Generalization for Stochastic Gradient Descent
- A Distributional Framework for Data Valuation
- Class-Weighted Classification: Trade-offs and Robust Approaches
- Extreme Multi-label Classification from Aggregated Labels
- structured prediction with partial labelling through the infimum loss
- Optimality and Approximation with Policy Gradient Methods in Markov Decision Processes
- Learning Mixtures of Linear Regressions in Subexponential Time via Fourier Moments
- Efficiently Learning Structured Distributions from Untrusted Batches

先修课程与考核方式

- 数学分析
 - 高等代数
 - 计算机编程
-
- Home work: 30% (每周1次, 下一周上课之前交)
 - Mid-Term exam: 10%
 - Final exam: 60%

Ch1-1 随机事件及其运算



必然现象与随机现象

自然界所观察到的现象：必然现象 随机现象

- 必然现象：在一定条件下必然发生的现象，其特征是条件完全决定结果

- 太阳从东边升起
- 水往低处流
- 可导的函数必连续



- 随机现象：在一定条件下可能出现、可能不出现的现象，其特征是条件不能完全决定结果

- 婴儿的诞生
- 流星殒落
- 蝴蝶效应 ...



随机现象的例子

例1：在相同条件下掷一枚均匀的硬币

结果：可能是**正面**、也可能**反面**



例2：过马路交叉口时,可能遇上的交通指挥灯

结果：**红、黄、绿**



例3：含正品和次品的产品中任取一产品

结果：**正品、次品**



随机现象的必然性与偶然性

- 随机现象揭示了条件和结果之间的非确定性联系, 其数量关系无法用函数确切的描述
- 随机现象在一次观察中出现什么结果具有偶然性, 即多种可能的结果中不能确定到底是哪一种结果
- 随机现象是否是无规律可言 **不是**

大量重复试验或观察中, 结果的出现具有一定的统计规律性

□ **偶然性**: 对随机现象做一次观察, 观察结果不可预知

□ **必然性**: 对随机现象做大量观察, 观察结果具有一定的规律性, 即**统计规律性**

概 率: 研究随机现象统计规律性的学科

概率论普遍存在于生活

法国数学家拉普拉斯：对生活的大部分，最重要的问题实际上只是概率问题.

图灵奖得主 **Y. LeCun**：历史上多数研究成果的出现是偶然事件
... 所有努力都是为了提升概率.

现实生活中的每个人：所有的努力都是为了提高成功的概率.

随机试验

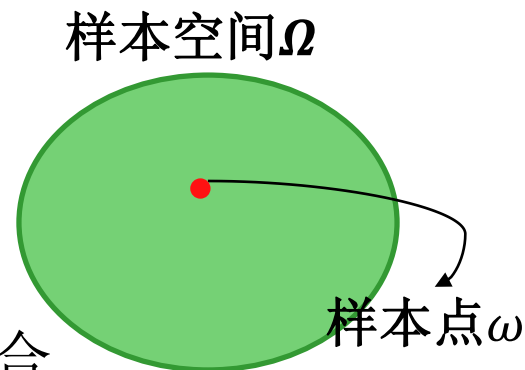
- **随机现象**：具有**不确定性**（或偶然性）的现象
- **试验**：对某随机现象的观察或测量等
- **随机试验(用E表示)**：具备以下三个特点的试验
 - **可重复**：可在相同的条件下重复进行
 - **多结果**：结果不止一个，所有可能的结果事先已知
 - **不确定**：试验前无法预测/确定哪一种结果

例如

- ✓ E_1 ：抛一枚骰子，观察其出现的点数
- ✓ E_2 ：随机选取一盏电灯，测试其寿命

样本空间

- **样本点**：试验的每一种可能的结果
 - 记为 ω
- **样本空间**：试验中所有可能的结果组成的集合
 - 记为 Ω



E1: 抛一枚骰子，观察其出现的点数

样本空间 $\Omega =$

E2: 随机选取一盏电灯，测试其寿命

样本空间 $\Omega =$

样本空间类型

有限样本空间：有限个样本点

例：将一枚硬币抛掷两次，观察正面、反面出现的情况，该试验的样本空间 Ω

无限可列样本空间：样本点是无限的但可列的

例：中国一年内出生的婴儿数，其样本空间 Ω

不可列样本空间：样本点是无限的、且不可列的

例：随机选取一盏电灯，测试其寿命，则样本空间 Ω

随机事件

随机事件：样本空间 Ω 的子集，由单个或某些样本点 ω 的集合

- ◆ 本质是集合

- ◆ 一般用字母 A 、 B 、 C 等

称“**随机事件 A 发生**”当且仅当**试验的结果是子集 A 中的元素**

对试验 E ：抛两枚骰子



其样本空间 $\Omega = \{(i, j): i, j \in [6]\}$

- 随机事件 A ：点数相同， $A = ?$
- 随机事件 B ：点数和为偶数， $B = ?$

随机事件

- **基本事件**：由单个样本点构成的单点集合

对试验E：抛一枚骰子

– $A = \{\text{掷出1点}\}$ $B = \{\text{掷出基数点}\}$

- **必然事件**：试验中必定发生的事件，记为 Ω
- **不可能事件**：试验中不可能发生的事件，用 \emptyset 表示

对试验E：抛一枚骰子

- “抛出的点数小于7”的事件是必然事件
- “抛出的点数大于7”的事件是不可能事件

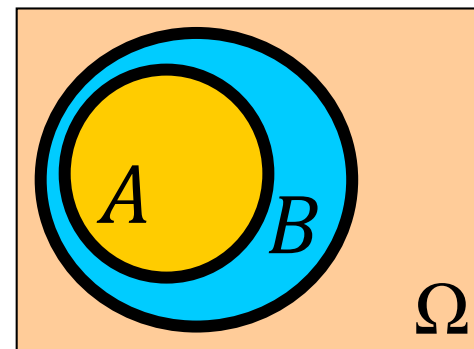
概率论与集合论之间的关系

记号	概率论	集合论
Ω	样本空间, 必然事件	全集
\emptyset	不可能事件	空集
ω	基本事件	元素
A	随机事件	子集

事件间的关系

包含： 若 A 发生必然导致 B 发生，称事件 B 包含事件 A ，记为 **$A \subset B$** 或 **$B \supset A$**

相等： 若 $A \supset B$ 且 $B \supset A$ ，记为 **$A = B$**

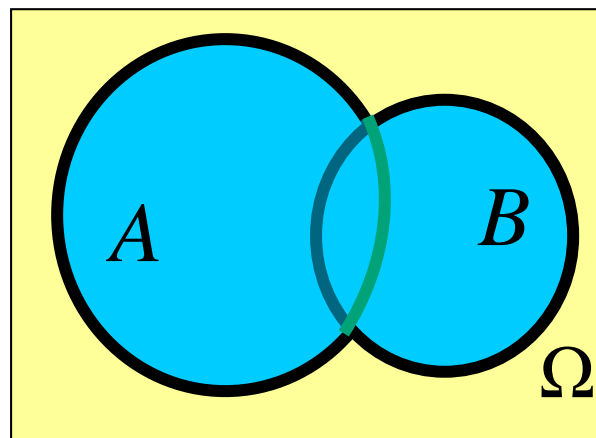


事件间的关系

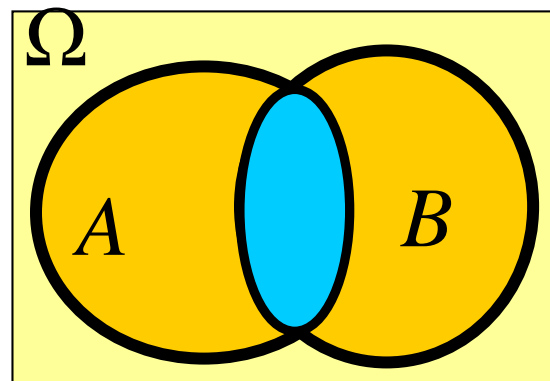
事件的并： 事件 A 和 B 至少发生一个的事件称为 A 和 B 的并/和，记为 **$A \cup B$**

事件 A_1, A_2, \dots, A_n 中至少有一个发生的事件称为 A_1, A_2, \dots, A_n 的并，记为

$$A_1 \cup A_2 \cup \dots \cup A_n = \bigcup_{i=1}^n A_i$$



事件的交： 事件 A 和 B 同时发生的事件称为 A 与 B 的交/积，记为 $A \cap B = AB$



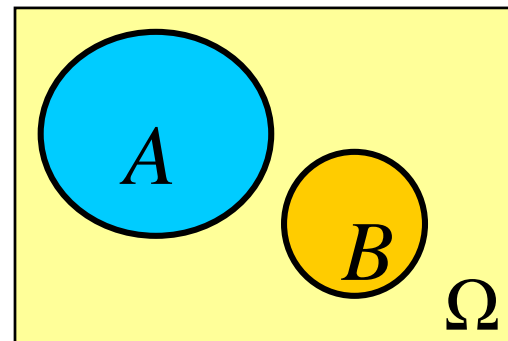
事件 A_1, A_2, \dots, A_n 同时发生的事件称为 A_1, A_2, \dots, A_n 的交，记

$$A_1 \cap A_2 \cap \dots \cap A_n = \bigcap_{i=1}^n A_i = A_1 A_2 \dots A_n$$

事件间的关系

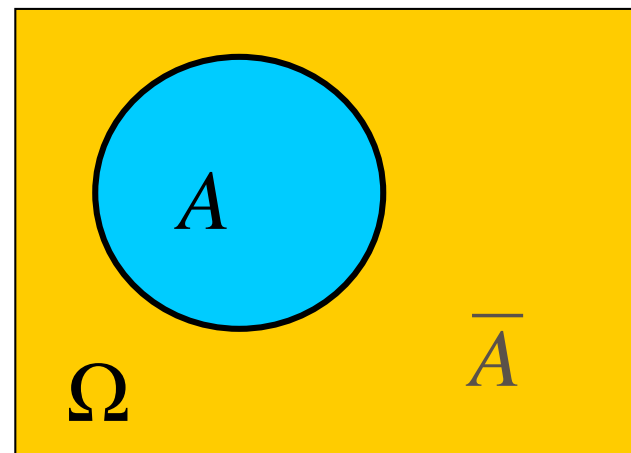
互斥/互不相容：若事件 A 和 B 不可能同时发生，称事件 A 和 B 互斥、或互不相容

$$B \cap A = \emptyset$$



对立/逆事件：事件 A 不发生的事件称为 A 的对立事件、或逆事件，记 \bar{A}

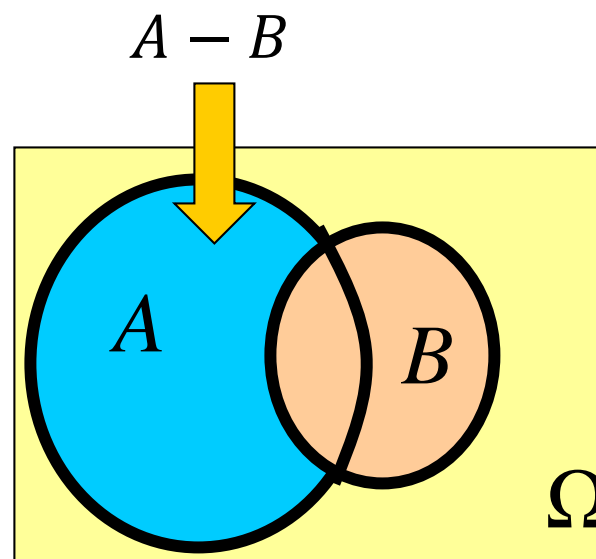
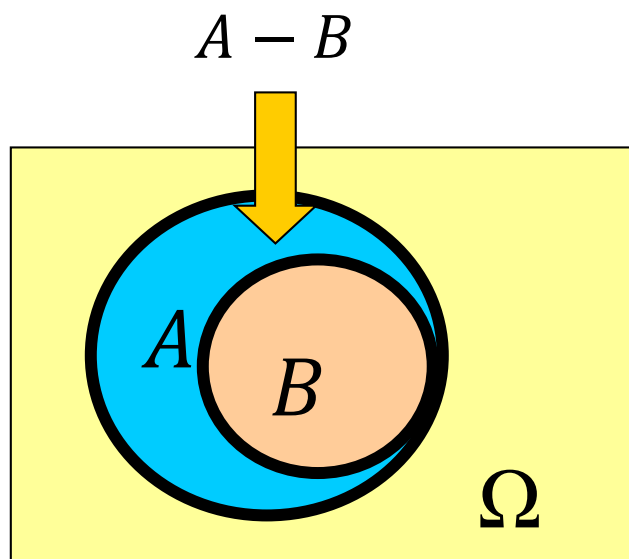
$$A \cap \bar{A} = \emptyset, \quad A \cup \bar{A} = \Omega$$



对立和互不相容事件之间的关系？

事件间的关系

事件的差： A 发生，而 B 不发生的事件称为 A 与 B 的差，记为 **$A - B$**



$$A - B = A - AB = A\bar{B} = (A \cup B) - B$$

概率论与集合论之间的对应关系

记号	概率论	集合论
\bar{A}	A 的对立事件	A 的补集
$A \subset B$	A 发生必然导致 B 发生	A 是 B 的子集
$A = B$	事件 A 与事件 B 相等	集合 A 与集合 B 相等
$A \cup B$	事件 A 与事件 B 的和	集合 A 与集合 B 的并集
$A \cap B$	事件 A 与 B 的积事件	集合 A 与集合 B 的交集
$A - B$	事件 A 与事件 B 的差	集合 A 与集合 B 的差集
$A \cap B = \emptyset$	事件 A 与事件 B 互不相容	集合 A 与集合 B 中没有相同的元素

事件的运算规律

- 幂等律: $A \cup A = A, A \cap A = A$
- 交换律: $A \cup B = B \cup A, A \cap B = B \cap A$
- 结合律: $(A \cup B) \cup C = A \cup (B \cup C)$
 $(A \cap B) \cap C = A \cap (B \cap C)$
- 分配律: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
 $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- 对偶律: $\overline{A \cup B} = \bar{A} \bar{B}, \overline{A \cap B} = \bar{A} \cup \bar{B}$

上述规律可推广到多个事件

例题

设A、B、C为三个随机事件，试用A、B、C表示如下随机事件：

- 事件A与B同时发生，而事件C不发生的事件：
- 三个事件中至少有一个发生的事件：
- 这三个事件中恰好有一个发生的事件：
- 这三个事件中至多有一个发生的事件：
- 这三个事件中至少有两个发生的事件：
- 这三个事件中至多有两个发生的事件：
- 这三个事件中恰好有两个发生的事件：

例题

设A、B、C为任意三个随机事件，证明

$$(\bar{A} \cup B)(A \cup \bar{B})(\bar{A} \cup \bar{B})(A \cup B) = \emptyset$$

$$(A - B) \cup (B - C) = (A \cup B) - BC$$