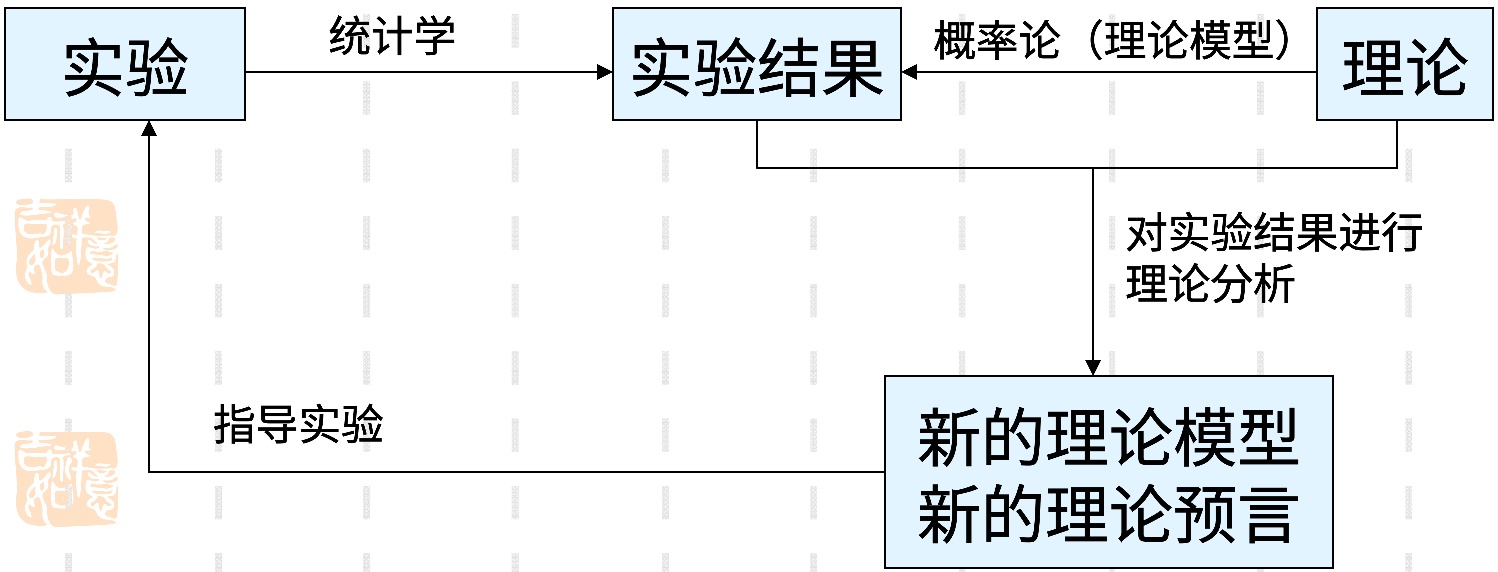
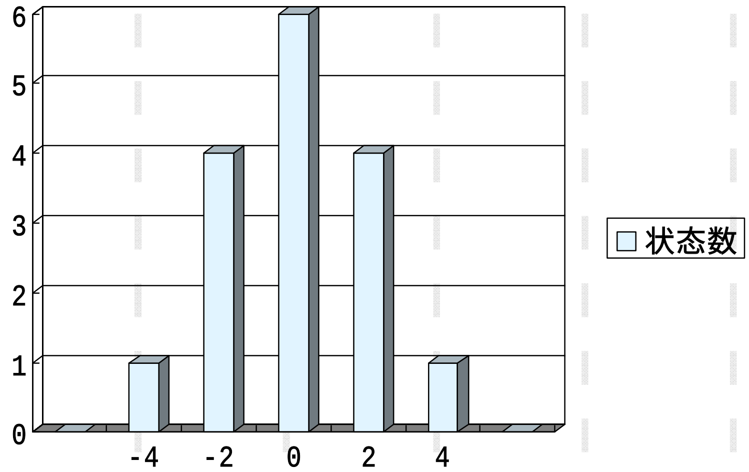
实验过程中的概率论与数理统计

- **Probability & Statistics**
  概率论和统计学（数理统计）——数学的两个分支
- 统计推断（**Statistical Reasoning**）
  *根据带随机性的**观测数据**，按照问题的条件，选用一定的**模型**，而对未知事物作出的、以**概率**形式表达的推断。*
  *例如，**温度**作为物体冷热程度的量度，并不描述组成物体的单个分子的能量，而是正比于分子动能的统计平均值。*
  与演绎推理相对。

实验与理论之间的关系



通过概率论将理论具体化为统计分布



二项式分布

组合态	态的数目(2^4)	磁矩
↑ ↑ ↑ ↑	1	-4
↑ ↑ ↑ ↓	4	-2
↑ ↑ ↓ ↓	6	0
↑ ↓ ↓ ↓	4	2
↓ ↓ ↓ ↓	1	4

理论 →概率和统计← 实验

设某一可观测量x和一组参数θ之间存在某种对应关系（未知）

理论：给定一组参数θ，可观测量x的预期分布形式什么？➔概率论的问题

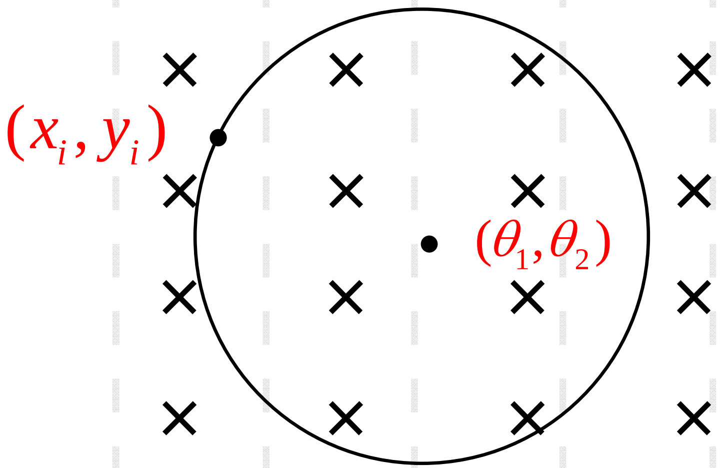
实验：给定一组x的观测值，确定θ的取值是什么？➔统计学的问题

统计学在粒子物理中的应用主要包括两部分：

参数估计（**Parameter Estimation**）：从一组观测值，用统计的方法求出未知参数θ的值及其误差。
径迹拟合（中，**P308**）

假设检验（**Hypothesis Testing**）：检验理论模型所预言的参数值是否与实验所得的值一致 。
反应角分布（中，**P369**）

径迹拟合——参数估计的例子



$$r_i^2 = (x_i - \theta_1)^2 + (y_i - \theta_2)^2$$

观测量：粒子坐标 (x_i, y_i)

参数：半径r_i，圆心坐标(θ_1, θ_2)

概率的基本概念

概率(Probability)的定义

物理学家的定义：频数极限

设在某实验中观测到了n个事例，其中种类为E的事例出现了r次，则某事例的种类为E的概率定义为：

$$p(E) = \frac{r}{n} \quad n \to \infty \qquad 0 \le p(E) \le 1$$

数学家的定义：利用集合理论（Set Thoery）

数学上采用集合空间上的概率测度来定义概率

定义Ω是所有可能的事件E_i的一个集合，其中E_i是互斥的（即 它们中的一个发生时，所有其它的事件都不发生），定义事件E_i发生的概率$p(E_i)$具有如下的性质：

$$\left.\begin{aligned} &a) \quad p(E_i) \ge 0; \\ &b) \quad p(E_i or E_j) = p(E_i) + p(E_j); \\ &c) \quad \sum_i p(E_i) = 1 \end{aligned}\right\} \quad 0 \le p(E) \le 1$$

如果p(E)=0, 则表示事件E总不发生；

如果p(E)=1, 则表示事件E总发生；

集合的概念怎样同物理的概率定义对应起来？

集合	➔	测量值的全体
元素	➔	一组测量中的一次测量值
子集	➔	符合特定条件的多次测量值(可能是相同的一些测量值)
概率	➔	频数

随机变量、样本空间

随机变量：

其值不能完全确定地预测的变量。

样本空间：

随机变量x的取值空间。

离散型随机变量：

若随机变量X只能取有限数目的值，则称x为离散型随机变量；其中，X的一个取值对应着集合中的一个元素，X的一种可能值的全体对应着集合中的一个子集.

若p_i为离散型随机变量x取值为x_i的概率：$p(x=x_i)=p_i$，则 $\sum_i p_i = 1$

连续型随机变量：

若随机变量x在有限取间内的取值是连续的，则称x为连续型的随机变量。连续型随机变量x的取值位于区间[x,x+dx]的概率定义为：

$$p(x \le X \le x + dx) = f(x)dx$$

其中，f(x)为概率密度函数➔p.d.f(probability density function),满足归一化条件

$$\int_\Omega f(x)dx = 1$$

概率的性质

以集合理论为基础介绍概率的一些运算规则。

一、集合（set）

集合是指一些具有相同性质的元素的全体。

集合A的元素(element)：

属于集合A的某一元素；

集合A的子集(subset)：

如果集合B的任一元素又是集合A的元素，则称B为A的子集；

集合的补集（complement）：

设A是样本空间Ω中的任一组元素的集合，则A的补集定义为Ω中所有不属于A的元素的集合，记为：$\overline{A}$

A和B的并集(union)：

属于A或属于B元素的集合，记为：$A \cup B$

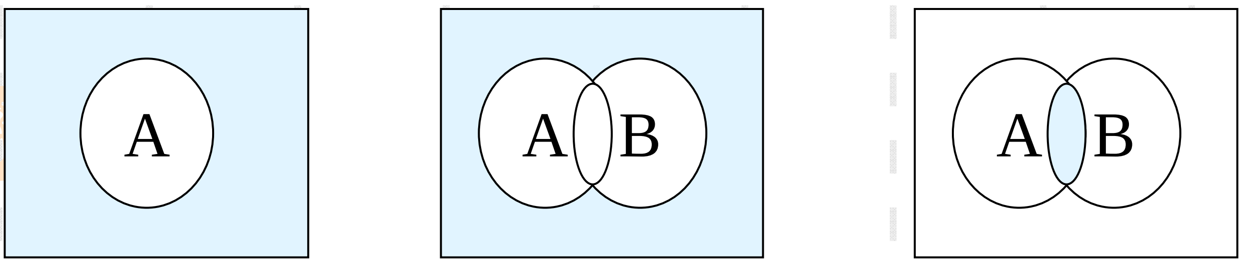
A和B的交集（intersection）：

既属于A又属于B的一些元素的集合，记为：$A \cap B$

如果 $A \cup B = \Omega$ ，则称A和B为完备集(Exhaustive Sets)；

如果 $A \cap B = 0$ ，则称A和B是互斥集(Exclusive Sets)——正交集。

维因图(Venn Diagram)



$\overline{A}$ $A \cup B$ $A \cap B$

例: PP相互作用事例的分类：

p+p→(2,4,6,…)个带电粒子+(0,1,2,..)个V^0粒子

定义:

A为至少有一个V^0的事例的集合;

B为具有两个带电粒子的事例的集合;

$\bar{A}$:没有V^0的事例的集合

$\bar{B}$:具有两个以上带电粒子的事例的集合

$A \cup B$:至少有一个V^0,或有两个带电粒子,或至少有一个V^0且有两个带电粒子的事例的集合;

$A \cap B$:有两个带电粒子且至少有一个V^0的事例的集合.

二、概率的加法定律(Addition rule of probability)

定义:

$p(A)$: 集合A中某一事件发生的概率;

$p(B)$: 集合B中某一事件发生的概率;

$p(A \cup B)$: 属于A或属于B的某一事件发生的概率;

$p(A \cap B)$: 既属于A又属于B的某一事件发生的概率;

加法定律:

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

推广到N个集合的情况:$A_1, A_2, \ldots A_N$,属于至少其中一个集合A_i的某一事件发生的概率（用Venn图来解释）——约当（Jordan）公式:

$$p(A_1 \cup A_2 \cup \ldots A_N) = S_1 - S_2 + S_3 - \ldots - (-1)^N S_N$$

其中:

$$S_1 = \sum_{i=1}^{n} p(A_i) \quad S_2 = \sum_{i<j=2}^{n} p(A_i \cap A_j)$$

$$S_3 = \sum_{i<j<k=3}^{n} p(A_i \cap A_j \cap A_k) \quad \cdots$$

三、条件概率(Conditional probability)

定义：

在事件A已经产生的情况下，事件B产生的概率，记为：p(B|A)

意义：

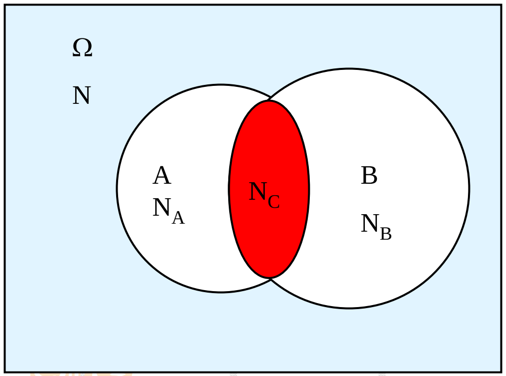
假设A和B是样本空间Ω的两个子集，如果我们只对A中的元素感兴趣，并将样本空间重定义为子集A，在新的样本空间A中子集B的概率称为B相对于A的条件概率。

条件概率与A和B的交集的关系：

P(B|A)由下式定义的

$$p(A \cap B) = p(B \mid A)\,p(A)$$

该式的意义可由下面的Venn图说明：



$$p(A)=\frac{N_A}{N} \qquad p(B)=\frac{N_B}{N}$$

$$p(A\cap B)=\frac{N_C}{N}$$

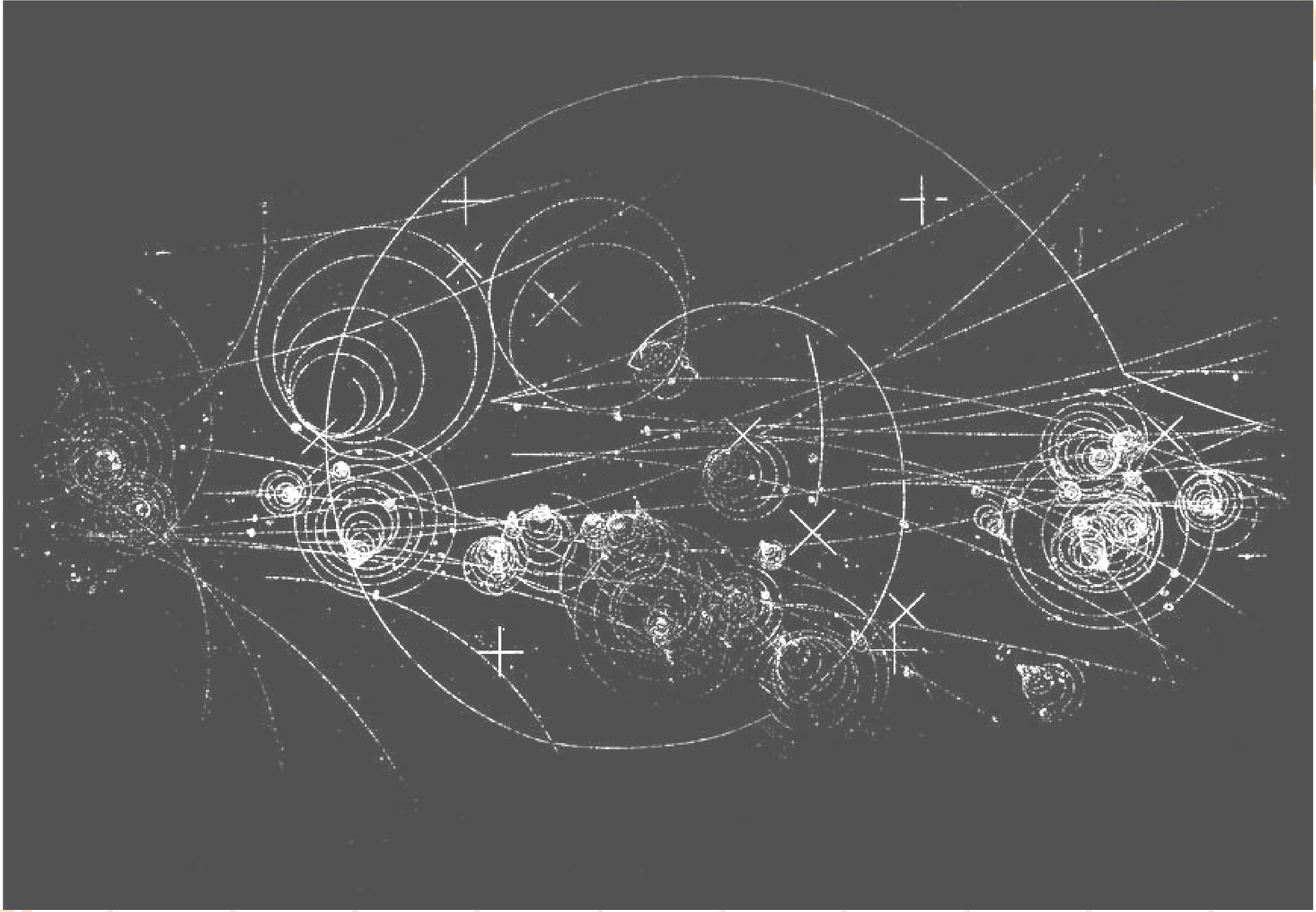
$$p(B\mid A)=\frac{N_C}{N_A}$$

$$p(A\mid B)=\frac{N_C}{N_B}$$

$$\therefore\ p(A\cap B)=\frac{N_C}{N}=\frac{N_C}{N_A}\cdot\frac{N_A}{N}=p(B\mid A)\,p(A)$$

$$=\frac{N_C}{N_B}\cdot\frac{N_B}{N}=p(A\mid B)\,p(B)$$

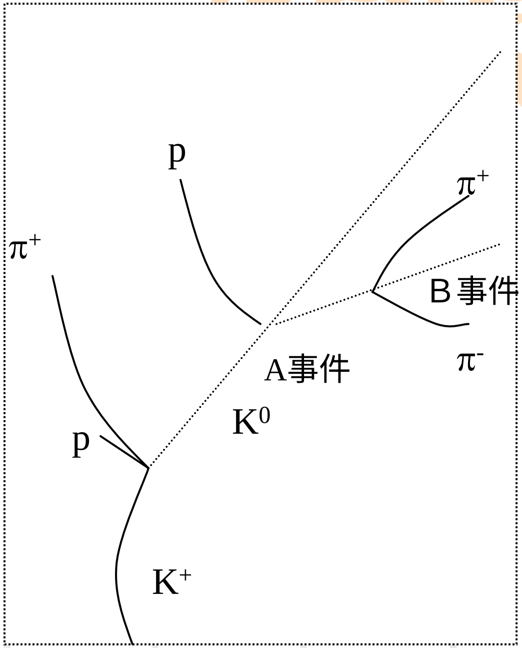
实验上，所有的概率都是条件概率，因为事例都是在一定的实验条件下获取的，只是由于这些条件对所有的事例都相同，因而被认为是无关紧要的。



Traces from a bubble chamber that was used from 1964 to 1972 for research into the smallest constituents of matter. It was installed in "DESY", Hamburg's first ring accelerator. Inside the bubble chamber there was a tank filled with liquid hydrogen. Tiny particles flying through the tank make the hydrogen boil. Hydrogen vaporizes and forms small bubbles along the paths of the particles. These traces were photographed. An analysis of the pictures yielded information about the types of the particles and their properties. (Source: DESY

条件概率的例子：K^0p散射截面

K^0的产生：$K^+ + p \rightarrow K^0\pi^+p$ ➔事例数为N
K^0的探测：$K^0 \rightarrow \pi^+\pi^-$ ➔事件B
感兴趣的相互作用：**$K^0p \rightarrow K^0p$ ➔事件A**
求：事件A的概率p(A)



只有在观测到事件B的条件下才能确定事件A的产生

$$p(A \cap B) = p(B \mid A)\, p(A)$$

$p(A \cap B)$:

既发生了K^0p散射又探测到了K^0衰变的概率

$$K^+ + p \rightarrow K^0\pi + p$$
$$\hookrightarrow K^0p \rightarrow K^0p$$
$$\hookrightarrow \pi^+\pi^-$$

➔探测到的事例数N_1➔ $\quad p(A \cap B) = \dfrac{N_1}{N}$

p(B|A): 在K^0p散射发生的条件下K^0衰变的概率

$$p(B\,|\,A)=\varepsilon\sigma Br(K^0\to\pi^+\pi^-)$$

其中： ε：探测效率

$Br(K^0\to\pi^+\pi^-)$：$K^0\to\pi^+\pi^-$ 的分支比

σ：K^0衰变的截面

K^0p散射截面：

$$p(A)=\frac{N_1}{\varepsilon N\sigma Br(K^0\to\pi^+\pi^-)}$$

四、独立性，乘法定律（Independence, Multiplication Rule）

如果事件B的产生与事件A的产生无关，则称A和B是不相关的（Independent），即

$$p(B \mid A) = p(B)$$

$$\therefore \quad p(A \cap B) = p(B \mid A) \cdot p(A) = p(A) \cdot p(B)$$

不相关事件的概率乘法定律

如果事件A和B是不相关的，则A和B都发生的概率等于A事件发生的概率乘以B事件发生的概率.

多个事件的不相关性

（定义更复杂）

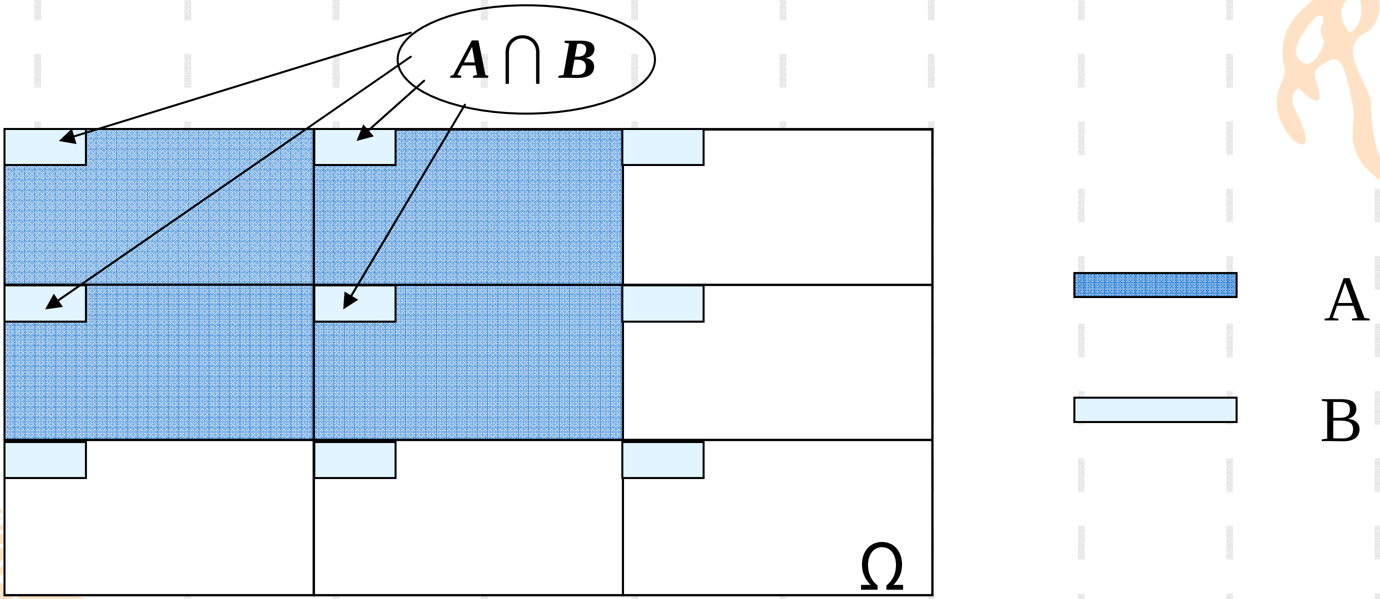
$$p(A_i \cap A_j) = p(A_i) \cdot p(A_j) \qquad (i \neq j;\ i, j = 1, \ldots, n)$$

$$p(A_i \cap A_j \cap A_k) = p(A_i) \cdot p(A_j) \cdot p(A_k)$$

……

$$p(\underbrace{A_i \cap A_j \cap \ldots \cap A_l}_{m}) = \underbrace{p(A_i) \cdot p(A_j) \ldots p(A_l)}_{m}$$

$$(i \neq j \neq \mathrm{k} \neq \ldots \neq l;\ i, j, k \ldots l = 1, \ldots, n)$$



A和B的不相关性（Independence）可以理解为上面的Venn图

一个推论：

$$P(B \mid \overline{A}) = P(B)$$

$$\because \quad \frac{N_B - N_c}{N - N_A} = \frac{N_B}{N}$$

不相关事件的概率加法定律

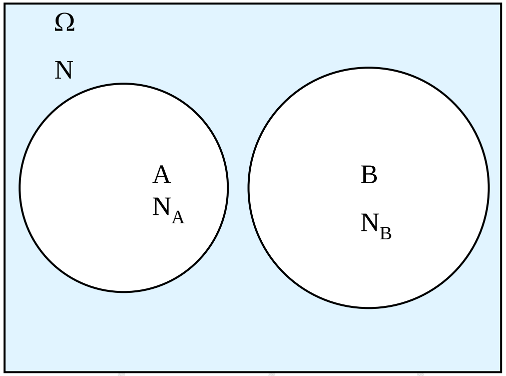
$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

$$= p(A) + p(B) - p(A) \cdot p(B)$$

互斥性（正交性）不等于不相关性（独立性）

对互斥事件：

$$p(A \cap B) = p(B|A) \cdot p(A) = 0$$



$N_C=0$

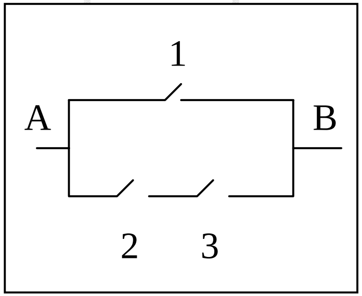
因为 $$p(B|A) = 0$$

但是 $$p(A) \cdot p(B) \neq 0$$

五、概率运算的几个例子：

例1：开关网络

α：每个开关接通的概率，每个开关的接通是互不相关的，求A、 B两端有电流通过的概率



E_i：第i个开关接通的事件：$p(E_i)=\alpha$， $i=1,2,3$
有电流流过的事件：2和3同时接通或1接通：

真值表——布尔代数

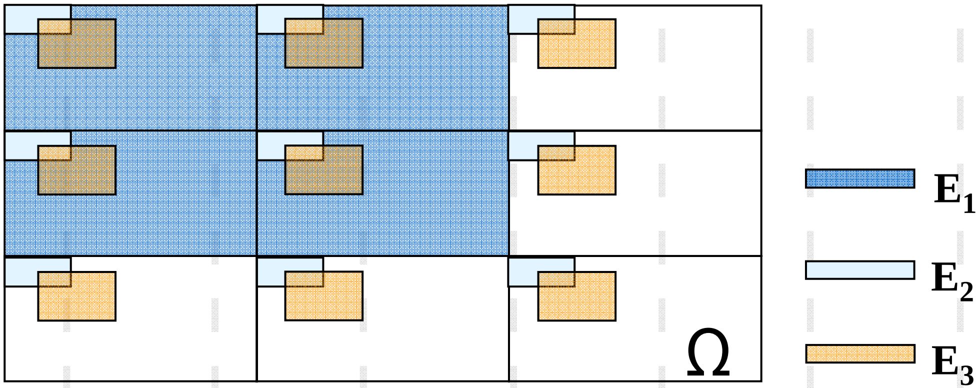
E_1	E_2	E_3	E
0	0	0	0
0	0	1	0
0	1	0	0
0	**1**	**1**	1
1	**0**	**0**	1
1	**0**	**1**	1
1	**1**	**0**	1
1	**1**	**1**	1

我们感兴趣的事件（布尔代数）：

$$\begin{aligned} \boldsymbol{E} &= \boldsymbol{E}_1+\bar{\boldsymbol{E}}_1(\boldsymbol{E}_2\boldsymbol{E}_3) \\ &= \boldsymbol{E}_1(1+(\boldsymbol{E}_2\boldsymbol{E}_3))+\bar{\boldsymbol{E}}_1(\boldsymbol{E}_2\boldsymbol{E}_3) \\ &= \boldsymbol{E}_1+\boldsymbol{E}_1(\boldsymbol{E}_2\boldsymbol{E}_3)+\bar{\boldsymbol{E}}_1(\boldsymbol{E}_2\boldsymbol{E}_3) \\ &= \boldsymbol{E}_1+(\boldsymbol{E}_2\boldsymbol{E}_3) \end{aligned}$$

$$\Rightarrow \boldsymbol{E}=\boldsymbol{E}_1\cup(\boldsymbol{E}_2\cap\boldsymbol{E}_3)$$

利用Venn图来理解:



$$p(E_1 \cup (E_2 \cap E_3))$$
$$= p(E_1) + p(E_2)p(E_3) - p(E_1)p(E_2)p(E_3)$$

$$
\begin{aligned}
P(E) &= p(E_1 \cup (E_2 \cap E_3)) \\
&= p(E_1) + p(E_2 \cap E_3) - p(E_1 \cap (E_2 \cap E_3)) && \leftarrow \text{概率加法} \\
&= p(E_1) + p(E_2)p(E_3) - p(E_1)p(E_2)p(E_3) && \leftarrow \text{概率乘法} \\
&= \alpha + \alpha^2 - \alpha^3
\end{aligned}
$$

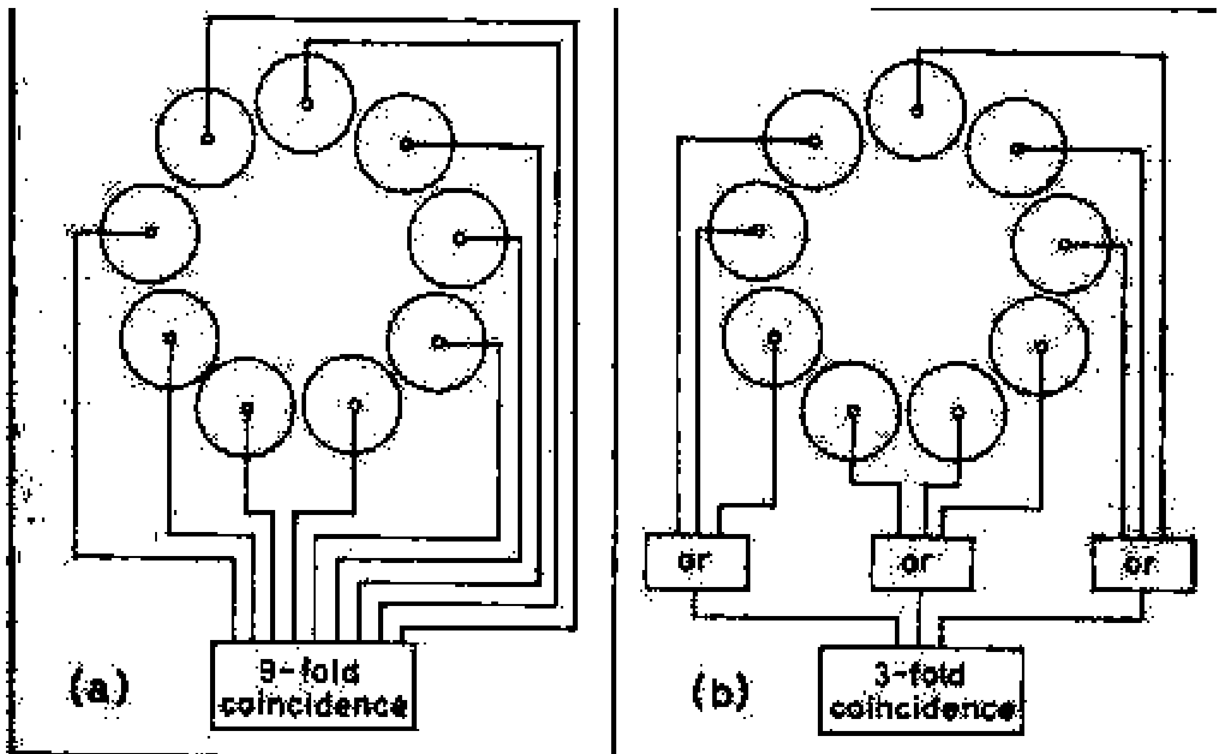
其中 $p(E_1 \cap (E_2 \cap E_3)) = p(E_1)p(E_2)p(E_3)$

包含在“多个事件不相关性”的定义中。

例2：Cerenkov计数器的效率

9只光电倍增管环绕契伦科夫计数器的轴线排列成一圈，当有带电粒子沿计数器轴线穿过时，每个光电倍增管都可探测到粒子所发出的契伦科夫光。

设每只PMT的探测效率为ε=0.93，且每只PMT对契伦科夫光的探测是相互独立的，求：契伦科夫计数器的效率P



E：某只PMT有信号输出的事件，p(E)= ε=0.93

a) 如果要求9只PMT都有输出：

$$p=(p(E))^9=\varepsilon^9=0.52$$ ➔概率乘法定律

b) 将9只PMT分成3组，每组有3只，如果每组至少有一只有信号，则认为该组有信号输出，最后要求三组都有信号

$$\begin{aligned}
p_1 &= p(E \cup E \cup E) \\
&= p(E) + p(E \cup E) - p(E \cap (E \cup E)) \\
&= p(E) + p(E) + p(E) - p(E \cap E) - p(E)p(E \cup E) \\
&= 3p(E) - p(E)p(E) - p(E)[p(E) + p(E) - p(E \cap E)] \\
&= 3p(E) - 3p(E)p(E) + p(E)p(E)p(E) \\
&= 3\varepsilon - 3\varepsilon^2 + \varepsilon^3
\end{aligned}$$

$$\therefore \quad p = p_1 \cdot p_2 \cdot p_3 = (3\varepsilon - 3\varepsilon^2 + \varepsilon^3)^3 = 0.999$$

例3：π^0的探测效率

$\pi^0 \rightarrow \gamma\ \gamma$

$\gamma \rightarrow e^+e^-$ ➔对产生过程

$\gamma \rightarrow e^+e^-$

α：发生对产生的概率，**1-** α：不发生对产生的概率

E：γ 发生对产生的事件，**p(E)**= α, 两γ发生对产生的事件是互不相关的

a)探测到两个γ的概率： $p(2\gamma)=p(E)p(E)=\alpha^2$

b)没有γ被探测到的概率：$p(0\gamma)=(1-\alpha)^2$

c)只有一个γ被探测到的概率：

E_1:第一个γ被探测到，第二个γ没有被探测到

E_2:第二个γ被探测到，第一个γ没有被探测到

$p(E_1)=p(E_2)=\alpha(1-\alpha)$

E_1和E_2是互斥的事件：两事件同时发生的概率为零

$$p(1\gamma)=p(E1\cup E2)=p(E1)+p(E2)=2\alpha(1-\alpha)$$

显然： $p(2\gamma)+p(1\gamma)+p(0\gamma)=1$

d)至少探测到一个γ的概率：

$$
\begin{aligned}
\boldsymbol{p}(\geq 1\gamma) &= \boldsymbol{p}(\boldsymbol{E}\cup\boldsymbol{E}) \\
&= \boldsymbol{p}(\boldsymbol{E})+\boldsymbol{p}(\boldsymbol{E})-\boldsymbol{p}(\boldsymbol{E}\cap\boldsymbol{E}) \\
&= 2\alpha-\alpha^{2} = \boldsymbol{p}(2\gamma)+\boldsymbol{p}(1\gamma)
\end{aligned}
$$

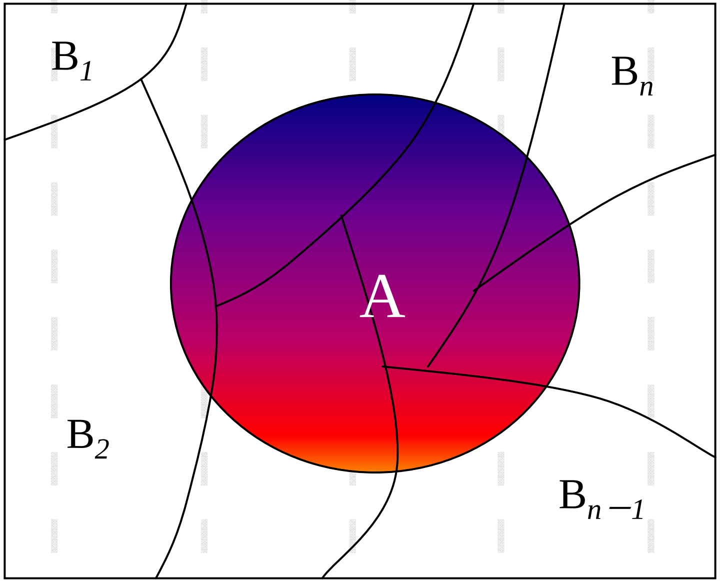
六、边缘概率（Marginal probability)

实验所获取的事例通常可按不同的标准进行分类，如果忽略掉某些分类标准而只考虑在某一种分类标准下某事件出现的概率，则称这种概率为边缘概率

例：粒子的单举产额

设实验事例按A、B两种分类标准可分为：$A_1, A_2, \ldots . A_m$　$B_1, B_2, \ldots . B_n$，且

$$\sum_{i=1}^{m} p(A_i) = 1 \quad \sum_{i=1}^{n} p(B_i) = 1$$



则Ai的边缘概率定义为

$$p(A_i) = \sum_{j=1}^{n} p(A_i \cap B_j)$$

同样，Bi的边缘概率定义为

$$p(B_i) = \sum_{j=1}^{n} p(B_i \cap A_j)$$

例：pp相互作用事例的分类

p+p→(2,4,6,…)个带电粒子+(0,1,2,..)个V^0粒子

分类标准A:事例中有V^0

分类标准B:事例中带电粒子的数目

A	A_1	A_2	A_3	A_4	
	K_s^0	Λ	$K_s^0K_s^0$	$K_s^0\Lambda$	
B	B1	B2	B3	B4	B5
	2	4	6	8	10

如果实验中已测出了概率 $p(A_i \cap B_j)$ ，如

$p(A_1 \cap B_1)$ ：具有两个带电粒子且具有1个 K_s^0 的概率

则边缘概率P(A_1),即有一个 K_s^0 而不管有几个带电粒子的事例的概率为

$$p(A_1) = \sum_{j=1}^{5} p(A_1 \cap B_j)$$

贝叶斯(Bayes)定理

定理：设样本空间Ω被分成了n个互斥的完备事件组Bi，即

$$B_1+B_2+\ldots+B_n=\Omega$$

$$\sum_{i=1}^{n} p(B_i)=1$$

若A也是属于Ω的一个事件组，则有

$$p(B_i\mid A)=\frac{p(A\mid B_i)p(B_i)}{\sum_{j=1}^{n} p(A\mid B_j)p(B_j)}$$

证明：利用概率运算定律

1.根据条件概率的定义

$$p(A\cap B_i)=p(B_i\mid A)p(A)=p(A\mid B_i)p(B_i)$$

$$\therefore \quad p(B_i\mid A)=\frac{p(A\mid B_i)p(B_i)}{p(A)}$$

2.根据边缘概率的定义

$$p(A)=\sum_{j=1}^{n} p(A\cap B_j)=\sum_{j=1}^{n} p(A\mid B_j)p(B_j)$$

代入上式即得贝叶斯定理

贝叶斯定理给出了两个条件概率$p(B_i|A)$和$p(A|B_i)$之间的关系

贝叶斯定理的例子

三个抽屉分别装有金币和银币

B_1：两个金币

B_2：一个金币和一个银币

B_3:两个银币

随机地选一个抽屉并从中取出一个钱币，假定取出的是一个金币，求同一抽屉中另一个钱币是金币的概率

A:第一次取出金币的事件，另一个钱币也是金币条件要求只能选取B_1,即要求的是在A发生的条件下，选择抽屉B_1的概率： $p(B_1|A)$

在选定抽屉的情况下，取出一个金币的概率

$P(A|B_1)=1$, $p(A|B_2)=0.5$, $p(A|B_3)=0$

选择某一抽屉的概率

$P(B_1)=p(B_2)=p(B_3)=1/3$

由贝叶斯定理

$$p(B_1|A)=\frac{p(A|B_1)p(B_1)}{\sum_{j=1}^{3}p(A|B_j)p(B_j)}=\frac{1\cdot\frac{1}{3}}{1\cdot\frac{1}{3}+\frac{1}{2}\cdot\frac{1}{3}+0\cdot\frac{1}{3}}=\frac{2}{3}$$