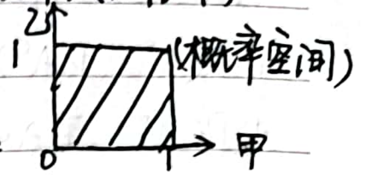
概率模型

随机试验 trial　　样本空间 Ω (outcome)　　信念分布 belief

样本空间的试验结果必须满足互斥且完整.(可能是有限个或无限个)

· 样本空间(离散/连续):



例 连续模型. 甲和乙约定见面, 每个人随机迟到 0-1 小时.

分配概率:

(定义) 事件: 样本空间的子集, 某些试验结果的集合, 可以用"是否发生"描述.

若试验结果记为 ω, 事件记为 A: $\omega\in A$ 表示试验中事件发生, $\omega\notin A$ 则不发生.

概率公理 1°(非负性) 对任意事件 A, $P(A)\geq 0$.

用于验证是否为概率模型　2°(归一化) $P(\Omega)=1$.

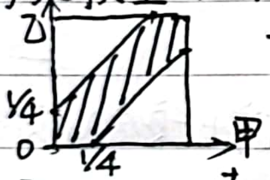
3°(可加性) 若 $A\cap B=\emptyset$, 那么 $P(A\cup B)=P(A)+P(B)$.

证明 $P(A)\leq 1$: $A\cup A^c=1$, $A\cap A^c=\emptyset$. $P(A)+P(A^c)=1$ 且 $P(A)\geq 0$, $P(A)\leq 1$.

补充: $P(\{s_1,\dots s_k\}) = P(\{s_1\})+\dots+P(\{s_k\}) = P(s_1)+\dots+P(s_k)$. (有限可加)

概率的实质: 为事件确定一个非负数 $P(A)$, 相当于一个映射.

例 约会模型进阶: 若甲、乙最多等待对方15分钟, 则他们能相遇的概率多大?



阴影部分表示有概率相遇. $P(\text{相遇}) = \frac{7}{16}$.

(悖论原因: 没有明确"等可能"概念)

Bertrand 悖论: 在半径为1的圆内任取一条弦, 求弦长大于 $\sqrt{3}$ 的概率.

如果认为弦的端点等可能落在圆周上: 固定一个端点, 从角度考虑 $P=1/3$.

如果认为弦中点等可能地落在与之垂直的直径上: $P=1/2$.

如果认为弦中点等可能落在圆内, 考虑面积: $P=1/4$.

公理3 从有限可加推广到可列可加: 若 $A_1, A_2, \dots$ 互不相交, $P(A_1\cup A_2\cup\dots)=P(A_1)+P(A_2)+\dots$

可列的实质: 与自然数集一一对应, 可数无穷.

概率性质 (只能使用推广后的3条公理)

性质1. $P(\emptyset)=0$.　证明: $\Omega=\Omega+\emptyset+\emptyset+\dots$, $P(\Omega)=P(\Omega)+P(\emptyset)+\dots$,

$0=P(\emptyset)+P(\emptyset)+\dots$, 则 $P(\emptyset)=0$.

Date. /

性质2° 有限可加性 事件 A_i, $i=1,\dots,n$, 且 $A_i\cap A_j=\phi$, $i\neq j$, 则 $P(\bigcup_{i=1}^{n}A_i)=\sum_{i=1}^{n}P(A_i)$.

(证明) 令 $A_{n+1}=A_{n+2}=\dots=\phi$, 则 $\bigcup_{i=1}^{n}A_i=\bigcup_{i=1}^{\infty}A_i$, 且 $i\neq j$ 时 $A_i\cap A_j=\phi$,

由公理3和性质1, $P(\bigcup_{i=1}^{n}A_i)=P(\bigcup_{i=1}^{\infty}A_i)=\sum_{i=1}^{\infty}P(A_i)=\sum_{i=1}^{n}P(A_i)$.

性质3° $P(A^c)=1-P(A)$. $1=P(\Omega)=P(A\cup A^c)=P(A)+P(A^c)$.

性质4° 若 $A\subset B$, 则 $P(B\setminus A)=P(B)-P(A)$.

证明: $B=A\cup(B\setminus A)$ 且 $A\cap(B\setminus A)=\phi$. 用性质2°可得.

推论: (概率单调性) 若 $A\subset B$, 则 $P(A)\le P(B)$; 对任意事件, $0\le P(A)\le 1$.

性质5° $P(A\cup B)=P(A)+P(B)-P(A\cap B)$.

证明: $P(A\cup B)=P(A)+P(B\setminus A)=P(A)+P(B\setminus(A\cap B))=P(A)+P(B)-P(A\cap B)$.

推论: 有限次可加性 finite subaddictivity / Boole's inequity: $P(\bigcup_{i=1}^{n}A_i)\le\sum_{i=1}^{n}P(A_i)$

性质6° the inclusion-exclusion formula 容斥恒等式

$$P(\bigcup_{i=1}^{n}A_i)=\sum_{1\le i\le n}P(A_i)-\sum_{1\le i<j\le n}P(A_i\cap A_j)+\sum_{1\le i<j<k\le n}P(A_i\cap A_j\cap A_k)+\dots(-1)^{n-1}P(A_1\cap A_2\dots\cap A_n).$$

证明: (利用互斥. 归纳法) 记 $P_k=\sum_{1\le j_1<\dots<j_k\le n}P(A_{j_1}\dots A_{j_k})$, $P(\bigcup_{i=1}^{n}A_i)=\sum_{k=1}^{n}(-1)^{k-1}P_k$.

归纳法: $k=n-1 \Rightarrow$ 时 $P(\bigcup_{i=1}^{n}A_i)=P(\bigcup_{i=1}^{n-1}A_i)+P(A_n)-P(\cap(\bigcup_{i=1}^{n-1}A_i\cdot A_n))$

其中 $(\bigcup_{i=1}^{n-1}A_i)\cap A_n=\bigcup_{i=1}^{n-1}(A_i\cap A_n)=\bigcup_{i=1}^{n-1}B_i$. (交集变并集)

$q_k=\sum_{1\le j_1<\dots<j_k\le n-1}P(B_{j_1}\dots B_{j_k})=\sum_{1\le j_1<\dots<j_k\le n-1}P(A_{j_1}\dots A_{j_k}A_n)$. (替换形式)

$P(\bigcup_{i=1}^{n}A_i)=\sum_{k=1}^{n-1}(-1)^{k-1}P'_k+\sum_{k=0}^{n-1}(-1)^{k}q_k$. 相当于: 前 $n-1$ 项的概率 + 一定含 A_n 项的概率

$\to P_k=P'_k+q_{k-1}$. $\quad =\sum_{k=1}^{n}(-1)^{k-1}q_{k-1}$.

$\to =\sum_{k=1}^{n-1}(-1)^{k-1}(P'_k+q_{k-1})+\underbrace{(-1)^{n-1}P_n}_{(-1)^{n-1}q_{n-1}}=\sum_{k=1}^{n-1}(-1)^{k-1}P_k+(-1)^{n-1}P_n$.

性质7° (Bonferroni's inequity) $P(\bigcup_{i=1}^{n}A_i)\ge\sum_{1\le i\le n}P(A_i)-\sum_{1\le i<j\le n}P(A_i\cap A_j)$

(Kounias inequity) $P(\bigcup_{i=1}^{n}A_i)\le\min_{k}\{\sum_{1\le i\le n}P(A_i)-\sum_{i:i\neq k}P(A_i\cap A_k)\}$.

性质8° (可列次可加性 σ-subaddictivity) $P(\bigcup_{i=1}^{\infty}A_i)\le\sum_{i=1}^{\infty}P(A_i)$

证明: 令 $B_1=A_1$, $B_n=A_n\setminus\bigcup_{i=1}^{n-1}A_i$, 则 $\bigcup_{i=1}^{\infty}A_i=\bigcup_{i=1}^{\infty}B_i$ 且 $B_i\cap B_j=\phi$, $i\neq j$;

由公理(3)和单调性, $P(\bigcup_{i=1}^{\infty}A_i)=P(\bigcup_{i=1}^{\infty}B_i)=\sum_{i=1}^{\infty}P(B_i)\le\sum_{i=1}^{\infty}P(A_i)$.

性质9° $P(\bigcap_{i=1}^{\infty}A_i)\ge1-\sum_{i=1}^{\infty}P(A_i^c)$, 特别地, $P(A_1\cap A_2)\ge1-P(A_1^c)-P(A_2^c)$.

证明: $P(\bigcap_{i=1}^{\infty}A_i)=1-P((\bigcap_{i=1}^{\infty}A_i)^c)=1-P(\bigcup_{i=1}^{\infty}A_i^c)\ge1-\sum_{i=1}^{\infty}P(A_i^c)$,

令 $A_3=A_4=\dots\Omega$ 即可得第2个结论.

处理证明3大技巧：互斥，求补，归纳法　　问题解决思路 $\begin{cases}\text{序贯树}\\\text{递归+边界条件}\end{cases}$

匹配问题补充

例 翻牌匹配问题（古典概型） 一共有n张牌，n个数，依次翻开，有一张匹配即可.

$P(A_1)=\frac{1}{n}$（第一张匹配） $P(A_1A_2)=\frac{1}{n(n-1)}$（第1、2张均匹配）…

由对称性，有两张匹配的概率为 $P(A_jA_k)=\frac{1}{n(n-1)}$.

$P(A_1\cup A_2\cdots\cup A_n)=\frac{1}{n}C_n^1-\frac{1}{n(n-1)}C_n^2+\frac{1}{n(n-1)(n-2)}C_n^3+\cdots=1-\frac{1}{2!}+\frac{1}{3!}-\frac{1}{4!}+\cdots\frac{(-1)^{n-1}}{n!}$

实际上是 $1-e^{-1}$ 的Taylor展开. （收敛到定值）

随机抽样与随机分配

性质：有放回/无放回，有序/无序. 含有M个~~箱子~~球的箱子里抽n次.

有序对

· 放回有序：$\Omega=\{w:w=(a_1,\dots a_n),a_i\in\{1,2,\dots M\}\}$，$\#(\Omega)=M^n$.

放回无序：$\Omega=\{w:w=[a_1,\dots a_n],a_i\in\{1,2,\dots M\}\}$，$\#(\Omega)=C_{M+n-1}^{n}=\binom{M+n-1}{n}$.

不放回有序：$\Omega=\{w:w=(a_1,\dots a_n),a_k\neq a_l,k\neq l,a_i\in\{1,2,\dots M\}\}$，$\#(\Omega)=P_M^n=\frac{M!}{(M-n)!}$

不放回无序：$\Omega=\{w:w=[a_1,\dots a_n],a_i\in\{1,2,\dots M\}\}$ $(a_k\neq a_l,k\neq l)$ $\#(\Omega)=\binom{M+n-1}{n}$.

随机抽样与随机分配的对应：相当于按箱分配球，有序⇔球不同，无序⇔相同；

放回抽样⇔可以一箱多球，不放回抽样⇔每箱最多一球

古典概型补充说明

证明结论 $\binom{m}{n}=\binom{m-1}{n-1}+\binom{m-1}{n}$.

方法1：$\binom{m}{n}=\sum_{k=n-1}^{m-1}\binom{k}{n-1}$，可以证明有放回无序抽样的结论 $\binom{M+n-1}{n}$.

M个球里抽n个（有放回无序）：假设在M-1个球时成立，抽任意n'个球有 $\binom{M+n'-2}{n'}$.

情景理解：在前M-1个球中抽j次 $\binom{j+M-2}{j}$，抽第M个球(n-j)次.

$\sum_{j=0}^{n}\binom{j+M-2}{j}=\sum_{j=0}^{n}\binom{j+M-2}{M-2}$. 接下来换元 $\begin{cases}M-2=\hat{n}-1\\k=j+\hat{n}-1\end{cases}$，原式 $=\sum_{k=\hat{n}-1}^{n+\hat{n}-1}\binom{k}{\hat{n}-1}=\binom{\hat{m}}{\hat{n}}$.

记 $\hat{m}=n+\hat{n}$，那么原式 $=\binom{n+M-1}{M-1}$ 得证.

例：m个球r种颜色，同色球相同（每种球分别有 $n_1,n_2,\dots n_r$ 个）求m个球的排列数.

（等价地，把m个球放入r个盒子，使 m_j 个球进入第j个盒子）

结论：$(m!)/(n_1!n_2!\dots n_r!)$

古典概型技巧：适度借助常识；用最简单情形验证；将（所有）物体编号；

情景证明(story proof)

Date.　　　/

有限可加→可列可加，但不能推广到不可列无限可加.

否则：以在区域内找落点为例，$1=P(\Omega)=P(\bigcup_{x,y}\{(x,y)\})=\sum_{x,y}P(x,y)=0$.

声明：$P(A)=1$ 只能说"几乎必然"发生，而非绝对一定（比如 $x=0$ 在 $\mathbb{R}$ 轴上）

· 不是任意样本空间的子集都可以分配概率.

$\forall x,y\in[0,1]$，定义等价关系 $x\sim y$ 当且仅当 $x-y\in\mathbb{Q}$，将区间 $[0,1]$ 分割为了一些互不相交的等价类. 存在 $[0,1]$ 的一个子集 E，由每个等价类中选取一个元素构成.

记 $\{r_i\}=[-1,1]\cap\mathbb{Q}$，令 $E_n=r_n+E$（E 的每个元素加 r_n）. $m\neq n$ 时 $E_m\cap E_n=\phi$，因为若 $\exists x\in E_m\cap E_n$，$x-r_m\in E$，$x-r_n\in E$，$(x-r_m)-(x-r_n)\in\mathbb{Q}$，不符合 E 构造. 又 $[0,1]\subset\bigcup_{n=1}^{\infty}E_n\subset$

（平移不改变概率）

假设 $P(E)$ 存在，则 $P(E_n)=P(E)$，那么 $0<P([0,1])\leq P(\bigcup_{n=1}^{\infty}E_n)\leq P([-1,2])<\infty$.

而由于 $\{E_n\}$ 互斥，$P(\bigcup_{n=1}^{\infty}E_n)=\sum_{n=1}^{\infty}P(E)$，$P(E)=0$，矛盾. 故 E 不可分配概率.

· 已知：有限范围内 $+\cup\cap$ 运算结果可以度量，现在希望推广到可列.

将可列转化到有限次交并补运算：$B_n=A_n-(A_1+A_2+\dots A_{n-1})=A_n\cap(A_1\cup\dots A_{n-1})^c$

那么可列并 $\bigcup_{i=1}^{\infty}A_i=\bigcup_{i=1}^{\infty}B_i$，可列并可以合理分配.

可列交可由可列并取补得到，不改变其能合理分配性质.

概率空间

定义事件域 $\mathcal{F}$，表示 Ω 的某些子集的集合. 若 $\mathcal{F}$ 满足以下3个条件：

$1°\ \Omega\in\mathcal{F}$. $2°$ 若 $A\in\mathcal{F}$ 则 $A^c\in\mathcal{F}$. $3°$ 若 $A_n\in\mathcal{F}$，$n=1,2,\dots$ 则 $\bigcup_{n=1}^{\infty}A_n\in\mathcal{F}$.

则称 $\mathcal{F}$ 是 Ω 上的事件域或 σ-域或 σ-代数，$\mathcal{F}$ 中的元素为事件，

称 $(\Omega,\mathcal{F})$ 是可测空间（measurable space）.

σ-代数例：$\mathcal{F}=\{\Omega,\phi\}$，平凡的 σ-代数；$\mathcal{F}=\{\Omega$ 的所有子集$\}$，最大的 σ-代数

$\mathcal{F}=\{\Omega,\phi,A,A^c\}$ 是含 A 的最小 σ-代数，记作 $\mathcal{F}=\sigma(\{A\})$.

注意：Ω 的子集未必是事件，只有 $\mathcal{F}$ 中元素为事件；$\mathcal{F}$ 对集合的各种运算都封闭，包括极限

例 若 A,B 是 Ω 的两个子集且 $B\neq A^c$，则由 $\mathcal{A}=\{A,B\}$ 生成的 σ-域 $\mathcal{F}$ 是 $\{\Omega,\phi,A,A^c,B,B^c,AB,AB^c,A^cB,A^cB^c,A\cup B,A^c\cup B,A\cup B^c,A^c\cup B^c,AB\cup A^cB^c,AB^c\cup A^cB\}$.

封闭性补充：$\mathcal{F},\mathcal{G}$ 为两个 σ-域，但 $\mathcal{F}\cup\mathcal{G}$ 不一定是 σ 域.

若 $\mathcal{F}$ 和 $\mathcal{G}$ 是两个 σ-域，且 $\mathcal{F}\subset\mathcal{G}$ 则 $\mathcal{F}\cup\mathcal{G}$ 是 σ-域.

(二元 → 三元)

定义：概率或概率测度(probability measure) $(\Omega,\mathcal{F})$ 是可测空间，$\mathbb{P}$ 是 $\mathcal{F}$ 上函数．
若满足 1°(非负性) $\forall A\in\mathcal{F}$，$\mathbb{P}(A)\geq 0$． 2° $\mathbb{P}(\Omega)=1$ (完全性)
3°(可列可加性) 对 $\mathcal{F}$ 中互不相交(disjoint，互斥)的事件列，$\mathbb{P}(\bigcup_{n=1}^{\infty}A_n)=\sum_{n=1}^{\infty}\mathbb{P}(A_n)$．
称 $\mathbb{P}$ 为 $\mathcal{F}$ 上的概率测度/概率，$(\Omega,\mathcal{F},\mathbb{P})$ 为概率空间．
说明：当 Ω 有限或可列，存在概率分配使 Ω 的任意子集都可测；若 Ω 不可列，则
其一定存在不可测子集，不能被当作事件．但即使事件存在，也不一定能计算概率．
规范表述：若 $\mathbb{P}(A)=1$，则 A 以概率1发生或几乎处处发生，也叫几乎必然，记作 a.s(almost surely)．

2]. 事件域 $\mathcal{F}$ / 可测空间 $(\Omega,\mathcal{F})$ ≠ 测度空间 $(\Omega,\mathcal{F},P)$：
事件域包含全部我们关心的事件，且每一个事件都可以分配概率；对同一个可测空间/事件域，可以构造不同的概率空间．概率计算较困难时可以用上下界逼近．

概率的连续性

若 $A_1\subset A_2\subset A_3\ldots$ 称事件列 $\{A_1,\ldots\}$ 单调递增；$A_1\supset A_2\supset A_3\ldots$ 称单调递减 (单调序列)

单调有界 $[0,1]$ ⇒ 有极限．单增 $\lim_{n\to\infty}A_n=\bigcup_{i=1}^{\infty}A_i$，单减 $\lim_{n\to\infty}A_n=\bigcap_{i=1}^{\infty}A_i$．

定理：对单调事件列 $\{A_n\}$，无论单调增还是单调减，$P(\lim_{n\to\infty}A_n)=\lim_{n\to\infty}P(A_n)$．

证明：$\{A_i\}$ 单调增：$\bigcup_{i=1}^{\infty}A_i=A_1\cup\bigcup_{i=2}^{\infty}(A_i-A_{i-1})$ (拆成互斥)

$\mathbb{P}(\lim_{n\to\infty}A_n)=\mathbb{P}(A_1)+\sum_{i=2}^{\infty}(\mathbb{P}(A_i)-\mathbb{P}(A_{i-1}))=\mathbb{P}(A_1)+\lim_{n\to\infty}\sum_{i=2}^{n}(\mathbb{P}(A_i)-\mathbb{P}(A_{i-1}))=\lim_{n\to\infty}\mathbb{P}(A_n)$．

$\{A_i\}$ 单调减：基本同上，拆成互斥后有其补集单调增．

$1-\mathbb{P}(\lim_{n\to\infty}A_n)=\mathbb{P}(\lim_{n\to\infty}(\Omega-A_n))=1-\lim_{n\to\infty}\mathbb{P}(A_n)$，$\mathbb{P}(\lim_{n\to\infty}A_n)=\lim_{n\to\infty}\mathbb{P}(A_n)$．

(定义) 事件列的上极限与下极限　上极限 $\limsup_{n\to\infty}A_n=\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k=\{\omega\in\Omega:\omega$ 属于无穷多个 $A_k\}$

(理解) $\forall n$，$\exists k\geq n$ s.t. $\omega\in A_k$．下极限 $\liminf_{n\to\infty}A_n=\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}A_k$．

(理解) $\exists n_0$，$\forall k\geq n_0$，$\omega\in A_k$．即 ω 属于所有 A_k (无穷多)，除了有限个之外．

例：考虑一种情况：5个人共同上课，$\omega_1=$ V V V ...，$\omega_2=$ O O V V V ...，
$\omega_3=$ V O V V ...，$\omega_4=$ V V O O O ...，$\omega_5=$ V O V O V O ...

那么上极限为 $\{\omega_1,\omega_2,\omega_3,\omega_5\}$，下极限为 $\{\omega_1,\omega_2,\omega_3\}$．
下极限可以看作一个核，其范围小于上极限，上极限相对松散．
$\limsup A_n$ 和 $\liminf A_n$ 都存在；$\lim A_n$ 不一定存在，必须要求上下极限相等．
补性质：$\limsup A_n+\liminf A_n^c=\Omega$．

Date. /

定理：任意事件列$\{A_i\}$，$P(\limsup_{n\to\infty} A_n) \geq \limsup_{n\to\infty} P(A_n)$，$P(\liminf_{n\to\infty} A_n) \leq \liminf_{n\to\infty} P(A_n)$.
($\cap$+单减) (有界)

(证明) 记$B_n = \bigcup_{k=n}^{\infty} A_k$，则$\{B_n\}$是单调递减的. 有$P(\limsup_{n\to\infty} A_n) = P(\lim_{n\to\infty} B_n) = \lim_{n\to\infty} P(B_n)$.

$P(B_n) \geq P(A_k)$，$\forall k \geq n$，即$P(B_n) \geq \sup_{k\geq n} P(A_k)$. $\limsup_{n\to\infty} A_n = \lim_{n\to\infty} \bigcap_{n=1} \bigcup_{k=n} A_k = \bigcap_{n=1}^{\infty} B_n = \lim_{n\to\infty} B_n$

此时$\sup_{k\geq n}$是关于n单调的数列，不是事件列. 单调性保证了极限存在.

定理：Borel-Cantelli引理(BC引理)：若$\sum_{j=1}^{\infty} P(A_n) < \infty$，则$P(\limsup_{n\to\infty} A_n) = 0$.

2° 若$\{A_i\}$互独立且$\sum_{j=1}^{\infty} P(A_j) = \infty$，则$P(\limsup_{n\to\infty} A_n) = 1$.

试验：抛掷一枚均匀硬币无穷多次，记H_n表示第n次抛硬币得到正面.

$P(H_n) = \frac{1}{2}$，$P(\limsup_{n\to\infty} H_n) = 1$，即出现无穷多次正面的概率为1.

而$P(\liminf H_n) = 1 - P(\limsup H_n^c) = 0$，即几乎不可能某时刻后一直为正面.

·可以证明：$\liminf A_n + \limsup A_n^c = \Omega = \liminf A_n^c + \limsup A_n$.

条件概率

(定义) 概率空间$(\Omega, \mathcal{F}, P)$，$A, B \in \mathcal{F}$且$P(A) > 0$，则在A发生的条件下，B发生的条件概率为$P(B|A) = \frac{P(A\cap B)}{P(A)}$.

$P(A) = 0$时相应的条件概率无定义，有时可根据需要再定义$P(B|A)$.

定理 (对应公理 非负、归一、可加) $A \in \mathcal{F}$且$P(A) > 0$：

1° $\forall B \in \mathcal{F}$，$P(B|A) \geq 0$.　　2° $P(\Omega|A) = 1$.

3° 对互斥事件列$\{B_i\}$，$P(\bigcup_{i=1}^{\infty} B_i | A) = \sum_{i=1}^{\infty} P(B_i|A)$.

用$P_A(\cdot)$表示A发生的条件下的条件概率，则$(\Omega, \mathcal{F}, P_A)$也是一个概率空间. 若概率完全集中在A上，也可以将A看成新的样本空间$(A, A\cap\mathcal{F}, P_A)$. 所有概率性质对条件概率均适用.

条件概率依然是一个概率测度 ⇒ 概率空间.

· 可以改写记号 $P_A(\cup B_i) = \sum_{i=1} \frac{P(B_i\cap A)}{P(A)} = \sum P_A(B_i)$.

大小关系：$P(B|A)$和$P(B)$的大小关系不确定. ($B\subset A$时$P(B)$小；互斥时$P(B|A)=0$)

· 乘法公式 $P(\bigcap_{i=1}^{n} A_i) = P(A_1) \prod_{i=2}^{n} P(A_i | A_1 \cap A_2 \ldots \cap A_{i-1})$.

公式可以任意换序，所以其实是n!个公式，视情况选用.

例：写信问题，n封信随机装入n个写有不同地址的信封，考虑地址正确概率.

记A_i表示第i个信封上所写地址正确，$\bigcup_{i=1}^{n} A_i$表示至少有一个正确.

$q_0=1-P(\bigcup_{i=1}^{n}A_i)$ 表示没有一个正确. 对 $\forall i\le n$, $P(A_i)=\frac{1}{n}=\frac{(n-1)!}{n!}$.

$q_{0,n}\to n$ 个均不匹配. 现在想求 $P(\bigcup_{i=1}^{n}A_i)$: $\forall 1\le i<j\le n$, $P(A_iA_j)=\frac{1}{n(n-1)}=\frac{(n-2)!}{n!}$,

$\forall 1\le i<j<k\le n$, $P(A_iA_jA_k)=\frac{(n-3)!}{n!}$ …以此类推, $P(\bigcup_{i=1}^{n}A_i)=1-\frac{1}{2!}+\frac{1}{3!}-\frac{1}{4!}\dots+\frac{(-1)^{n-1}}{n!}$

$q_0=\frac{1}{2!}-\frac{1}{3!}+\dots\frac{(-1)^n}{n!}=\sum_{k=2}^{n}\frac{(-1)^k}{k!}$, 记为 $q_0(n)$. $\lim_{n\to\infty}q_0(n)=\frac{1}{e}=0.37$.

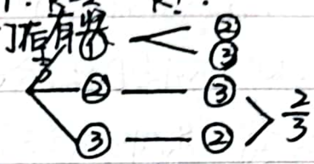
考虑恰好有 r 封匹配(即有 $n-r$ 封不匹配, $q_{0,n-r}$)的概率. 不妨先设正确的就是前 r 封,

事件相当于"从 n 个不同编号的球中无放回地摸 r 个恰好依次摸得前 r 号": 前 r 封 $\frac{1}{n(n-1)\dots(n-r+1)}$

后 $n-r$ 封 $q_{0(n-r)}=\sum_{k=2}^{n-r}\frac{(-1)^k}{k!}$; 从 n 个信封里取 r 个组合有 $\binom{n}{r}$ 种取法, 故所求概率应为

$q_r(n)=\binom{n}{r}\cdot\frac{1}{n(n-1)\dots(n-r+1)}\sum_{k=2}^{n-r}\frac{(-1)^k}{k!}=\frac{1}{r!}\sum_{k=2}^{n-r}\frac{(-1)^k}{k!}$.

·用序贯树解决 Monty Hall 问题:



不妨设选择门①

换前: 1/3 换后: 2/3.

全概率公式: 将 Ω 分割成一组 $\{A_i\}$ (互斥), 则 $P(B)=P(A_1)P(B|A_1)+\dots P(A_n)P(B|A_n)$.

特别地, $P(B)=P(A)P(B|A)+P(A^c)P(B|A^c)$, 有限并可换成可列并.

例: 袋中有 n 个白球, m 个黑球, 无放回连抽 k 个, 问第 k 次为黑的概率. $(\frac{m}{m+n})$

设 A_k 表示第 k 次抽白; 用归纳法: 假设前 A_{k-1} 步都不改变黑球概率.

$P(A_k)=P(A_1)P(A_k|A_1)+P(A_1^c)P(A_k|A_1^c)=\frac{m-1}{m+n-1}\cdot\frac{m}{m+n}+\frac{m}{m+n-1}\cdot\frac{n}{m+n}=\frac{m}{m+n}$.

(例: 敏感问题调查)

Bayes 准则 (可用于推理, 先验概率, 后验概率)

定理 概率空间 $(\Omega,\mathcal{F},P)$, $\{A_i\}$ 是 Ω 的一个分割, 则 $P(A_i|B)=\frac{P(A_i)P(B|A_i)}{\sum_{j=1}^{n}P(A_j)P(B|A_j)}$.

特别地, $P(B)>0$ 时, $P(A|B)=\frac{P(A)P(B|A)}{P(A)P(B|A)+P(A^c)P(B|A^c)}$.

信息的先后顺序不影响结果, 即 Bayes 准则具有一致性.

理解: 已知 $P(B|A_i)$, $\forall i$, 求 $P(A_i|B)$. (在 B 发生的情况下修正"信念程度")

某一类敏感问题在人群中占比为 p, 现在袋中放入占比为 p_0 的红球和 $q_0=1-p_0$ 的白球, 被调查者取到红球就讲真话, 取到白球就讲假话, 仅回答是/否, 抽到球的颜色也是保密的. 若每个对象应回答"是"的概率为 p_1, 求 p_0.

$p_1=q_0+(p_0-q_0)p\Rightarrow p=\frac{p_1-q_0}{p_0-q_0}$. 实际问题中 p_1 未知, 可以由调查结果近似估计得到.

$|p_0-q_0|$ 越大, 所得结论越可靠, 但也越不易被调查对象接受.

Date.　　/

事件的独立性

定义：独立性(independence) 若概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 内，$A, B \in \mathcal{F}$，若 $\mathbb{P}(A\cap B)=\mathbb{P}(A)\mathbb{P}(B)$，则称A与B相互独立.（或 $\mathbb{P}(A|B)=A$）

后者：B的发生不改变A的概率；前者写成了A,B对称的形式.

显然，不可能事件、必然事件与任何事件独立.

若A与B独立，则A与 B^c、A^c 与B，A^c 与 B^c 相互独立. 独立≠互斥.

同一试验、不同试验都可以有相互独立的事件.

(定义) 条件独立性(conditional independence) 给定 $C\in\mathcal{F}$，$\mathbb{P}(C)>0$，若 $\mathbb{P}(A\cap B|C)=\mathbb{P}(A|C)\mathbb{P}(B|C)$，则称A与B在给定C之下条件独立.（$\mathbb{P}_C(A\cap B)=\mathbb{P}_C(A)\mathbb{P}_C(B)$）

相当于更换了概率空间. 注意：条件独立与独立没有相互蕴含关系，不能互推.

即 若用⫫符号表示独立，$A⫫B \leftrightarrow A⫫B|C$.

(定义) 一组事件的独立性 $A_1,\dots A_n\in\mathcal{F}$，若对任意非空子集 $S\subset\{1,2,\dots,n\}$，都有 $\mathbb{P}(\bigcap_{i\in S}A_i)=\prod_{i\in S}\mathbb{P}(A_i)$，则称这组事件相互独立.

注意：任意非空子集可拆≠两两相互独立(可能随数量出错)

(定义) 两两独立性 pairwise independence. $\forall i,j\in\{1,2,\dots,n\}$ $\mathbb{P}(A_iA_j)=\mathbb{P}(A_i)\mathbb{P}(A_j)$.

例：某地区某时期每人咽试子样本中含病毒概率为p，混合100人样本，求能检出病毒阳性概率. 记 A_i 为第i个人咽试子中含病毒.

$\mathbb{P}(\bigcup_{i=1}^{100}A_i)=1-\mathbb{P}(\bigcap_{i=1}^{100}A_i^c)=1-(1-p)^{100}$

若 $p=0.4\%$，检出概率为0.33；若 $p=0.0004\%$，则检出概率降至0.04%.

例（赌徒破产模型）甲有本金a元，每局赢的概率是1/2，无论输赢，赌注都为1元，输光后也停止赌博，甲决定再赢b元后停止，求甲输光的概率 $q(a)$.

用递归方式解. 边界 $q(0)=1$，模型相当于古典概型. A表示第1局赢，B_k 表示有k元本金时输光. $q(a+b)=0$. $q(k)=\mathbb{P}(B_k)=\mathbb{P}(A)\mathbb{P}(B_k|A)+\mathbb{P}(A^c)\mathbb{P}(B_k|A^c)$

$q(k)=\frac{1}{2}[\mathbb{P}(B_{k+1})+\mathbb{P}(B_{k-1})]\Rightarrow q(k)$ 为等差数列，$q(n)=1+n(q(1)-1)$.

代 $q(a+b)=1+(a+b)(q(1)-1)=0$，有 $q(1)-1=-\frac{1}{a+b}$. 那么 $q(a)=\frac{b}{a+b}$

说明：本金有限的情况下，贪心越大，输光的概率越大；一直赌下去必定输光.

↗ 严格定义：$(\Omega,\mathcal{F})$为可测空间，若Ω上函数$X(\omega)$满足：对$\forall x\in\mathbb{R}$，$\{\omega|X(\omega)\le x\}\in\mathcal{F}$，则称$X(\omega)$为$(\Omega,\mathcal{F})$上的随机变量。

随机变量

实质：给每个可能的试验结果分配一个数，即$\Omega\to\mathbb{R}$的函数，可能离散/连续。

（也是映射，但范围更大；概率的范围仅限$[0,1]$）

随机性来自试验结果的生成，而不是映射。同一样本空间可有多个随机变量。

记自变元为ω，有$X=X(\omega)$，更多关注经X映射后的概率空间$(\mathbb{R},\mathcal{B},\mathbb{P})$。（信念转移）

随机变量的相互独立性：设$X_1,\dots X_n$是$(\Omega,\mathcal{F})$上的随机变量，若对任意实数$x_1,\dots x_n$都有$\mathbb{P}(X_1\le x_1,\dots X_n\le x_n)=\mathbb{P}(X_1\le x_1)\dots\mathbb{P}(X_n\le x_n)$，则相互独立。

· 不是任意函数都能做随机变量（不一定能分配概率）

要求：（数值化）$X:\Omega\to\mathbb{R}$，（可度量）$X^{-1}(A)\in\mathcal{F}$。

（定义）设$(\Omega,\mathcal{F})$为可测空间，若Ω上函数$X(\omega)$满足：对$\forall \omega\in\mathbb{R}$，$\{\omega|X(\omega)\le x\}\in\mathcal{F}$，则称$X(\omega)$为$(\Omega,\mathcal{F})$上的随机变量（random variable, rv），可简记为X。

性质：$X^{-1}(A\cap B)=X^{-1}(A)\cap X^{-1}(B)$，$X^{-1}(A\cup B)=X^{-1}(A)\cup X^{-1}(B)$，$X^{-1}(A^c)=[X^{-1}(A)]^c$（交并补不变）

对$X^{-1}[(-\infty,x]]\in\mathcal{F}$：交并补得到闭区间，用连续取并逼近得到开区间和孤立点。

全体事件域$\mathcal{B}=\sigma((-\infty,x])=\sigma((-\infty,x))\dots$可以通过类似区间生成。

习惯记号：$X,W,Y,Z\dots$或ξ,ζ,η，表示序列用下标。$\{X\le x\}$代替事件$\{\omega|X(\omega)\le x\}$。

Borel集：用C表示$\mathbb{R}$上左开右闭的子区间全体，$C=\{(a,b]\}$。令$\mathcal{B}=\sigma(C)$，称作Borel域；

定理 X是可测空间上的随机变量，对任意Borel集A，$\{X\in A\}\overset{def}{=}\{\omega|X(\omega)\in A\}\in\mathcal{F}$。

（对任何Borel集A，$\{X\in A\}$都是事件，于是可计算概率$\mathbb{P}(X\in A)$，保证了其可测性）

Borel可测函数定义（映射到实数轴上的所有正常集合）

定理：若X是可测空间上的随机变量，$g(x)$是可测函数，则$Y=g(x)$是$(\Omega,\mathcal{F})$上的随机变量。

构造与生成：连续函数、阶梯函数、单调函数及它们的线性组合都可测，多元同理。

四则运算、求极限（lim sup, inf），max、min、sin、e^x等变换；函数复合不改变可测性。

· 随机变量的独立性

（定义）$X_1,\dots X_n$是$(\Omega,\mathcal{F})$上的随机变量，若$\forall x_1,\dots x_n\in\mathbb{R}$均有$\mathbb{P}(X_1\le x_1,\dots X_n\le x_n)=\mathbb{P}(X_1\le x_1)\dots\mathbb{P}(X_n\le x_n)$，则称其相互独立。

（与之前的独立性相比：从S任意子集切换到全集；事件→小于关系）

补充说明（通过单调+极限证明只需一个集合而非任意子集）

Date. /

定理：设 $X_1,\dots,X_n$ 相互独立，则对任何 Borel 集 $A_1,\dots,A_n$，事件 $\{X_1\in A_1\},\dots,\{X_n\in A_n\}$ 相互独立.

(定义) 若随机变量相互独立且有相同分布函数，则称其独立同分布 (independent and identically distributed, i.i.d)

定理 设随机变量 $X_1,\dots,X_n$ 相互独立，$g_1(x),\dots,g_n(x)$ 是一元实可测函数，$\varphi(x_1,\dots,x_k)$ 是 k元实可测函数，则 1° 随机变量 $g_1(X_1),\dots,g_n(X_n)$ 相互独立.

*2° 随机变量 $\varphi(X_1,\dots,X_k), X_{k+1},\dots,X_n$ 相互独立.

常见随机变量

离散(型)随机变量 (discrete r.v) 随机变量只取有限个或可列个值.

概率分布列 (probability mass function, PMF)：对离散型随机变量 X，称 $\mathbb{P}(X=x_k)=p_k,\ k\ge1$ 为 X 的概率分布，$\{p_k\}$ 为分布列.

分布列性质：$p_k\ge0$，$\sum_{k=1}^{\infty}p_k=1$. (由互斥，可列可加性天然成立)

<1> 两点分布 $B(1,p)$ / Bernoulli 分布：

X 的取值只有 0 或 1，概率分布 $\mathbb{P}(X=1)=p=1-\mathbb{P}(X=0)$. 记作 $X\sim B(1,p)$ 或 $X\sim B(p)$.

*<1> 二项分布 (Binomial 分布：$\mathbb{P}(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$，记作 $X\sim B(n,p)$.

相当于 Bernoulli 分布连续重复 n 次.

若 $X_1,\dots,X_n$ 相互独立且都服从 $B(p)$，则 $S=X_1+X_2+\dots X_n\sim B(n,p)$.

若 $X_1\sim B(n,p)$，$X_2\sim B(m,p)$，则 $X_1+X_2\sim B(m+n,p)$. (要求：X_1,X_2 相互独立)

(证明) $P(X+Y=k)=\sum_{j=0}^{k}P(Y=k-j)P(X=j)=\sum_j\binom{m}{k-j}\binom{n}{j}p^kq^{m-k+n}=\binom{m+n}{k}p^kq^{m-k+n}$.

定理 二项分布的最大可能值 k_0 存在，即二项分布的中心项：$b(k_0,n,p)=\max_{0\le k\le n}b(k,n,p)$.

且 $k_0=\begin{cases}(n+1)p \text{ 或 } (n+1)p-1, & \text{若 } (n+1)p \text{ 为整数} \quad (\text{两个最大项})\\ [(n+1)p], & \text{若 } (n+1)p \text{ 不为整数} \quad (\text{唯一最大项})\end{cases}$

*<2> 几何分布 / Geometric 分布

$\mathbb{P}(X=k)=(1-p)^{k-1}p,\ k=1,2,\dots$，$X\sim G(p)$. (连续若干次后发生)

定理：随机变量 $X\sim G(p)$ 的充要条件是无记忆性 (memoryless)，即 $\mathbb{P}(X=k+1\mid X>k)=\mathbb{P}(X=1)$.

在离散的条件下，无记忆性和几何分布可以互推，二者等价.

(证明) $\Rightarrow$: $P(X=k+1|X>k)=\frac{P(X=k+1, X>k)}{P(X>k)}=\frac{P(X=k+1)}{P(\bigcup_{j=k+1}^{\infty} X=j)}=\frac{(1-p)^k p}{\sum_{j=k+1}^{\infty}(1-p)^{j-1}p}=p=P(X=1)$.

$\Leftarrow$: 记 $r_k=P(X>k)$, $\{X=k+1\}=\{X>k\}-\{X>k+1\}$, 那么 $P(X=k+1)=r_k-r_{k+1}$. $r_0=1$.

$P(X=k+1|X>k)=1-\frac{r_{k+1}}{r_k}$, 代入并用 $P(X=1)=p$, 累加得 $P(X=k)=(1-p)^{k-1}p$.

帕斯卡分布 Pascal distribution

$P(X=k)=\binom{k-1}{r-1}(1-p)^{k-r}p^r$, 直到成功r次后停止.

负二项分布 $NB(r,p)$. 概率分布为 $P(Y=k)=\binom{k+r-1}{r-1}(1-p)^k p^r$, $k\in Z$. Y是停止时失败次数.

以后帕斯卡分布和负二项分布不再区分. 可以看成若干独立的几何分布之和.

<3> 超几何分布 (Hypergeometric distribution) 无放回,相当于抽取一组

$P(X=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$, $k\le\min(n,M)$, 记作 $X\sim HG(n,M,N)$.

应用: 质量检测; 捕获-再捕获模型(标志重捕法)

理解: N件产品中有M件次品,抽取n件,其中有k件次品的概率. 次品率 $p=\frac{M}{N}$.

在N足够大的情况下,无放回可以用有放回(二项分布)近似.

有放回 $\tilde{X}\sim B(n,\frac{M}{N})$, 无放回 $X\sim HG(n,M,N)$.

<4> 负超几何分布 (Negative Hd)

$P(X=k)=\frac{\binom{k+r-1}{k}\binom{N-k-r}{M-r}}{\binom{N}{M}}$, $k\le N-M$. $X\sim NH(r,M,N)$.

例 在包含N个元素的集合中有M个红球、N-M个黑球,每次无放回地抽取一个球,直到抽出r个. 求此时黑球为k个的概率.

先考虑第k+r个: $\frac{M}{N}$ 概率出红球, 前k+r-1个服从 $H(k+r-1,M-1,N-1)$ 超几何分布.

以上4种分布满足如下关系

$$
\begin{tikzcd}
 & \text{二项} \quad B(p) \arrow[d, "\perp\!\!\!\perp +"'] & \text{有放回 rep} & \text{无放回 no rep} & \\
P(\lambda) \arrow[dd, dashed] & B(n,p) \arrow[l, "n\to\infty"'] & & HG(n,M,N) \arrow[ll, "\frac{M}{N}\to p,\ N\to\infty", "\text{极限等价}"'] & \#\text{success in fixed \#trials} \\
 & NB(r,p) & & NHG(r,M,N) \arrow[ll] & \#\text{trials of fixed \#success} \\
E(\lambda) \arrow[d] & G(p) \arrow[l, dash, "\text{无记忆性}"'] \arrow[u, "+"] & & & \\
\Gamma(n,\lambda) & & & &
\end{tikzcd}
$$

第一行: 确定抽样数,求抽中次数; 第二行: 抽中若干次,求抽取次数.

Date. /

*<5> 泊松分布 (Poisson distribution)

$\mathbb{P}(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$，$X\sim P(\lambda)$，其中 λ 是正常数. $\frac{\lambda^k}{k!}$ 是核心项.

以放射性物质为例证明泊松分布：将时间 t 等分成 n 段，每段是 $\delta_n=\frac{t}{n}$ 秒. n 足够大

假定 1) 在 δ_n 内最多放出一个粒子，且放出的概率为 $p_n=\mu\delta_n=\frac{\mu t}{n}$，$\mu$ 为正常数.

2) 各时间段内是否放出粒子相互独立.

定义 $\lambda\triangleq\mu t=np_n$，相当于定义为 总时间内的平均个数 / 单位时间内的发生密度.

在上述假定下，放射性物质放出粒子数 $X\sim B(n,p_n)$.

$\mathbb{P}(X=k)=\lim_{n\to\infty}\binom{n}{k}p_n^k(1-p_n)^{n-k}=\lim_{n\to\infty}\frac{(\mu t)^k}{k!}\frac{n(n-1)\cdots(n-k+1)}{n^k}\left(1-\frac{\mu t}{n}\right)^{n-k}=\frac{(\mu t)^k}{k!}e^{-\mu t}$.

二项分布可以用 Poisson 分布近似 (n 很大，p 很小时) 特别是在 n 无法确定的情况下.

也不一定限制在一定时间内，如一张纸上的雨点数.

连续型随机变量

(定义) 对随机变量 X，若存在非负函数 $f(x)$ 使得对任意 $a<b$，(相当于信念分布)，

$\mathbb{P}(a<X\le b)=\int_a^b f(x)dx$，则称 X 是连续型随机变量 (continuous rv)，称 $f(x)$ 是 X 的概率密度函数，简称概率密度或密度 (probability density function, PDF)

$f(x)$ 性质 1° (归一性) $\int_{-\infty}^{+\infty}f(x)dx=1$

2° (单点) $\mathbb{P}(X=a)=0$. 推论：$\mathbb{P}(a<X\le b)=\mathbb{P}(a\le X\le b)$.

3° 对任意的 Borel 集 A，有 $\mathbb{P}(X\in A)=\int_A f(x)dx$.

理解连续性：$X(\omega)$ 不一定是连续函数，因为其样本空间无拓扑结构，不构成连续性.

$f(x)$ 也可能不是连续函数. 连续性随机变量 指它的分布函数绝对连续.

PDF 在其定义域内未必有界，如 $f(x)=\frac{1}{2\sqrt{x}}$ 也是密度函数.

常见连续型随机变量

<1> 均匀分布 $U(a,b)$：$f(x)=\begin{cases}\frac{1}{b-a}, & x\in(a,b)\\ 0, & x\notin(a,b)\end{cases}$ $X\sim U(a,b)$. $U\to$ uniform.

还可以写成示性函数 $f(x)=\frac{1}{b-a}I_{(a,b)}(x)=\frac{1}{b-a}I_{x\in(a,b)}$.

对任意 Borel 集 A，若 A 的测度 $m(A)=\int_A 1dx<\infty$，可以类似定义 A 上的均匀分布，

若 $f(x)=\begin{cases}\frac{1}{m(A)}, & x\in A\\ 0, & x\in A^c\end{cases}$. $X\sim U(A)$. 对任意的其它 Borel 集 B，

$\mathbb{P}(B)=\frac{m(A\cap B)}{m(A)}$ (可以用测度转化)

<2>指数分布（到发生某个事件所用时间）

$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geq 0\\ 0, & x<0.\end{cases}$ λ为正常数，$X\sim\mathcal{E}(\lambda)$，$\mathcal{E}\to$ Exponential. 均值为$\frac{1}{\lambda}$.

写作示性函数：$f(x)=\lambda e^{-\lambda x}I_{(0,+\infty)}=\lambda e^{-\lambda x}, x\geq 0$. 或 $f(x)=\frac{1}{\theta}e^{-\frac{x}{\theta}}I_{(0,+\infty)}$.

X超过某个值的概率随值增加指数递减，与几何分布非常相似.

定理：连续型随机变量X服从指数分布的充要条件是X无记忆性.

即 $\forall s,t\geq 0$，有 $\mathbb{P}(X>s+t\mid X>s)=\mathbb{P}(X>t)$.

(证明) ⇒：$X\sim\mathcal{E}(\lambda)$，$\mathbb{P}(X>x)=\int_x^{\infty}\lambda e^{-\lambda y}dy=e^{-\lambda x}$；

$\mathbb{P}(X>s+t\mid X>s)=\frac{\mathbb{P}(X>s+t)}{\mathbb{P}(X>s)}=\frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}=e^{-\lambda t}=\mathbb{P}(X>t)$

⇐：令 $G(x)=\mathbb{P}(X>x)$，$G(s+t)=G(s)G(t)$ 即 $\log G(s+t)=\log G(s)+\log G(t)$. $\log G(x)=\log G(1)\cdot x$.

$G(x)$单调非增，$\lambda=-\log\mathbb{P}(X>1)$. $\mathbb{P}(a<X\leq b)=e^{-\lambda a}-e^{-\lambda b}=\int_a^b\lambda e^{-\lambda y}dy$.

· 失效率：单位时间内失效的概率. $\mathbb{P}=\lim_{\Delta t\to 0}\frac{\mathbb{P}(t<X\leq t+\Delta t\mid X>t)}{\Delta t}=\lambda$.

由定义，$\lambda\triangleq\mu t$. 如失效6/h，等价于$\frac{1}{10}$/min，其实是线性的尺度转换.

若$X\sim\mathcal{E}(\lambda)$，$Y=\lambda X$，那么$Y\sim\mathcal{E}(1)$. λ称为尺度参数.（不同坐标系）rate parameter.

指数分布描述的是无老化时的寿命分布，是一种近似.

<3>正态分布 Normal distribution.

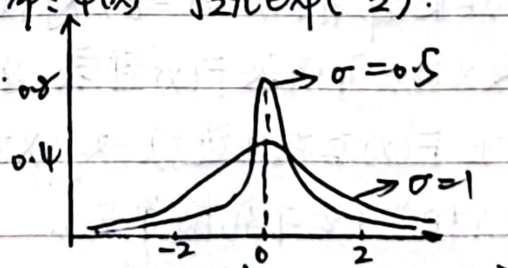
(一般)正态分布：$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$，$X\sim N(\mu,\sigma^2)$.

特别地，当$X\sim N(0,1)$时，X服从标准正态分布：$\varphi(x)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)$.

密度函数性质：(又称为钟型分布) 如右图.

关于$x=\mu$对称，$x=\mu\pm\sigma$处有拐点

$f(x)$：$x=(\sigma\sqrt{2\pi})^{-1}$处取最大值.



(可以调节μ以平移图像)

<4>其他连续型随机变量介绍

问题：PMF和PDF不能相互转换.

· 概率分布函数：对随机变量X，定义 $F(x)=\mathbb{P}(X<x)$，$x\in\mathbb{R}$.

为X的(概率)分布函数/累积分布函数(cumulative distribution function, CDF)

CDF与PMF：阶跃式函数，一定要画出端点的空心圆. 两者可以相互唯一确定.

$p_k=\mathbb{P}(X=x_k)$，$F(x)=\mathbb{P}\left(\bigcup_{j:x_j<x}\{X=x_j\}\right)=\sum_{j:x_j<x}p_j$.

Date.　　/

CDF与PDF: $F(x)=\mathbb{P}(X\le x)=\int_{-\infty}^{x} f(t)dt,\ x\in\mathbb{R}$.　　(f(x)可以不连续)

F(x)是连续函数，且在f(x)的连续点有$f(x)=F'(x)$，F(x)是f(x)的分布函数.

PDF可以唯一确定CDF，但CDF不能唯一确定PDF（除非f(x)连续）

标准正态分布：（专用符号$\varphi(x)$，$\Phi(x)$）$\Phi(x)=\int_{-\infty}^{x}\varphi(t)dt=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$.

由对称性，$\Phi(x)+\Phi(-x)=1$，没有解析解，参考正态分布表.

对一般的正态分布：$X\sim N(\mu,\sigma^2)$，$\mathbb{P}(X\le a)=\Phi(\frac{a-\mu}{\sigma})$.

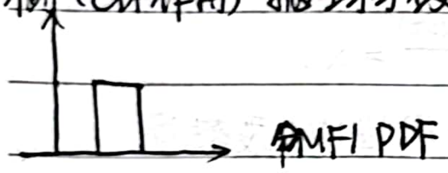
3σ准则：$\mathbb{P}(|X-\mu|\le 3\sigma)=99.73\%$.

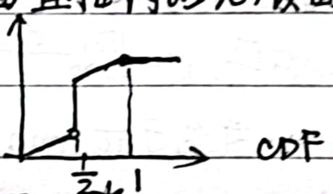
双向互推　CDF

PMF　　PDF

PDF可推CDF，f(x)连续时CDF可推PDF.

·例（CDF作用）抛均匀硬币，正面直接得0.5元，反面转转盘得0到1元的均匀分布.





·分布函数的性质：1° F(x)单调不降

2° $F(-\infty)=\lim_{x\to-\infty}F(x)=0$，$F(+\infty)=\lim_{x\to+\infty}F(x)=1$

3° F(x)是右连续的.

可以通过CDF直接判断分布的类型和随机变量的类型.

若CDF中有单点概率不为0情况，则对应变量一定不是连续型随机变量.

·若对$\forall x\in\mathbb{R}$都有$\mathbb{P}(X=x)=0$，则$F(x)=\mathbb{P}(X\le x)$是连续的.　(若F不连续↑)

定理：设X的分布函数F(x)连续，数集A中任两点之间距离大于某个正数δ.

若在A^c上$F'(x)$存在且连续，则X的密度函数：$f(x)=\begin{cases}F'(x), & x\notin A\\ 0, & x\in A\end{cases}$.

(数集A相当于若干离散单点).

(证明) $\forall a<b$，不妨设$a_0=a<a_1<\dots a_k<b=a_{k+1}$，$(a,b)$中只含A中的$a_1,\dots,a_k$.

此时 $\mathbb{P}(a<X\le b)=\sum_{j=0}^{k}\mathbb{P}(a_j<X\le a_{j+1})=\sum_{j=0}^{k}[F(a_{j+1})-F(a_j)]=\sum_{j=0}^{k}\int_{a_j}^{a_{j+1}}f(t)dt=\int_a^b f(t)dt$.

此时f(x)是X的概率密度.

说明：F连续不能保证X是连续型随机变量，但F不连续时X一定不是连续型.

(反例)若F是二项分布的随机变量的分布函数时，除有限个点外，$F'(x)=0$，所以

此时X的密度不存在.