Peking University

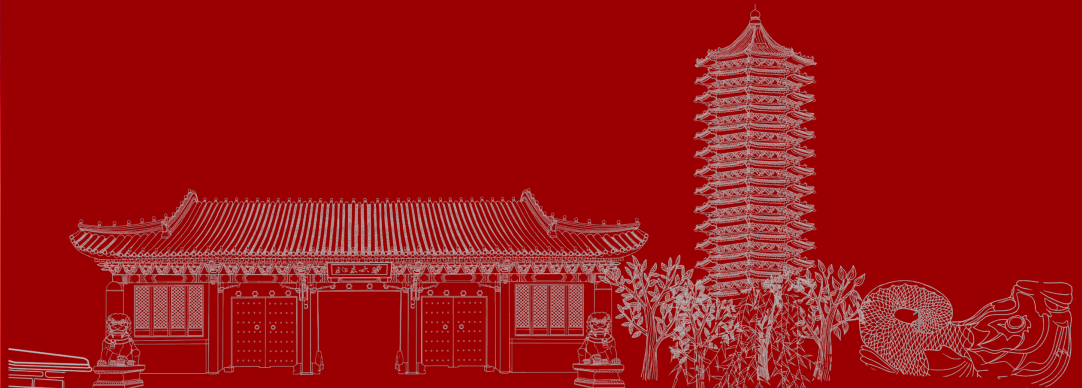
北京大学
PEKING UNIVERSITY

人工智能引论

2. AI 数学基础

授课教师 张牧涵

2024年02月22日



目录

- 样本空间、随机事件
- 古典概率、条件概率
- 全概率公式、Bayes公式
- 事件的独立性、随机变量
- 数学期望、方差和标准差
- 协方差、相关系数

概率论基础

1.0 确定性现象和随机现象

确定性现象：在一定条件下必然发生的现象，在实验或者观察前即可预知确切结果

随机现象：在一定条件下可能出现多种结果，而在实验或者观察前不能预知确切的结果的现象。

确定性现象例子：

- 在地面上抛一个球必然落回地面
- 标准大气压下，水加热到100摄氏度必然沸腾
- 太阳必然从东方升起
- 在AI引论课上，只完成一个lab必然不及格
- ……

随机现象例子：

- 掷三个骰子，刚好出现666
- 买50张彩票中了5元钱
- 2026年世界杯阿根廷卫冕成功
- 一个网站在一段时间内的点击数达到1000
- 薛定谔的猫还活着 (?)
- 在AI引论课期末考试拿到了95分 (???)
- ……

概率论基础

1.0 确定性现象和随机现象

确定性现象：在一定条件下必然发生的现象，在实验或者观察前即可预知确切结果

随机现象：在一定条件下可能出现多种结果，而在实验或者观察前不能预知确切的结果的现象。

单次实验结果具有不确定性，但是大量重复实验具有**统计规律性**的现象

随机现象统计规律的例子：

- 掷10000次硬币，出现正面的次数接近5000次
- 家园食堂的人数在每天十一点和十二点附近有两个高峰
- 封闭空间内气体分子速度的分布随温度变化
- ……

概率论基础

1.0 确定性现象和随机现象

确定性现象：在一定条件下必然发生的现象，在实验或者观察前即可预知确切结果

随机现象：在一定条件下可能出现多种结果，而在实验或者观察前不能预知确切的结果的现象。

单次实验结果具有不确定性，但是大量重复实验具有**统计规律性**的现象

概率论：揭示**随机现象统计规律性**的数学学科

概率论基础

1.1 样本空间与样本点

样本空间：随机试验 E 所有可能的结果组成的集合称为样本空间，记为 Ω.

样本点：样本空间的元素，即随机试验 E 的可能结果，记为 ω，$\Omega = \{\omega|\omega\text{为样本点}\}$.

随机试验 E_1：将一枚硬币投掷2次，正面记为 H，反面记为 T.

样本空间为：$\Omega_1 = \{HH, HT, TH, TT\}$

随机试验 E_2：投掷一颗骰子，观察出现的点数。

样本空间为：$\Omega_2 = \{1, 2, 3, 4, 5, 6\}$

有限样本空间

随机试验 E_3：观察一个新灯泡的寿命。可能的结果有无穷多个。

样本空间为：$\Omega_3 = \{t|0 \le t < +\infty\}$

无限样本空间

不包含任何样本点的空间，叫做空集，记为$\emptyset$

PEKING UNIVERSITY

1.2 随机事件

随机事件：**满足某些条件**的样本点组成的集合。它是样本空间 Ω 的子集，记为 $A, B, \ldots$

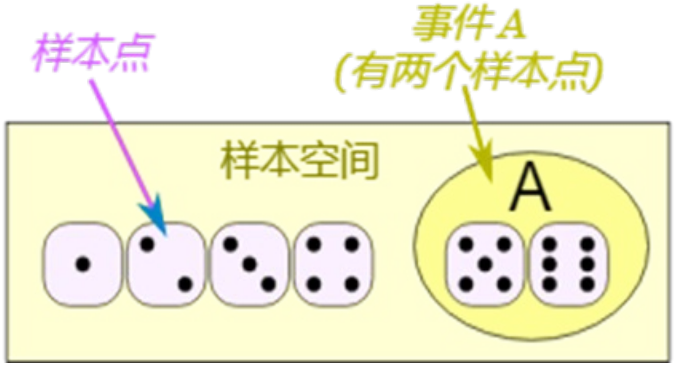
例如：对于试验 E_1，以下A, B, C都是随机事件：

$$A = \{\text{至少出一个正面}\} = \{HH, HT, TH\}$$

$$B = \{\text{两次出现同一面}\} = \{HH, TT\}$$

$$C = \{\text{恰好出现一次正面}\} = \{HT, TH\}$$

例如：对于试验E_2,以下A, B都是随机事件：



$$A = \{\text{掷出奇数点}\} = \{1, 3, 5\}$$

$$B = \{\text{掷出偶数点}\} = \{2, 4, 6\}$$

必然事件？
不可能事件？

随机事件发生 ⟷ 样本空间中的一个样本点发生

概率论基础

1.2 随机事件的关系

包含关系 $A \subseteq B$：事件A发生则事件B一定发生

事件的并 $A \cup B$：事件A与事件B至少有一个发生

事件的交 $A \cap B$：事件A与事件B同时发生

对立事件 $\bar{A}$：$\bar{A} \cup A = \Omega$，$\bar{A} \cap A = \emptyset$

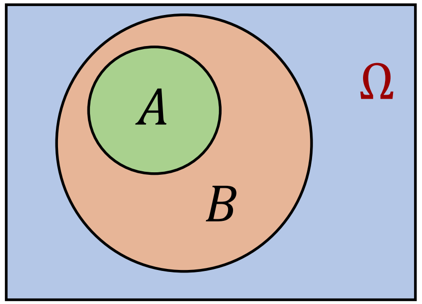
互斥事件A，B：$A \cap B = \emptyset$

例如：$\boldsymbol{A} = \{\text{掷出奇数点}\} = \{\mathbf{1}, \mathbf{3}, \mathbf{5}\}$，

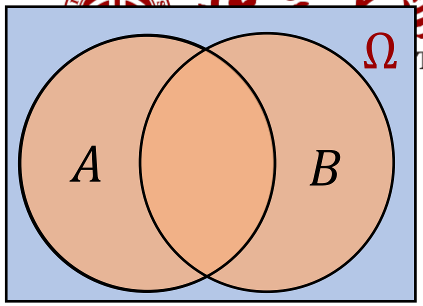
$\boldsymbol{B} = \{\text{掷出偶数点}\} = \{\mathbf{2}, \mathbf{4}, \mathbf{6}\}$，

$\boldsymbol{C} = \{\text{掷出2点}\} = \{\mathbf{2}\}$， $\boldsymbol{D} = \{\text{掷出5点}\} = \{\mathbf{5}\}$

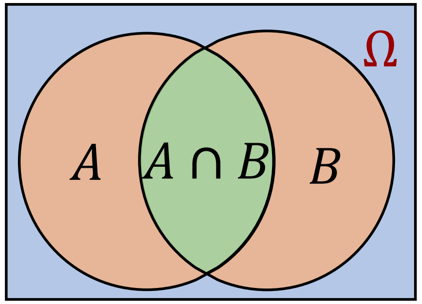
思考： $\boldsymbol{A}$和$\boldsymbol{B}$， $\boldsymbol{A}$和$\boldsymbol{C}$， $\boldsymbol{A}$和$\boldsymbol{D}$ 分别什么关系？



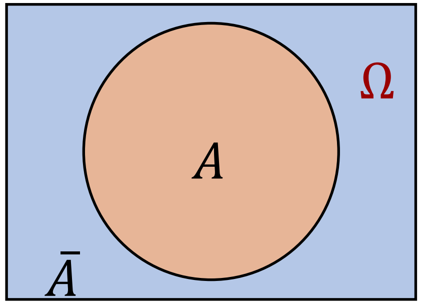
包含关系 $A \subseteq B$



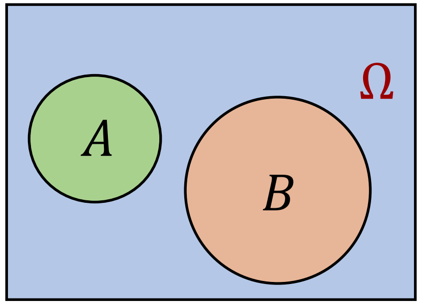
事件的并 $A \cup B$



事件的交 $A \cap B$



对立事件$\bar{A}$



互斥事件

北京大学 PEKING UNIVERSITY

1.3 古典概型

古典概型：设 Ω 为随机试验 E 的样本空间，若：

① **有限性**：只有**有限**个试验结果（样本）$\{\omega_1, \omega_2, \cdots, \omega_n\}$；

② **等可能性**：每个试验结果（样本）在一次试验中出现的可能性**相等**；

则称具有以上两个特征的随机试验数学模型为**古典概型**。

古典概型概率：设古典概型试验 E 所有可能结果（样本）有 n 个，为 $\{\omega_1, \omega_2, \cdots, \omega_n\}$，事件 A 包含其中 m 个结果（样本），则

$$P(A) = \frac{m}{n} = \frac{A\text{包含的试验结果个数}}{\text{试验结果的总数}} = \frac{A\text{包含}\Omega\text{中样本点的个数}}{\Omega\text{中样本点的总数}}$$

$$P(\Omega) = 1 \qquad P(\emptyset) = 0$$

PEKING UNIVERSITY

1.3 古典概型

例：将一颗骰子连掷两次，试求下列事件的概率。

（1）两次掷得的点数之和为8；

（2）第二次掷得3点。

解：

设 A 表示“点数之和为8”事件，B 表示“第二次掷得3点”事件。

样本空间

$\Omega=\{(1,1),(1,2),\cdots,(1,6),(2,1),\cdots,(2,6),\cdots,(6,1),(6,6)\}$共36种等概率结果

$$A=\{(2,6),(3,5),(4,4),(5,3),(6,2)\}$$

$$B=\{(1,3),(2,3),(3,3),(4,3),(5,3),(6,3)\}$$

所以 $P(A)=\frac{5}{36}$，$P(B)=\frac{6}{36}=\frac{1}{6}$

概率论基础

1.3 古典概型

不同关系的两个随机事件各自概率的关系：

包含关系 $A \subseteq B$：$P(B) \geq P(A)$

事件的并 $A \cup B$：$P(A \cup B) \geq P(A), P(A \cup B) \geq P(B)$

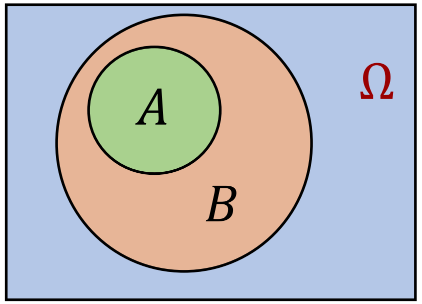
事件的交 $A \cap B$：$P(A \cap B) \leq P(A), P(A \cap B) \leq P(B)$

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

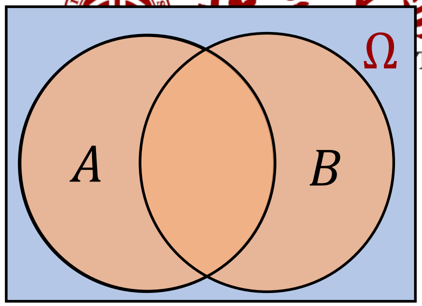
互斥事件 A，B：$P(B) \leq 1 - P(A)$

对立事件 $\bar{A}$：$\boxed{P(\bar{A}) = 1 - P(A)}$

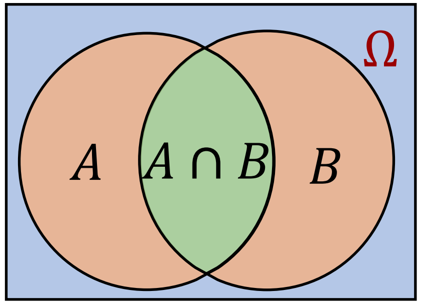
$$P(A \cap B) + P(\bar{A} \cap B) = P(B)$$



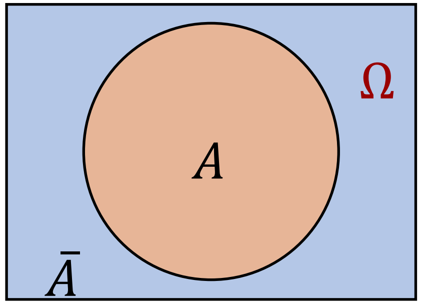
包含关系 $A \subseteq B$



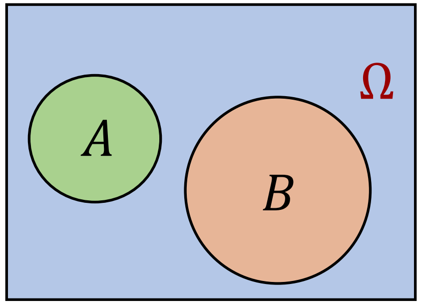
事件的并 $A \cup B$



事件的交 $A \cap B$



对立事件$\bar{A}$



互斥事件

概率论基础

1.4 条件概率

研究：

在事件B**已经出现**的条件下，事件A发生的概率，记作 $\boldsymbol{P(A|B)}$.

问题：

由于附加了条件，$P(A)$ 与 $P(A|B)$ 意义不同，一般 $P(A|B) \neq P(A)$，

它们之间有什么关系？

概率论基础

1.4 条件概率

例：

掷一颗均匀骰子，$A=\{掷出2点\}$，$B=\{掷出偶数点\}$，$P(A|B)=?$

解：

掷一颗骰子可能的结果有6种，并且6种结果等可能，$P(A)=\frac{1}{6}$

由于已知事件 B 已经发生，所以此时试验所有可能结果只有3种，而事件 A 包含的基本事件只占其中一种，故 $P(A|B)=\frac{1}{3}$.

上例中，$P(A|B)\neq P(A)$

原因：

“事件 B 已发生”这个新条件改变了不加条件的样本空间。

PEKING UNIVERSITY 北京大学

1.4 条件概率

研究：

在事件B**已经出现**的条件下，事件A发生的概率，记作 $\boldsymbol{P(A|B)}$.

问题：

由于附加了条件，$P(A)$ 与 $P(A|B)$ 意义不同，一般 $P(A|B) \neq P(A)$，

它们之间有什么关系？

投掷一颗骰子
事件 $\boldsymbol{A}$ = {掷出1点}
事件 $\boldsymbol{B}$ = {掷出偶数点}
$P(A|B)=0$，$P(A) = 1/6$

$P(A|B) < P(A)$

连续投掷一颗骰子两次，
事件 $\boldsymbol{A}$ = {第一次掷出1点}
事件 $\boldsymbol{B}$ = {第二次掷出1点}
$P(A|B)=1/6$，$P(A) = 1/6$

$P(A|B) = P(A)$

投掷一颗骰子
事件 $\boldsymbol{A}$ = {掷出2点}
事件 $\boldsymbol{B}$ = {掷出偶数点}
$P(A|B) = 1/3$，$P(A) = 1/6$

$P(A|B) > P(A)$

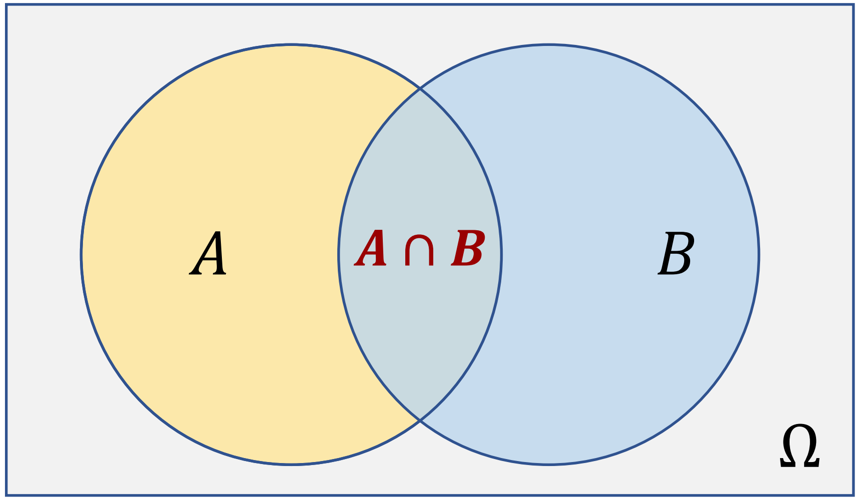
1.4 条件概率

条件概率定义：

设 A，B为两个事件，$P(B)>0$，则称 $\frac{P(AB)}{P(B)}$为在事件 B 发生的条件下事件 A 发生的概率.

记作：$P(A|B)=\frac{P(AB)}{P(B)}$.

注意我们将 $P(\boldsymbol{A}\cap\boldsymbol{B})$ 记为 $P(AB)$



1.4 条件概率

条件概率定义：

设 A，B为两个事件，$P(B)>0$，则称 $\frac{P(AB)}{P(B)}$ 为在事件 B 发生的条件下事件 A 发生的概率.

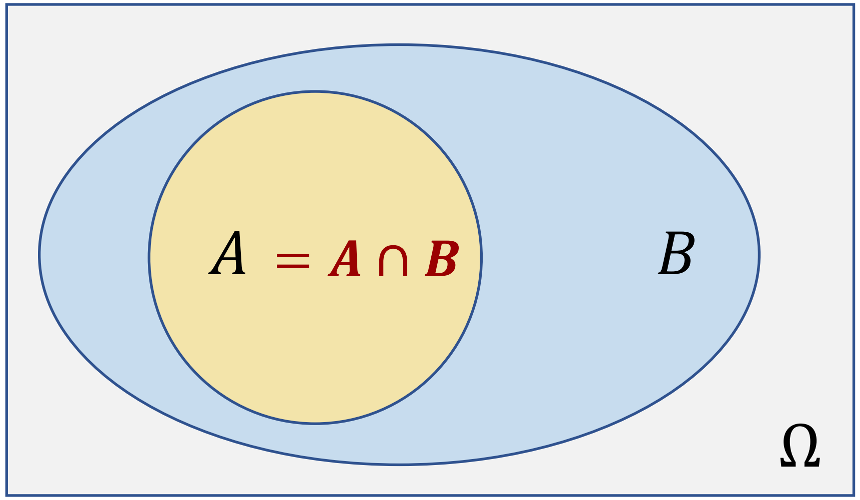
记作：$P(A|B)=\frac{P(AB)}{P(B)}$.

注意我们将 $P(\boldsymbol{A}\cap\boldsymbol{B})$ 记为 $P(AB)$

一般地，条件概率与无条件概率之间的大小无确定关系

若有 $A\subseteq B$，则：

$$P(A|B)=\frac{P(AB)}{P(B)}=\frac{P(A)}{P(B)}\geq P(A)$$



PEKING UNIVERSITY

1.4 条件概率

例：

某人外出旅游两天，需知道两天的天气情况，据预报，第一天下雨的概率为0.6，第二天下雨的概率为0.3，两天都下雨的概率为0.1. 求当第一天下雨时，第二天不下雨的概率。

解：

设 A 与 B 分别表示第一天与第二天下雨。

故 $P(A) = 0.6$，$P(B) = 0.3$，$P(AB) = 0.1$， $P(\bar{B}) = 1 - P(B)$=0.7.

第一天下雨时，第二天不下雨的概率为 $P(\bar{B}|A)$.

$$P(\bar{B}|A) = 1 - P(B|A) = 1 - \frac{P(AB)}{P(A)} = 1 - \frac{0.1}{0.6} = \frac{5}{6}$$

1.4 条件概率

乘法公式：

对任意事件 A，B

$$P(AB) = P(A)P(B|A) \quad (P(A) > 0)$$

$$P(AB) = P(B)P(A|B) \quad (P(B) > 0)$$

推广：

3个事件时：$P(ABC) = P(AB)P(C|AB) = P(A)P(B|A)P(C|AB)$ （其中$P(AB) > 0$）

n个事件时：$P(A_1A_2\cdots A_n) = P(A_1)P(A_2|A_1)\cdots P(A_n|A_1A_2\cdots A_{n-1})$ （其中$P(A_1A_2\cdots A_{n-1}) > 0$）

PEKING UNIVERSITY

1.5 事件的独立性

思考：

如果 $P(B|A) \neq P(B)$，则说明 A 事件的发生对 B 事件的发生产生了影响；否则说明 A 事件的发生对 B 事件的发生没有产生影响，后者就称 A 事件和 B 事件相互独立。

当 $P(B) \neq 0$ 时，$P(A|B) = \frac{P(AB)}{P(B)}$，且 $P(A) = P(A|B)$

$\Leftrightarrow P(AB) = P(A)P(B).$

相互独立：

设 A，B 是两事件，如果满足：

$$P(AB) = P(A)P(B)\text{，}$$

则称 A，B 为相互独立的事件，简称 A，B 独立。

1.5 事件的独立性

多个事件相互独立：

一般地，设 $A_1, A_2, \cdots, A_n$ 为 n 个事件，如果对于任意的 k，$(1 \le k \le n)$，任意 $1 \le i_1 < i_2 < \cdots < i_k \le n$，满足

$$P(A_{i1}A_{i2}\cdots A_{ik}) = P(A_{i1})P(A_{i2})\cdots P(A_{ik})$$

则称 $A_1, A_2, \cdots, A_n$ 为相互独立的事件。

两两独立：

设 A，B，C 是三事件，如果满足等式

$P(AB) = P(A)P(B)$，$P(BC) = P(B)P(C)$，$P(AC) = P(A)P(C)$，

则称 A，B，C 为两两独立的事件。

例：若 A，B，C 两两独立，并且满足

$$P(ABC) = P(A)P(B)P(C),$$

则称 A，B，C 为相互独立的事件。

（1）两两独立的事件组**不一定**相互独立。
（2）相互独立的事件组**一定**两两独立。

两两独立和**相互独立**之间的关系？

1.6 例子

例：一个盒子中有6个白球、4个黑球，从中不放回地每次任取1个，连取2次，求第二次取到白球的概率。

解：

设事件 A 为第二次取到白球，事件 B 为第一次取到白球，则 $\bar{B}$ 为第一次取到黑球。

因为 AB 与 $A\bar{B}$ 互斥，且 $A = AB \cup A\bar{B}$，所以

$$
\begin{aligned}
P(A) &= P(AB) + P(A\bar{B}) \\
&= P(B)P(A|B) + P(\bar{B})P(A|\bar{B}) \\
&= \frac{6}{10} \times \frac{5}{9} + \frac{4}{10} \times \frac{6}{9} \\
&= \frac{3}{5}
\end{aligned}
$$

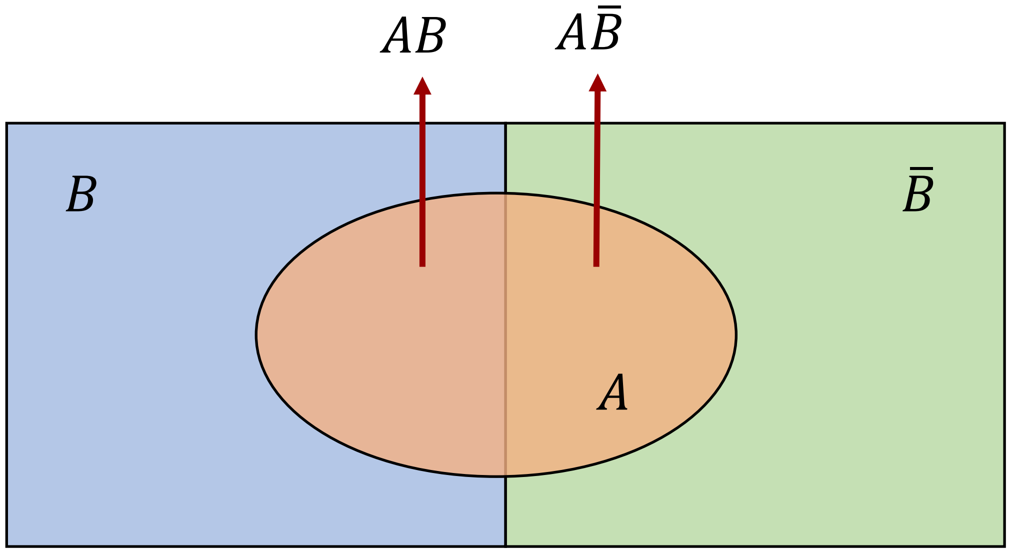
互斥事件的并的概率等于各互斥事件概率的和（概率的加法公理）

概率的乘法公式

概率论基础

1.6 例子

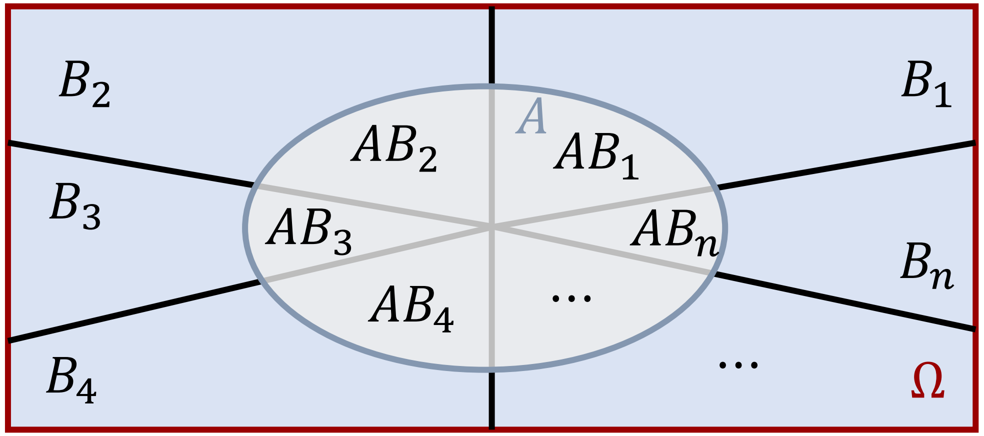
$$
\begin{aligned}
P(A) &= P(AB \cup A\bar{B}) \\
&= P(AB) + P(A\bar{B}) \\
&= P(B)P(A|B) + P(\bar{B})P(A|\bar{B})
\end{aligned}
$$



PEKING UNIVERSITY

1.6 例子

$$P(A)=\sum_{i=1}^{n}P(AB_i)$$



概率论基础

1.6 全概率公式

完备事件组：

设 $B_1, B_2, \cdots, B_n$ 为样本空间 Ω 的一组事件，如果满足

（1）$B_i B_j = \emptyset, \ \ 1 \le i \ne j \le n$

（2）$\bigcup_{i=1}^{n} B_i = \Omega$

则称这组事件为**完备事件组**。

全概率公式：

设 Ω 为随机试验的样本空间，$B_1, B_2, \cdots, B_n$ 为 Ω 中的一个完备事件组，满足 $\forall\, 1 \le i \le n, P(B_i) > 0$，则对 Ω 中的任一事件 A，有

$$P(A) = \sum_{i=1}^{n} P(B_i) P(A|B_i)$$

北京大学 PEKING UNIVERSITY

1.6 全概率公式

全概率公式：

$$P(A)=\sum_{i=1}^{n}P(B_i)P(A|B_i)$$

- $P(B_i)$ 为**先验概率**（Prior Probability）：它由以往经验得到，一般地，它是事件 A 的原因
- 全概率公式是**由因求果**

PEKING UNIVERSITY 北京大学

1.6 全概率公式

例：

一批麦种，其中一、二等品分别占90%和10%，并且它们结出的麦粒数为50以上的概率分别为0.5和0.15，现从中任取一个麦种，求结出的麦粒数为50以上的概率是多少？

解：

设 $A=\{$结出的麦粒数分别为50以上$\}$，

$B_1=\{$取到一等品麦种$\}$，$B_2=\{$取到二等品麦种$\}$，

$P(B_1)=0.9$，$P(B_2)=0.1$，$P(A|B_1)=0.5$，$P(A|B_2)=0.15.$

因此由**全概率公式**，可得

$$\begin{aligned}P(A)&=P(B_1)P(A|B_1)+P(B_2)P(A|B_2)\\&=0.9\times 0.5+0.1\times 0.15=0.465\end{aligned}$$

PEKING UNIVERSITY

1.6 全概率公式

例：

某厂由甲、乙、丙三个车间生产同一种产品，每个车间的产量分别占全厂的25%，35%，40%，各车间产品的次品率分别为5%，4%，2%，求该厂产品的次品率。

解：

设 B_1，B_2，B_3 分别表示产品来自甲、乙、丙车间，A 表示取到次品，

$$
\begin{aligned}
P(A) &= \sum_{i=1}^{3} P(B_i) \cdot P(A|B_i) \\
&= 0.25\times0.05 + 0.35\times0.04 + 0.4\times0.02 \\
&= 0.0345
\end{aligned}
$$

概率论基础

1.7 贝叶斯（Bayes）公式

例：

根据收集到的数据，已知：$P(\text{男性}|\text{糖尿病}) = 0.2$， $P(\text{糖尿病}) = 0.05$， $P(\text{男性}) = 0.5$，如何计算在已知一个人是男性的情况下，他患有糖尿病的概率，即 $P(\text{糖尿病}|\text{男性})$？

解：

$$
\begin{aligned}
P(\text{糖尿病}|\text{男性}) &= \frac{P(\text{男性且患糖尿病})}{P(\text{男性})} \\
&= \frac{P(\text{男性}|\text{糖尿病}) \cdot P(\text{糖尿病})}{P(\text{男性})} \\
&= \frac{0.2 \times 0.05}{0.5} \\
&= 0.02
\end{aligned}
$$

PEKING UNIVERSITY

1.7 贝叶斯（Bayes）公式

Bayes公式：

设 $B_1, B_2, \cdots, B_n$ 构成一组完备事件组，$P(B_i) > 0$，

则对任一事件 A 有：$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{P(A)} = \frac{P(B_i)P(A|B_i)}{\sum_{j=1}^{n} P(B_j)P(A|B_j)}$.

- 称 $P(B_i|A)$ 为**后验概率**（Posterior Probability）
- 已知结果发生，对导致结果发生的因素的可能性大小重新修正
- 贝叶斯（Bayes）公式体现了**执果求因**

1.7 Bayes公式

例：某厂由甲、乙、丙三个车间生产同一种产品，每个车间的产量分别占全厂的25%，35%，40%，各车间产品的次品率分别为5%，4%，2%，现从中任取一件，已知取出的是次品，问该次品为甲车间产品的概率是多大？

解：设 B_1，B_2，B_3 分别表示产品来自车间甲、乙、丙车间，A表示取到次品，则所求概率为 $P(B_1|\mathrm{A})$，由贝叶斯公式得

$$P(B_1|A)=\frac{P(B_1)P(A|B_1)}{P(A)}=\frac{P(B_1)P(A|B_1)}{\sum_{j=1}^{3}P(B_j)P(A|B_j)}$$

$$=\frac{0.25\times 0.05}{0.25\times 0.05+0.35\times 0.04+0.40\times 0.02}\approx 36.23\%$$

1.7 Bayes公式

例：艾滋病普查：使用一种血液试验来检测人体内是否携带艾滋病病毒。设这种试验的假阴性比例为5%（即在携带病毒的人中，有5%的试验结果为阴性），假阳性比例为1%（即在不携带病毒的人中，有1%的试验结果为阳性）。据统计，人群中携带病毒者约占0.1%，若某人的血液试验结果呈阳性，试问该人携带艾滋病病毒的概率。

解：设“携带病毒”为 A “试验呈阳性”为 B，则

$P(A)=0.001$，$P(\bar{A})=0.999$，$P(B|A)=0.95$，$P(B|\bar{A})=0.01$，

$$P(A|B)=\frac{P(AB)}{P(B)}=\frac{P(AB)}{P(AB)+P(\bar{A}B)}=\frac{P(A)P(B|A)}{P(A)P(B|A)+P(\bar{A})P(B|\bar{A})}\approx 0.087$$

概率论基础

1.8 随机变量

随机变量：

设 Ω 是试验 E 的样本空间，随机变量 $X = X(\omega),\ \omega \in \Omega$，是一个单值实函数，将样本空间的样本点映射为实值。随机变量的取值是依靠样本点的概率随机的。随机变量常用大写字母 X, Y, Z 等表示，$\{\omega | X(\omega) \le x\} = \{X \le x\}$

分布函数：

设 X 是一个随机变量，则函数

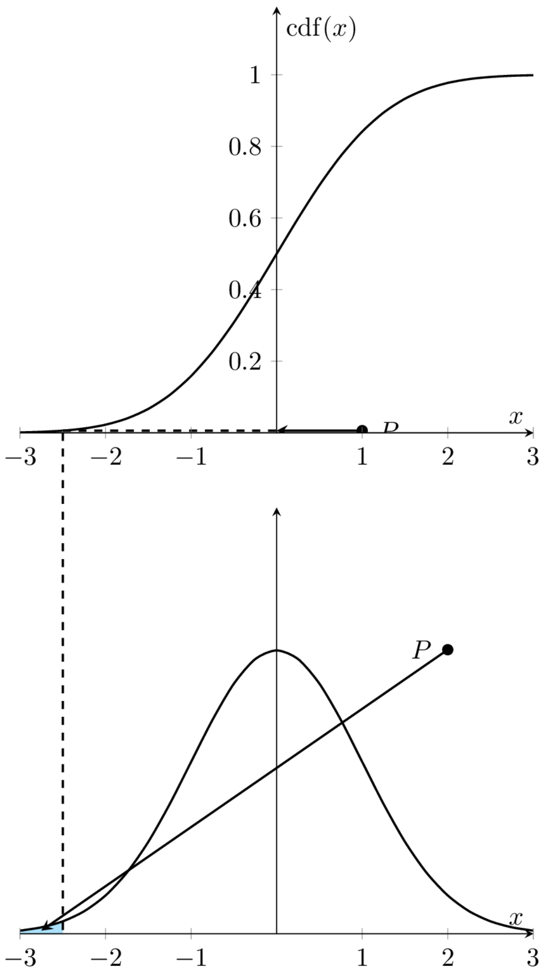
$$F(x) = P\{X \le x\}, \qquad x \in (-\infty, +\infty)$$

称为随机变量 X 的累积分布函数（Cumulative Distribution Function, CDF），简称分布函数。

注：由分布函数的定义有对于任意函数 x_1, x_2 $(x_1 < x_2)$，有

$$P\{x_1 < X \le x_2\} = P\{X \le x_2\} - P\{X \le x_1\} = F(x_2) - F(x_1)$$

分布函数可以完整地描述随机变量取值的概率规律



1.8 随机变量

离散型随机变量：

取值为有限个或者无穷可列个的随机变量。

分布律：

设离散型随机变量 X 的所有可能取值为 x_k，$k=1,2.\cdots$，则称

$$P\{X=x_k\}=p_k, \qquad k=1,2,\cdots$$

为 X 的分布律，也称概率函数。

用分布律表示分布函数：

$$F(x)=P\{X\le x\}=\sum_{x_i\le x}P\{X=x_i\}=\sum_{x_i\le x}p_i$$

1.8 随机变量

一维连续型随机变量：

设随机变量 X 的分布函数为 $F(x)$，若存在非负可积函数 $f(x)$，使对于任意实数 x，有

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

则称 X 为连续型随机变量，并称 $f(x)$ 为 X 的概率密度函数（Probability Density Function, PDF），简称概率密度。

注：连续型随机变量的分布函数 $F(x)$ 必为连续函数。

$$1 = \int_{-\infty}^{+\infty} f(t)dt$$

1.8 随机变量

例：设随机变量 X 的概率密度函数为 $f(x)$，试确定常数 A 以及 X 的分布函数。

$$f(x)=\begin{cases}Ae^{-3x}, & x\geq 0\\ 0, & x<0\end{cases}$$

解：由于 $1=\int_{-\infty}^{+\infty}f(x)dx=\int_{0}^{+\infty}Ae^{-3x}dx=\frac{1}{3}A$，可知 $A=3$，即

$$f(x)=\begin{cases}3e^{-3x}, & x\geq 0\\ 0, & x<0\end{cases}$$

从而 X 的分布函数为 $F(x)=\int_{-\infty}^{x}f(t)dt=\begin{cases}1-e^{-3x}, & x\geq 0\\ 0, & x<0\end{cases}$

PEKING UNIVERSITY

1.9 数学期望

离散型随机变量数学期望：

设离散型随机变量 X 的分布律为 $P(X=x_k)=p_k,\ k=1,2,\cdots$

若级数 $\sum_{k=1}^{+\infty} x_k p_k$ 绝对收敛，则称其为 X 的数学期望，或称为理论均值，记作 $E(X)$，即

$$E(X)=\sum_{k=1}^{+\infty} x_k p_k .$$

注：

（1）上述定义要求级数绝对收敛的目的在于使数学期望唯一，由级数的概念可知，当级数绝对收敛时，可以保证其和不受次序变化的影响。

（2）数学期望是一个确定的数，失去了随机性。

（3）数学期望的“线性”性质： $E(\mathrm{a}X+\mathrm{b}Y)=\mathrm{a}E(X)+\mathrm{b}E(Y)$，a和b是常数

1.9 数学期望

连续型随机变量数学期望：

设连续型随机变量 X 的概率密度函数为 $f(x)$，若广义积分 $\int_{-\infty}^{+\infty} xf(x)dx$ 绝对收敛，则称此积分为 X 的数学期望，记作 $E(X)$，即 $E(X)=\int_{-\infty}^{+\infty} xf(x)dx.$

例：

随机变量 X 的概率密度函数为 $f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geq 0\\ 0, & x<0\end{cases}$，求 $E(X)$.

解：

$$E(X)=\int_{-\infty}^{+\infty} xf(x)dx=\int_{0}^{+\infty} x\lambda e^{-\lambda x}dx=\frac{1}{\lambda}$$

1.9 方差和标准差

方差：

设 X 为随机变量，若 $E\{[X-E(X)]^2\}$ 存在，则称其为 X 的方差，记为 $D(X)$ 或 $Var(X)$，即

$$D(X)=Var(X)=E\{[X-E(X)]^2\}$$

注：

（1）方差的算术平方根 $\sqrt{D(X)}$ 称为 X 的标准差或均方差，记为 $\sigma(X)$.

（2）方差刻画随机变量取值对于其数学期望的平均偏离程度，

若 X 的取值比较集中，则方差 $D(X)$ 较小；

若 X 的取值比较分散，则方差 $D(X)$ 较大。

（3）方差 $D(X)$ 是一个确定的数，失去了随机性。

PEKING UNIVERSITY

1.9 方差和标准差

方差：

$$D(X) = Var(X) = E\{[X - E(X)]^2\}$$

计算：

（1）若 X 为离散型，其概率分布为 $P\{X = k\} = p_k,\ (k = 1, 2, \cdots)$，

$$D(X) = \sum_k [x_k - E(X)]^2 p_k\,.$$

（2）若 X 为连续型随机变量，其概率密度为 $f(x)$，

$$D(X) = \int_{-\infty}^{+\infty} [x - E(x)]^2 \cdot f(x) dx\,.$$

（3）$D(X) = E\{[X - E(X)]^2\} = E\{X^2 - 2XE(X) + [E(X)]^2\}$

$$= E(X^2) - 2E(X)E(X) + [E(X)]^2 = E(X^2) - [E(X)]^2$$

PEKING UNIVERSITY 北京大学

1.10 协方差

协方差：

若随机变量 X 的期望 $E(X)$ 和 Y 的期望 $E(Y)$ 存在，则称

$$Cov(X,Y) = E\{[X - E(X)][Y - E(Y)]\}$$

为 X 与 Y 的协方差。外层期望 E 对 X,Y 的联合分布求积分。

注：由定义得 $Cov(X,X) = D(X)$，$Cov(X,Y) = Cov(Y,X)$.

计算： $\boldsymbol{Cov(X,Y) = E(XY) - E(X)E(Y)}$

推导：

$$\begin{aligned} Cov(X,Y) &= E\{[X - E(X)][Y - E(Y)]\} \\ &= E[XY - E(X)Y - E(Y)X + E(X)E(Y)] \\ &= E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y) \end{aligned}$$

概率论基础

1.11 相关系数

相关系数：

若随机变量 X，Y的方差和协方差均存在，

且 $D(X)>0$，$D(Y)>0$，则

$$\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

称为 X，Y 的相关系数。

相关系数的性质：

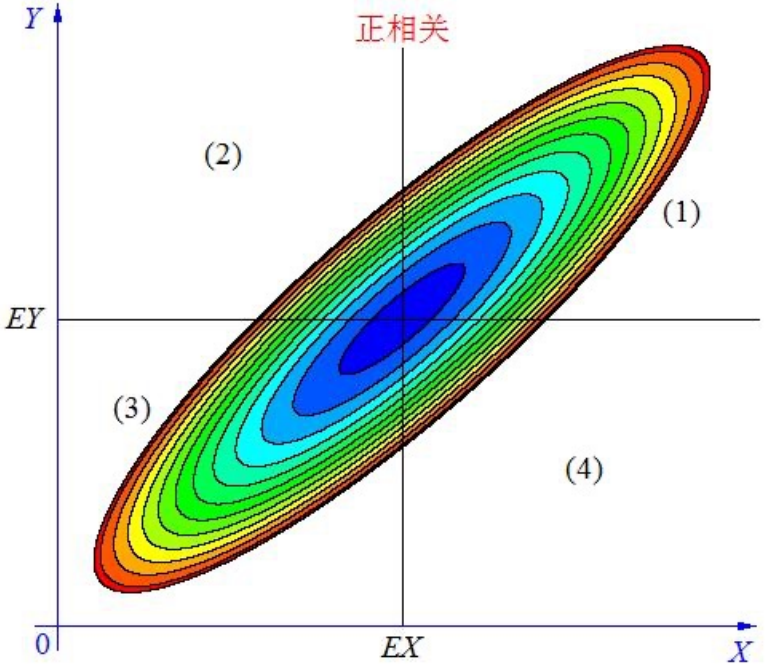
（1）相关系数反映随机变量之间的线性相关程度。

（2）若 X 与 Y 相互独立，则 $\rho_{XY}=0.$

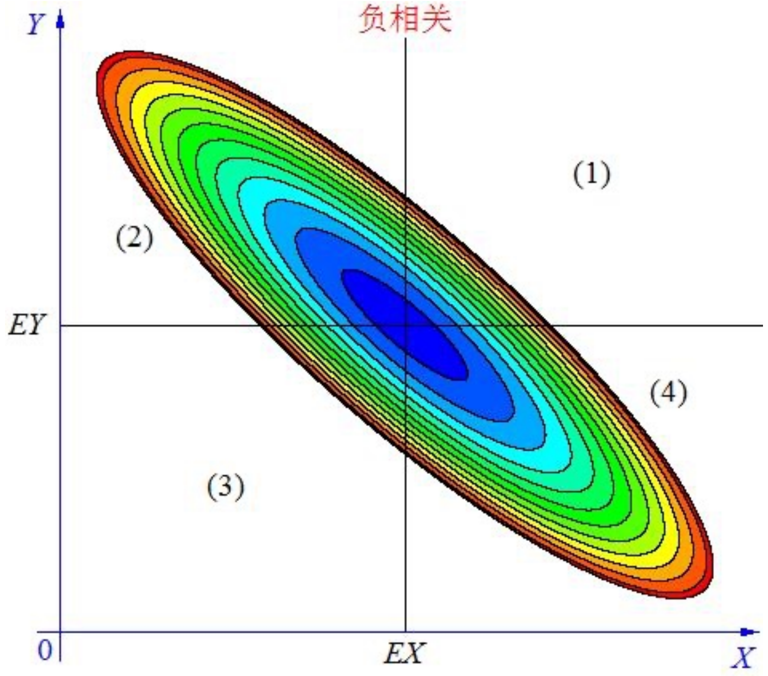
（3）$|\rho_{XY}|\le 1.$

（4）$|\rho_{XY}|=1 \iff \exists$ 常数 a,b，使得 $P\{Y=a+bX\}=1.$

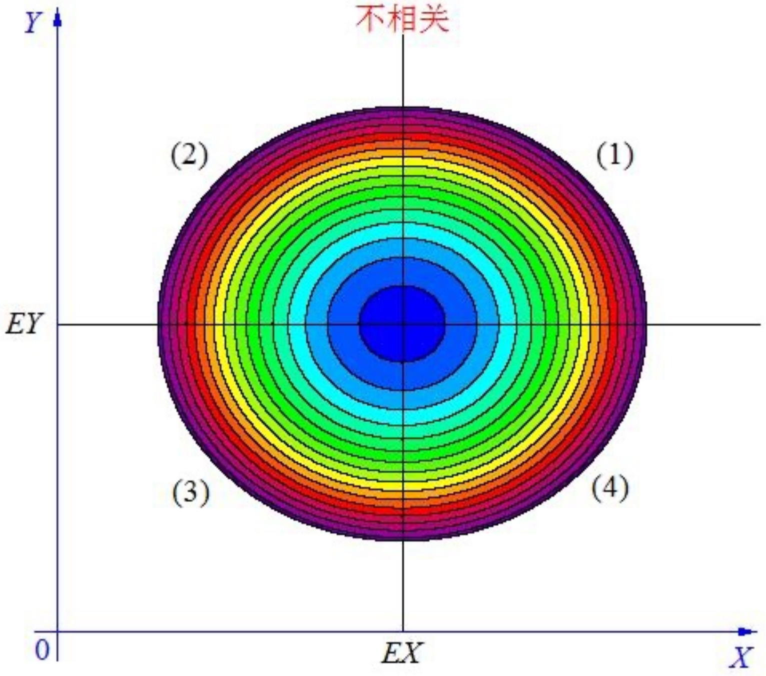
1.11 协方差和相关系数可视化



(a) *X* 与 *Y* 正相关



(b) *X* 与 *Y* 负相关



(c) *X* 与 *Y* 不相关

1.12 离散型随机变量示例

例：将扔一个骰子所得结果记为随机变量 X，另一个随机变量 $Y = X + 1$，求X和Y期望、方差、协方差、相关系数。

解：首先，写出两个变量的分布列，分别求 X 和 Y 的期望

$$\mathrm{E}(X) = 1 * \frac{1}{6} + 2 * \frac{1}{6} + 3 * \frac{1}{6} + 4 * \frac{1}{6} + 5 * \frac{1}{6} + 6 * \frac{1}{6} = \frac{7}{2}$$

$$\mathrm{E}(Y) = 2 * \frac{1}{6} + 3 * \frac{1}{6} + 4 * \frac{1}{6} + 5 * \frac{1}{6} + 6 * \frac{1}{6} + 7 * \frac{1}{6} = \frac{9}{2}$$

X	1	2	3	4	5	6
Y	2	3	4	5	6	7
$P(X)$	1/6	1/6	1/6	1/6	1/6	1/6

PEKING UNIVERSITY

1.12 离散型随机变量示例

例：将扔一个骰子所得结果记为随机变量 X，另一个随机变量 $Y = X + 1$ ，求 X 和 Y 期望、方差、协方差、相关系数。

解：根据公式求 X 和 Y 的方差

$$\mathrm{D}(X) = \frac{25}{4} * \frac{1}{6} + \frac{9}{4} * \frac{1}{6} + \frac{1}{4} * \frac{1}{6} + \frac{1}{4} * \frac{1}{6} + \frac{9}{4} * \frac{1}{6} + \frac{25}{4} * \frac{1}{6} = \frac{35}{12}$$

同样地，可计算得

$$\mathrm{D}(Y) = \frac{35}{12}$$

X	1	2	3	4	5	6
$X - E(X)$	$-5/2$	$-3/2$	$-1/2$	$1/2$	$3/2$	$5/2$
$[X-E(X)]^2$	$25/4$	$9/4$	$1/4$	$1/4$	$9/4$	$25/4$
$P(X)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

PEKING UNIVERSITY

1.12 离散型随机变量示例

例：将扔一个骰子所得结果记为随机变量 X，另一个随机变量 $Y = X + 1$ ，求 X 和 Y 期望、方差、协方差、相关系数。

解：最后求协方差与相关系数，将 X 和 Y 的每个取值对应相乘得到XY的分布列，$E(XY) = \frac{56}{3}$

协方差：

$$Cov(X,Y) = E(XY) - E(X)E(Y) = \frac{35}{12}$$

相关系数：

$$\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}} = 1$$

X	1	2	3	4	5	6
Y	2	3	4	5	6	7
XY	2	6	12	20	30	42
$P(XY)$	1/6	1/6	1/6	1/6	1/6	1/6

概率论小结

- 样本空间与样本点
- 随机事件及其关系
- 古典概率、条件概率
- 事件的独立性、全概率公式、贝叶斯公式
- 随机变量
- 数学期望、方差和标准差
- 协方差、相关系数

谢
谢

北京大學
PEKING UNIVERSITY