概率的公理化：S为样本空间，F为某些子集组成事件域，定义在F上的实值函数P满足：① $\forall A\in F, P(A)\geq 0$ ② $P(S)=1$ ③ 对互不相容的事件$A_1,A_2,A_3,\cdots$，有$P(\bigcup_{i=1}^{+\infty}A_i)=\sum_{i=1}^{+\infty}P(A_i)$。则称$P(A)$为事件$A$的概率，$(S,F,P)$为概率空间。

一般加法公式：$P(\bigcup_{i=1}^{n}A_i)=\sum_{i=1}^{n}P(A_i)-\sum_{1\leq i<j\leq n}P(A_iA_j)+\cdots+(-1)^{n-1}P(A_1\cdots A_n)$

全概率公式：$B_1,B_2,\cdots B_n$为S的划分，即B_i互不相容，且$\bigcup_{i=1}^{n}B_i=S$。

若$P(B_i)>0$，则 $P(A)=\sum_{i=1}^{n}P(B_i)\cdot P(A|B_i)$。

条件概率版：$P(C)>0$，$P(A|C)=\sum_{i=1}^{n}P(B_i|C)\cdot P(A|B_iC)$。

贝叶斯公式：$P(A|B)=P(B|A)\cdot\frac{P(A)}{P(B)}$

相互独立：n个事件$A_1,\cdots A_n$，对于$\{1,\cdots,n\}$的任意子集ℓ，都有$P(\bigcap_{i\in S}A_i)=\prod_{i\in S}P(A_i)$。　相互独立$\Rightarrow$两两独立。

布尔不等式：$P(\bigcup_i A_i)\leq\sum_i P(A_i)$.（归纳法证明）

期望：对离散随机变量X，若$\sum_i|x_ip_i|<\infty$，则$E(X)=\sum_i x_ip_i$

证明存在某种分法：$P(X\geq E(X))>0$，$P(X\leq E(X))>0$ ← 反证 $P(X\leq E(X)+\varepsilon)\leq\frac{1}{2}$

拆分子问题：（换视角）$X=\sum_i \mathbb{1}_{A_i}\Rightarrow E(X)=\sum_i E(\mathbb{1}_{A_i})$

马尔可夫不等式：若X为非负随机变量，$E(X)>0$，$a>0$，

则 $P(X\geq a\cdot E(X))\leq\frac{1}{a}$　或：$P(X\geq a)\leq\frac{E(X)}{a}$

推广：令$Y=e^{\lambda X}(\lambda>0)\Rightarrow P(X\geq t)\leq\frac{E(e^{\lambda X})}{e^{\lambda t}}$.

切比雪夫不等式：若$\sigma(X)>0$，$\forall c>0$，

（令$Y=(X-E(X))^2$），则 $P(|X-E(X)|\geq c\cdot\sigma(X))\leq\frac{1}{c^2}$.

也即 $P(|X-E(X)|\geq a)\leq\frac{Var(X)}{a^2}$

方差：若$E[(X-E(X))^2]$存在．$Var(X)=E[(X-E(X))^2]$

性质：① $Var(aX+b)=a^2Var(X)$，$\sigma(aX+b)=|a|\cdot\sigma(X)$

② $Var(X)=E(X^2)-E^2(X)$ ③ $Var(X\pm Y)=Var(X)+Var(Y)$　X,Y独立

协方差 $Cov(X,Y)=E[(X-E(X))\cdot(Y-E(Y))]$

$=E(XY)-E(X)E(Y)$.

性质：$Var(A+B)=Var(A)+Var(B)+2Cov(A,B)$

推广：$Var(\sum_{i=1}^{n}X_i)=\sum_{i=1}^{n}Var(X_i)+2\sum_{1\leq i<j\leq n}Cov(X_i,X_j)$.

$Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$

边际分布列：联合分布列$\Rightarrow$单变　$(P(Y=y_i)>0)$

条件分布列：$p_{i|j}=P(X=x_i|Y=y_j)=\frac{P(X=x_i,Y=y_j)}{P(Y=y_j)}$.

两两独立：$\forall x,y$，$P(X=x,Y=y)=P(X=x)P(Y=y)$.

相互独立$\Rightarrow E(XY)=E(X)E(Y)$.（可推广）

$\Rightarrow Var(X\pm Y)=Var(X)\pm Var(Y)$.

重期望公式：$E(E(X|Y))=E(X)$

$X_1,X_2,\cdots X_n$为i.i.d.随机变量N取正整数

作业题：且与$\{X_n\}$独立$\Rightarrow E(\sum_{i\leq N}X_i)=E(N)\cdot E(X_1)$

$Var(\sum_{i\leq N}X_i)=E(N)\cdot Var(X_1)+Var(N)\cdot[E(X_1)]^2$

X为离散随机变量，X仅取非负整数值，则$E(X)=\sum_{x=0}^{+\infty}P(X>x)$.（交换积分）

X为连续随机变量，X仅取非负实数值，则$E(X)=\int_0^{+\infty}P(X>x)dx$.

设X为连续随机变量，若函数$y=g(x)$严格单调，且反函数$h(y)$有连续导数，则$Y=g(X)$的概率密度函数为

$f_Y(y)=f_X(h(y))|h'(y)|$，当$y\in(\alpha,\beta)$．α,β为对应$g(\pm\infty)$.

$k^2=k(k-1)+k$.　证

二项分布 $X\sim B(n,p)$．n:次数，p:向上概率．

$E(X)=np$，$E(X^2)=n(n-1)p^2+np$，$Var(X)=np(1-p)$.

可加性：$X\sim B(n,p)$，$Y\sim B(m,p)\Rightarrow X+Y\sim B(m+n,p)$.

泊松分布 $X\sim\pi(\lambda)$，$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$，$k\in N$.

$E(X)=\lambda$，$E(X^2)=\lambda^2+\lambda$，$Var(X)=\lambda$.

可用于近似n很大，p很小的二项分布．$\lim_{n\to+\infty}n\cdot p_n=\lambda$.

可加性：$X\sim P(\lambda_1)$，$Y\sim P(\lambda_2)$，$X+Y\sim P(\lambda_1+\lambda_2)$.

超几何分布，$X\sim h(n,N,M)$，N中有M个坏，抽n个．坏的个数

$P(N=k)=\binom{M}{k}\binom{N-M}{n-k}/\binom{N}{n}$，$k=0,1,\cdots,\min\{n,M\}$.

$E(X)=\frac{nM}{N}$，$E(X^2)=\frac{M(M-1)n(n-1)}{N(N-1)}-\frac{nM}{N}$

$Var(X)=\frac{nM(N-M)(N-n)}{N^2(N-1)}$.

几何分布　**无记忆性**．$P(X>m+n|X>n)=P(X>m)$.

$X\sim G(p)$．伯努利实验．首次出现．

$P(X=k)=p(1-p)^{k-1}$，$E(X)=\frac{1}{p}$，$E(X^2)=\frac{2-p}{p^2}$，$Var(X)=\frac{1-p}{p^2}$

负二项分布　**无记忆性**：$NB(a,p)+NB(b,p)\sim NB(a+b,p)$

$X\sim NB(r,p)$．结果A发生r次时次数n　$\boxed{p^r=(1-(1-p))^r}$

$P(X=k)=\binom{k-1}{r-1}p^r\cdot(1-p)^{k-r}$　$k=r,r+1,\cdots$．正则性（r个$G(p)$之和）

$E(X)=\frac{r}{p}$，$Var(X)=\frac{1-p}{p^2}$，$E(X^2)=\frac{n^2+n-np}{p^2}$（分类讨论）

（右连续）

连续随机变量：随机X，$\exists f(x)$，s.t. $F(x)=\int_{-\infty}^{x}f(t)dt$

对$f(x)$ ① 非负 ② 正则 $\int_{-\infty}^{+\infty}f(x)=1$．导数存在点，$F'(x)=f(x)$.

若$\int_{-\infty}^{+\infty}f(x)|x|dx<+\infty$，则$E(x)=\int_{-\infty}^{+\infty}xf(x)dx$（绝对收敛）

给定X.g，求$E(Y=g(x))$ $E(g(X))=\int_{-\infty}^{+\infty}f_X(x)\cdot g(x)dx$.

$P(X\leq E(x))>0$．连续$\Rightarrow$若$P(X\leq E(x))=0$，$\exists\varepsilon$，s.t. $P(X\leq E(X)+\varepsilon)\leq\frac{1}{2}$.

均匀分布 $X\sim U(a,b)$．$x\in(a,b)$，$f(x)=\frac{1}{b-a}$.

$E(x)=\frac{a+b}{2}$，$E(X^2)=\frac{b^3-a^3}{3(b-a)}$，$Var(X)=\frac{(b-a)^2}{12}$.

$F(X)=0\ (x<a)$；$\frac{x-a}{b-a}\ (a\leq x<b)$；$1\ (x\geq b)$.

bias $\leq 2\sigma$: 95%　$\leq 3\sigma$: 99%.

标准正态分布 $X\sim N(0,1)$．$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$

$\int_{-\infty}^{+\infty}e^{-t^2}dt=\sqrt{\pi}$，$E(X)=0$，$E(X^2)=1$，$Var(X)=1$

正态分布 $X\sim N(\mu,\sigma^2)$．$E(X)=\mu$.

$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$，$Var(X)=\sigma^2$．对称性

可加性：$X+Y\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$

指数分布 $X\sim Exp(\lambda)$．$\lambda>0$．$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geq 0\\ 0, & x<0\end{cases}$

$F(X)=\begin{cases}1-e^{-\lambda x}, & x\geq 0\\ 0, & x<0\end{cases}$　$E(X)=\frac{1}{\lambda}$，$E(X^2)=\frac{2}{\lambda^2}$，$Var(X)=\frac{1}{\lambda^2}$（分部积）

伽玛分布　**无记忆性**：$P(X>s+t|X>s)=P(X>t)$

$X\sim\Gamma(\alpha,\lambda)$．$\alpha,\lambda>0$．$\Gamma(\alpha)=\int_0^{+\infty}x^{\alpha-1}e^{-x}dx$

$\Gamma(n)=(n-1)!$，$\Gamma(\frac{1}{2})=\sqrt{\pi}$，$\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$

$f(x)=\begin{cases}\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, & x\geq 0\\ 0, & x<0\end{cases}$　$E(X)=\frac{\alpha}{\lambda}$，$E(X^2)=\frac{\alpha(\alpha+1)}{\lambda^2}$，$Var(X)=\frac{\alpha}{\lambda^2}$

可加性：$X+Y\sim Ga(\alpha_1+\alpha_2,\lambda)$　$\alpha=1\Rightarrow$指数分布

$\alpha=\frac{n}{2}$，$\lambda=\frac{1}{2}\Rightarrow$自由度为$n$的**$\chi^2$（卡方）分布**．

$Z_i\sim N(0,1)$，$\chi^2(k)=\sum_{i=1}^{k}Z_i^2\sim Ga(\frac{k}{2},\frac{1}{2})$.

泰勒公式

$e^x=1+x+\frac{1}{2!}x^2+\frac{1}{3!}x^3+\cdots+\frac{1}{n!}x^n+o(x^n)$

$\sin x=x-\frac{1}{3!}x^3+\frac{1}{5!}x^5+\cdots+(-1)^{n-1}\frac{1}{(2n-1)!}x^{2n-1}$

$\cos x=1-\frac{1}{2!}x^2+\frac{1}{4!}x^4+\cdots+(-1)^n\frac{1}{(2n)!}x^{2n}$

$(1+x)^\alpha=1+\alpha x+\frac{\alpha(\alpha-1)}{2}x^2+\cdots+\frac{\prod_{\alpha}^{\alpha-n+1}}{n!}x^n$.

$\frac{1}{1+x}=1-x+x^2+\cdots+(-1)^nx^n$ | $\frac{1}{1-x}=1+x+\cdots+x^n$.

$\ln(1+x)=x-\frac{1}{2}x^2+\frac{1}{3}x^3+\cdots+(-1)^n\cdot\frac{1}{n}x^n$.

$\ln(1-x)=-x-\frac{1}{2}x^2-\frac{1}{3}x^3-\cdots-\frac{1}{n}x^n$

（可复用）

$C_n^k=\frac{k}{n}C_{n-1}^{k-1}$　$C_n^k=C_{n-1}^{k}+C_{n-1}^{k-1}$　$(x+y)^n=\sum_{k=0}^{n}\binom{n}{k}x^ky^{n-k}$.

$A_n^k=nA_{n-1}^{k-1}$　$A_n^k=kA_{n-1}^{k-1}+A_{n-1}^{k}$　$\sum_{i=0}^{k}\binom{n}{i}\binom{m}{k-i}=\binom{n+m}{k}$

柯西不等式

$(\sum_{i=1}^{n}a_ib_i)^2\leq(\sum_{i=1}^{n}a_i^2)(\sum_{i=1}^{n}b_i^2)$.

令$a_i=\sqrt{p_i}$，$b_i=x_i\sqrt{p_i}\Rightarrow E^2(X)\leq E(X^2)$

积分微分 $da^x=a^x\ln a\,dx$　$I_n=\int_0^{+\infty}e^{-x}\cdot x^ndx=nI(n-1)$

$d\sec x=\sec x\tan x\,dx$　$d\arctan x=\frac{1}{1+x^2}dx$

$d\csc x=-\csc x\cot x\,dx$　$d\,\mathrm{arccot}\,x=-\frac{1}{1+x^2}dx$

$d\arcsin x=\frac{1}{\sqrt{1-x^2}}dx$　$d\,\mathrm{arcsec}\,x=\frac{1}{x\sqrt{x^2-1}}dx$

$d\arccos x=-\frac{1}{\sqrt{1-x^2}}dx$　$d\,\mathrm{arccsc}\,x=\frac{1}{-x\sqrt{x^2-1}}dx$

$d\tan x=\sec^2x\,dx$　$d\cot x=-\csc^2x\,dx$

$\int\frac{1}{x}=\ln|x|$ | $\int\tan x=-\ln|\cos x|$ | $\int\cot x=\ln|\sin x|$

$\int\sec x=\ln|\sec x+\tan x|$　万能公式：$t=\tan\frac{x}{2}$.

$\int\csc x=\ln|\cot x-\csc x|$ | $\int\frac{1}{\sqrt{x^2\pm a^2}}=\ln|x\pm\sqrt{a^2\pm x^2}|$

$\int\arccos x=x\arccos x-\sqrt{1-x^2}$ | $\int\arctan x=x\arctan x-\frac{1}{2}\ln(1+x^2)$

$\int\mathrm{arccsc}\,x=x\,\mathrm{arccsc}\,x+\ln|x+\sqrt{x^2-1}|$ | $\int\frac{1}{\sqrt{1-x^2}}=\arcsin x=-\arccos x$

$\int\frac{1}{1+x^2}=\arctan x=-\mathrm{arccot}\,x$ | $\int\frac{1}{x\sqrt{x^2-1}}=\mathrm{arcsec}\,x=-\mathrm{arccsc}\,x$

$\int a^x\ln a=a^x$ | $\int\sqrt{x^2\pm a^2}=\frac{x}{2}\sqrt{x^2\pm a^2}\pm\frac{a^2}{2}\ln|x+\sqrt{x^2\pm a^2}|$

$\int\sqrt{a^2-x^2}=\frac{x}{2}\sqrt{a^2-x^2}+\frac{a^2}{2}\arcsin\frac{x}{|a|}$ | $\int\frac{1}{(x+a)^2+b^2}=\frac{1}{b}\arctan\frac{x+a}{b}$

$\int\frac{1}{x^2-a^2}=\frac{1}{2a}\ln\left|\frac{x-a}{x+a}\right|$ | $\int\frac{1}{\sqrt{a^2-x^2}}=\arcsin\frac{x}{a}$

伯努利不等式：$(A+B)^n\geq A^n+nA^{n-1}B$　$(A>0, A+B>0)$

均值不等式：$\left(\sum_i\frac{1}{x_i}\right)^{-1}\leq\sqrt[n]{\prod_i x_i}\leq\frac{1}{n}\sum_i x_i\leq\sqrt{\frac{1}{n}\sum_i x_i^2}$.

极限&无穷小 $\lim_{n\to\infty}\sqrt[n]{n}=1$ / $\lim_{n\to\infty}(1+n)^{\frac{1}{n}}=1$ / $\lim_{n\to\infty}(1+\frac{x}{n})^n=e^x$.

$n!\sim\sqrt{2\pi n}\left(\frac{n}{e}\right)^n[1+o(\frac{1}{n})]$ / $\lim_{n\to\infty}\frac{n}{\sqrt[n]{n!}}=e$ / $\lim_{x\to 0}(1+x)^\alpha-1\sim\alpha x$

$\lim_{x\to 0}a^x-1\sim x\ln a$ / $\lim_{x\to 0^+}\ln x=o(x^{-\alpha})\ (\alpha>0)$ / $\lim_{x\to 0}e^{-\frac{1}{x^2}}=o(x^n)$.

无穷大估计：$n\to\infty$时，$\ln n\ll n^\varepsilon\ll a^n\ll n!\ll n^n$　$(\varepsilon>0, a>1)$

期望放缩：$E[(X-E(X))^2]\geq\sum_{x\geq m}p(x)(x-E(X))^2\geq\sum_{x\geq m}p(x)(m-E(X))^2$

$E(e^{Xt})=E\left(\sum_{n=0}^{\infty}\frac{1}{n!}(Xt)^n\right)=\sum_{n=0}^{\infty}\frac{t^n}{n!}E(X^n)$.

中心极限定理 $X_1,X_2,\cdots X_n$ i.i.d.，$E(X_i)=\mu$，$Var(X_i)=\sigma^2$，

$S_n=\sum_{i=1}^{n}X_i$，当$n\to+\infty$，$\frac{S_n-n\mu}{\sqrt{n}\sigma}\sim N(0,1)$．即$S_n\sim N(n\mu,n\sigma^2)$

常见题型 ① 已知$p_X(x)$，$Y=g(x)$，求$p_Y(y)$．先算$P(Y\leq y)$，再求导．

②（换视角）n不同球，有放回取m次．X=取出不同编号数量$\Rightarrow A_i=\mathbb{I}$($i$取出过)

$X=\sum_i A_i$，$E(X)=\sum_i E(A_i)=n\cdot(1-(1-\frac{1}{n})^m)$．③（巧取$A_i$）．~，直至所有球至少取1次，$X$=取总次数$\Rightarrow X_i$表示取到$i-1$个不同，再取到$i$个不同所加次数

$X_2\sim(\frac{n-1}{n})$，$X_i\sim G(\frac{n-i+1}{n})$，$E(X_i)=\frac{n}{n-i+1}$，$E(X)=\sum E(X_i)=n\cdot(\frac{1}{n}+\frac{1}{n-1}+\cdots+1)$.

④ 进店，$X\sim\pi(\lambda)$，$Y=B(X,p)$，$Z=B(X,1-p)\Rightarrow Y.Z$独立，$Cov(X,Y)=Var(Y)=\lambda p$

⑤ 随机图模型，X_i：i认识人；A_{ab}：$a.b$认识，$E(X_i)=E(\sum_{k\neq i}\mathbb{1}(A_{ik}))=\frac{n-1}{2}$.

$X_iX_j=[\sum_{k\neq i}\mathbb{1}(A_{ik})][\sum_{\ell\neq j}\mathbb{1}(A_{\ell j})]=\sum_{k\neq i}\sum_{\ell\neq j}\mathbb{1}(A_{ik})\cdot\mathbb{1}(A_{\ell j})=\sum_{k\neq i}\sum_{\ell\neq j}\mathbb{1}_{A_{ik}\cap A_{\ell j}}=\cdots P(A_{ik}\cap A_{j\ell})$

$E(X_iX_j)=1\cdot\frac{1}{2}+[(n-1)^2-1]\cdot\frac{1}{4}=\frac{(n-1)^2+1}{4}$；$Cov(X_i,X_j)=\frac{n-1}{4}\ (i=j)$ / $E(X_iX_j)-E(X_i)E(X_j)=\frac{1}{4}$

⑥ 分治：$X\sim G(p)\Rightarrow E(X)=p+(1-p)[E(X)+1]$，$Y=\mathbb{1}_{X=1}$．⑦ 球与桶 n个球，等可能放入m桶，X_i：第i桶中数$\Rightarrow X_i\sim B(n,\frac{1}{m})$．⑧ 换视角 $1,\cdots,n$打乱，π_i为乱后编号，X：局部最大值$\Rightarrow A_i$：是~，$E(X)=\sum_{2}^{n-1}E(\mathbb{1}_{A_i})=\frac{n-2}{3}$（先构序再放）

⑨ 概率证法 $P(A>0)$，则$A\neq\phi$．⑩ $E[g(x)]=E\{g(x)\cdot I(|x|\leq M)\}+E\{g(x)\cdot I(|x|>M)\}$

⑪ $E(e^{tS_n})=E(e^{t\sum_i X_i})=[E(e^{tX_i})]^n$．$P(\frac{S_n}{n}\geq a)=P(S_n\geq na)=P(e^{\cdots})$

$\leq\frac{E(e^{tS_n})}{e^{tna}}$，最小化右式$\Rightarrow$求导（取对数）⑪ $e=2.718$　by Arthals

多维连续随机变量

边际分布： $F(x,+\infty)$

边际密度： $f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\,dy$

条件密度： $f(x|y)=\frac{f(x,y)}{f_Y(y)}$

相互独立： $f(x,y)=f_X(x)\cdot f_Y(y)$

二维正态分布 $\begin{cases}u=\dfrac{\frac{x-\mu_1}{\sigma_1}-\rho\frac{y-\mu_2}{\sigma_2}}{\sqrt{1-\rho^2}}\\ v=\dfrac{y-\mu_2}{\sigma_2}\end{cases}\Rightarrow -\frac12(u^2+v^2)$

$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}+\frac{(y-\mu_2)^2}{\sigma_2^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\right]\right\}$

$\left|\frac{\partial(x,y)}{\partial(u,v)}\right|=\left|\frac{\partial(u,v)}{\partial(x,y)}\right|^{-1}=\sigma_1\sigma_2\sqrt{1-\rho^2}$ 相互独立 $\rho=0$

给定 $Y=y$，X的条件分布 $\sim N(\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2),\sigma_1^2(1-\rho^2))$

$Cov(X,Y)=\rho\sigma_1\sigma_2$ （证：换元u,v，利用对称性）

$Corr(X,Y)=Cov(X,Y)/\sigma_X\sigma_Y\in[-1,1]$

卷积：X,Y独立，$Z=X+Y\Rightarrow f_Z(z)=\int_{-\infty}^{+\infty}f_Y(z-x)f_X(x)dx$

若 $X_i\sim N(\mu_i,\sigma_i^2)\Rightarrow Z=\sum a_iX_i\sim N(\sum\mu_i,\sum a_i^2\sigma_i^2)$

若 X_i独立，$Y=\max(X_i)\Rightarrow F_Y(y)=\prod F_{X_i}(y)$

$Y=\min(X_i)\Rightarrow F_Y(y)=1-\prod[1-F_{X_i}(y)]$

若X,Y联合密度函数 $f(x,y)$，$\begin{cases}u=u(x,y)\\v=v(x,y)\end{cases}$ 有连续偏导数

且 $\begin{cases}x=x(u,v)\\y=y(u,v)\end{cases}$ 为唯一反函数，则 $U=u(X,Y),V=v(X,Y)$

的联合密度函数为 $f(x(u,v),y(u,v))\cdot|J|$

$J=\left|\frac{\partial(x,y)}{\partial(u,v)}\right|=\begin{vmatrix}\frac{\partial x}{\partial u}&\frac{\partial x}{\partial v}\\\frac{\partial y}{\partial u}&\frac{\partial y}{\partial v}\end{vmatrix}$

Chernoff界 $M_X(t)=E(e^{tX})=\sum_{i=0}^{+\infty}\frac{t^i}{i!}E(X^i)$

$t>0$ $P(X\ge k)=P(e^{tX}\ge e^{tk})\le M_X(t)\cdot e^{-tk}$

$t<0$ $P(X\le k)=P(e^{tX}\ge e^{tk})\le M_X(t)\cdot e^{-tk}$

Hoeffding引理：$a\le X\le b$，则 $E(e^{t(X-E(X))})\le e^{\frac{t^2(b-a)^2}{8}}$

Chernoff-Hoeffding不等式 证明 $M_{X-E(X)}(t)$

若 $X=\sum X_i$，X_i相互独立，且 $a\le X_i\le b$，则：

$P(X\ge E(X)+k)\le\exp(-\frac{2k^2}{n(b-a)^2})$ 令 $k=n\varepsilon$： $\exp(-\frac{2n\varepsilon^2}{(b-a)^2})$

$P(X\le E(X)-k)\le\exp(-\frac{2k^2}{n(b-a)^2})$

伯努利大数定律：$\lim_{n\to\infty}P(|\frac{n_A}{n}-p|<\varepsilon)=1$

马尔科夫大数定律：若 $\frac{1}{n^2}Var(\sum_{i=1}^nX_i)\to0$，

则$\{X_n\}$服从大数定律，$\lim_{n\to\infty}P(|\frac1n\sum X_i-\frac1n\sum E(X_i)|<\varepsilon)=1$

若 $X\sim N(\mu,\sigma^2)\Rightarrow\begin{cases}P(X-E(X)\ge k\sigma)\le\exp(-\frac{k^2}{2}) & \text{证：矩生成函数}\\P(X-E(X)\le k\sigma)\le\exp(-\frac{k^2}{2}) & \text{同左列}\end{cases}$

常用$M_X(t)$：$X\sim\pi(\lambda)$：$e^{\lambda(e^t-1)}$ / $X\sim N(\mu,\sigma^2)$：$e^{\mu t+\frac{\sigma^2t^2}{2}}$

$X\sim B(n,p)$：$(1-p+pe^t)^n$ / $X\sim Exp(\lambda)$：$\frac{\lambda}{\lambda-t}$ $(t<\lambda)$

$X\sim Ge(p)$：$\frac{pe^t}{1-(1-p)e^t}$ $(e^t<\frac{1}{1-p})$ / $X\sim\Gamma(\alpha,\beta)$：$(1-\frac{t}{\beta})^{-\alpha}$ $(t<\beta)$

辛钦大数定律 $\{X_n\}$ i.i.d. $\mu=E(X_i)$存在

则 $\lim_{n\to\infty}P(|\frac1n\sum X_i-\mu|<\varepsilon)=1$ $\quad\lim_{n\to\infty}P(\bigcup_{m=n}^{\infty}|X_m-X|\ge\varepsilon)=0$

依概率收敛： $\lim P(X_n=X)=1$ 次弱 p

几乎处处收敛： $P(\lim X_n=X)=1$ 强 a.s.

依分布收敛： $\lim F_n(X)=F(X)$ 最弱 d

依范数收敛： $\lim E(X-X_n)^p=0$ 强 L_p

a.s.⇒p：$\forall\varepsilon>0,\exists n,\forall k>n,|X_k-X|<\varepsilon$，此事件发生概率为1.

定义2：$\boxed{令n\to\infty,\forall\varepsilon>0,P(\bigcap_{k=n}^{\infty}|X_k-X|<\varepsilon)=1}$ F是累积分布函数

又 $1=\lim P(\bigcap_{k=n}^{\infty}|X_k-X|<\varepsilon)<\lim P(|X_n-X|<\varepsilon)$.

∴ $\lim P(X=X_n)=1$.

$L_p\Rightarrow d$：$\lim P(|X_n-X|\ge\varepsilon)\le\lim\frac{E(|X_n-X|^p)}{\varepsilon^p}=0$.

$p\Rightarrow d$：$F_n(x)=P(X_n\le x)=P(X_n\le x,X\le x+\varepsilon)+P(X_n\le x,X>x+\varepsilon)$

$\le F(x+\varepsilon)+P(|X-X_n|>\varepsilon)$

$F_X(x-\varepsilon)=P(X\le x-\varepsilon)=P(X\le x-\varepsilon,X_n\le x)+P(X\le x-\varepsilon,X_n>x)$

$\le F_n(x)+P(|X-X_n|>\varepsilon)$

∴ $F(x-\varepsilon)-P(|X-X_n|>\varepsilon)\le F_n(x)\le F(x+\varepsilon)+P(|X-X_n|>\varepsilon)$.

令$n\to\infty$. 由$F(x)$连续性，$F_n(x)\to F(x)$.

$d\Rightarrow p$：$P(X=c)=1$，即X的分布是一常数时，$d\Rightarrow p$.

$F_n(x)\to F(x)=I[X\ge c]$

$P(|X_n-c|\ge\varepsilon)\le F_n(c-\varepsilon)+1-F_n(c+\varepsilon)=0$

多维高斯分布 $X=(X_1,\cdots,X_n)$，$Cov(X)=E[(X-E(X))(X-E(X))^T]$

$=\begin{bmatrix}Var(X_1)&Cov(X_1,X_2)&\cdots\\Cov(X_2,X_1)&\cdots&\cdots\\\cdots&\cdots&Var(X_n)\end{bmatrix}$ 为半正定对称矩阵 $\alpha^TCov(X)\alpha=E[(\alpha^T(X-E(X)))^2]\ge0$

$X_1,X_2\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$ $\quad Cov(X)=B=\begin{pmatrix}\sigma_1^2&\rho\sigma_1\sigma_2\\\rho\sigma_1\sigma_2&\sigma_2^2\end{pmatrix}$

$B^{-1}=\frac{1}{1-\rho^2}\begin{pmatrix}\frac{1}{\sigma_1^2}&-\frac{\rho}{\sigma_1\sigma_2}\\-\frac{\rho}{\sigma_1\sigma_2}&\frac{1}{\sigma_2^2}\end{pmatrix}$ $\quad f(x)=\frac{1}{2\pi|B|^{\frac12}}\exp(-\frac12(x-\mu)^TB^{-1}(x-\mu))$

n维：$f(x)=\frac{1}{(2\pi)^{\frac n2}|B|^{\frac12}}\exp[-\frac12(x-\mu)^TB^{-1}(x-\mu)]$ $\quad X\sim N(\mu,B)$.

$\mu=0,B=I_n\Rightarrow f(x)=\prod_{i=1}^n\frac{1}{\sqrt{2\pi}}\exp(-\frac{(x_i-\mu_i)^2}{2})$. $X_1,\cdots X_n$相互独立.

性质：$X_i\sim N(\mu_i,B_{ii})$；若B为对角阵⇔X_i两两不相关. $X_i\sim N(0,1)$

$X\sim N(\mu,B)\Rightarrow AX+b\sim N(A\mu+b,ABA^T)$；$X\sim N(\mu,B)$ ⇓

$X\sim N(0,I_n)$，U正交，则 $UX\sim N(0,I_n)$；$B^{-\frac12}(X-\mu)\sim N(0,I_n)$

统计 k阶矩：$\frac1n\sum X_i^k$ / 中心矩：$\frac1n\sum(X_i-\bar X)^k$

无偏：$E(\hat\theta)=\theta$ / **渐进无偏**：$\lim_{n\to\infty}E(\hat\theta)=\theta$

均方误差：$MSE(\hat\theta)=E[(\hat\theta-\theta)^2]=Var(\hat\theta)+[E(\hat\theta-\theta)]^2$

一致估计：$\hat\theta_n\xrightarrow{P}\theta$，即 $\lim P(|\hat\theta_n-\theta|\ge\varepsilon)=0$

若 $MSE(\hat\theta_n)\to0\Rightarrow\hat\theta_n$为一致估计 / 一致+有界⇒渐无

最大似然估计：$\arg\max_\theta L(\theta)=P(X_1=x_1,\cdots X_n=x_n;\theta)$

不变性：给定 $\hat\theta_{MLE}$，若 $g(\theta)$有$g^{-1}(\theta)$，则 $\widehat{g(\theta)}_{MLE}=g(\hat\theta_{MLE})$

t分布 $X_1\sim N(0,1)$，$X_2\sim\chi^2(n)$. X_1,X_2独立.

$T=\frac{X_1}{\sqrt{X^2/n}}\sim t(n)$. $f(t)=\frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac n2)}(1+\frac{t^2}{n})^{-\frac{n+1}{2}}$

$f_1(t)=\frac{1}{1+t^2}\cdot\frac1\pi$. $f_\infty(t)\sim N(0,1)$.

区间估计 $\theta\in[\hat\theta_L,\hat\theta_H]$ 概率尽量大

置信区间：$P(\hat\theta_L<\theta<\hat\theta_H)\ge1-\alpha$. 可改为单侧

最小二乘估计 $\arg\min_{\alpha,\beta}Q(\alpha,\beta)=\sum(y_i-\alpha-\beta x_i)^2$

$S_{xx}=\sum(x_i-\bar x)(y_i-\bar y)$，$S_{xy}=\sum(x_i-\bar x)(y_i-\bar y)$. ∂Q 证

$\hat\beta=\frac{S_{xy}}{S_{xx}}$，$\hat\alpha=\bar y-\hat\beta\bar x$，$\sum(y_i-\hat y_i)=0$，$\sum\hat y_i(y_i-\hat y_i)=0$

$\hat\beta=\beta+\sum\varepsilon_i\cdot\frac{x_i-\bar x}{S_{xx}}$，$\hat\alpha=\alpha+\sum\varepsilon_i\cdot(\frac1n-\frac{x_i-\bar x}{S_{xx}}\cdot\bar x)$

$Cov(\hat\alpha,\hat\beta)=-\sigma^2\cdot\frac{\bar x}{S_{xx}}$ （消独立项）$E(\sum(y_i-\hat y_i)^2)=(n-2)\sigma^2$

$S^2=E(\frac{1}{n-2}\sum(y_i-\hat y_i)^2)$为$\sigma^2$无偏估计

$\sum(y_i-\bar y)^2=\sum(y_i-\hat y)^2+\sum(\bar y-\hat y_i)^2$ / $SST=SSE+SSR$

假设检验 H_0，H_1，W，σ^2是否已知

u：$\frac{\bar x-\mu_0}{\sigma/\sqrt n}\sim N(0,1)$ | $t=\frac{\bar x-\mu_0}{\sqrt{S^2}/\sqrt n}\sim t(n-1)$ | $\chi^2=\frac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$

① $H_0:\mu\le\mu_0$，$H_1:\mu>\mu_0$，$W:\{(\cdots)|u\ge c\}$ $\quad c=\Phi^{-1}(1-\alpha)$

② $H_0:\mu\ge\mu_0$，$H_1:\mu<\mu_0$，$W:\{(\cdots)|u\le c\}$ $\quad c=-\Phi^{-1}(1-\alpha)$

③ $H_0:\mu=\mu_0$，$H_1:\mu\ne\mu_0$，$W:\{(\cdots)||u|\ge c\}$ $\quad c=\Phi^{-1}(1-\frac\alpha2)$

④ $H_0:\mu=\mu_0$，$H_1:\mu\ne\mu_0$，$W:\{(\cdots)||t|\ge c\}$ $\quad c=\Phi^{-1}_{t(n-1)}(1-\frac\alpha2)$

⑤ $H_0:\sigma^2=\sigma_0^2$，$H_1:\sigma^2\ne\sigma_0^2$，$W:\{(\cdots)|\chi^2\le a或\chi^2\ge b\}$ $\begin{cases}a=F^{-1}(\frac\alpha2)\\b=F^{-1}(1-\frac\alpha2)\end{cases}$

作业 ① $X>0$. 离：$E(X)=\sum_0^{+\infty}P(X>x)$ / 连：$E(x)=\int_0^{+\infty}P(X>x)dx$

② $X\sim N(\mu,\sigma^2)$，$\bar X$与S^2独立. 证：取正交阵U，$U_{1j}=\frac{1}{\sqrt n}$

$Y=UX$，$E(Y)=(\frac{\mu}{\sqrt n},0,\cdots,0)^T$. $Cov(Y)=\sigma^2I$.

$\sum Y_i^2=\sum X_i^2$. $\bar X=Y_1/n$，$(n-1)S^2=\sum(X_i-\bar X)^2=\sum X_i^2-n\bar X^2$

$=\sum_2^nY_i^2\Rightarrow$与$Y_1(\bar X)$相互独立，$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$

③ 程序概率出错：$A_i=i$次运行返回错误结果，

$X=\sum\mathbb{1}_{A_i}$. $P(X\ge\frac n2)$. $X\sim B(n,p)$. Chernoff界.

④ 随机矩阵 $X_i\in\mathbb{R}^d$. $A\in\mathbb{R}^{k\times d}\sim N(0,1)$，$k=O(\log\frac{n}{\varepsilon^2})$.

至少1/2几率 $(1-\varepsilon)\|x_i-x_j\|^2\le\|\frac{1}{\sqrt k}A(x_i-x_j)\|^2\le(1+\varepsilon)\|x_i-x_j\|^2$.

$x\in\mathbb{R}^d$，$y_i\sim N(0,|x|_2^2)$，y_i,y_j独立. 令 $z=\frac{y}{|x|_2}\Rightarrow z\sim N(0,1)$

$|z|_2^2=\frac{|y|_2^2}{|x|_2^2}$，$E(|z|_2^2)=k$，$P(|z|_2^2\ge(1+\varepsilon)k)\le\exp(-\frac{k\varepsilon^2}{8})$. $P(|z|_2^2\le(1-\varepsilon)k)\le\exp(-\frac{k\varepsilon^2}{8})$

$P(|z|_2^2\notin(1\pm\varepsilon)k)\le2\exp(-\frac{k\varepsilon^2}{8})\Rightarrow P(|Ax|_2^2\notin(1\pm\varepsilon)k\cdot|x|_2^2)\le2\exp(-\frac{k\varepsilon^2}{8})$

$P(|\frac{1}{\sqrt k}Ax|_2^2\notin(1\pm\varepsilon)|x|_2^2)\le2\exp(-\frac{k\varepsilon^2}{8})$. $F(x)=\frac{1}{\sqrt k}Ax$，$E\doteq|F(x_i)-F(x_j)|_2^2\in\cdots$

对全部i,j成立. $P(\bar E)\le\frac{n(n-1)}{2}\cdot2\exp(-\frac{k\varepsilon^2}{8})\Rightarrow k=O(\frac{\log n}{\varepsilon^2})\Rightarrow P(E)\ge\frac12$.

⑤ 随机分组 $\forall S_{1\cdots m}\subseteq\{1,\cdots,n\}$，$\exists\chi:\{1,\cdots,n\}\to\{-1,+1\}$，s.t. $\forall\ 1\le i\le m$，

$disc_\chi(S_i)=|\sum_{j\in S_i}\chi(j)|\le O(\sqrt{n\log m})$. 证：$\forall i$. $E(\sum_{j\in S_i}\chi(j))=\sum_{j\in S_i}E(\chi(j))=0$

$P(\sum_{j\in S_i}\chi(j)\ge k)\le\exp(-\frac{k^2}{2n})$；$P(\sum_{j\in S_i}\chi(j)\le-k)\le\exp(-\frac{k^2}{2n})\Rightarrow P(disc_\chi(S_i)\ge k)\le2\cdot e^{-\frac{k^2}{2n}}$

$P(\max\{disc_\chi(S_i)\}\ge k)\le2m\cdot\exp(-\frac{k^2}{2n})$. $\Rightarrow k=O(\sqrt{n\log m})$，$RHS=\frac12$.

⑥ 老虎机 n台，t轮 $X_t\sim B(1,p_i)$. 求 $o\in\{1,\cdots,n\}$，$P(p_o\ge\max_ip_i-\varepsilon)\ge\frac23$.

每台条N次，$P(|\bar X_i-p_i|\ge\varepsilon)\le2\exp(-2N\varepsilon^2)\Rightarrow\varepsilon=\sqrt{\ln\frac2\alpha/2N}$

$P(|\bar X_i-p_i|\le\varepsilon)\ge1-\alpha$. 取 $\alpha=\frac{1}{3n}$，$N=O(\frac{\ln n}{\varepsilon^2})$

$P(|p_i-\bar X_i|\le\varepsilon)\ge1-\frac{1}{3n}$. $P(\forall i,|p_i-\bar X_i|\le\varepsilon)\ge\frac23$.

$p_o\ge\bar X_o-\varepsilon\ge\bar X_i-\varepsilon\ge p_i-2\varepsilon$.

欲 $P(p_o\ge\max_ip_i-\varepsilon)\ge1-\delta$：取 $\alpha=\frac\delta n$，$N=O(\frac{\ln\frac n\delta}{\varepsilon^2})$

则 $P(|p_i-\bar X_i|\le\varepsilon)\ge1-\frac\delta n$. $P(\forall i,|p_i-\bar X_i|\le\varepsilon)\ge1-\delta$.

⇒⑦ 给定$\{(x_i,y_i)\}$. $y_i=\alpha+\beta x_i+\varepsilon_i$，$\varepsilon_i\sim N(0,\sigma^2)$，$S^2=\frac{1}{n-2}\sum(y_i-\hat y_i)^2$.

Q为正交阵，$q_1=[\frac{1}{\sqrt n},\cdots\frac{1}{\sqrt n}]^T$，$q_2=[\frac{x_1-\bar x}{\sqrt{S_{xx}}},\cdots\frac{x_n-\bar x}{\sqrt{S_{xx}}}]^T$. q_{3^+}由Schmidt得到

$E(y_i)=\alpha+\beta x_i$，$Cov(z)=Q\sigma^2IQ^T=\sigma^2I\Rightarrow$各维独立$\Rightarrow z\sim n$维高斯

$z_1=\frac{1}{\sqrt n}\sum y_i=\sqrt n\bar y$，$z_2=\frac{1}{\sqrt{S_{xx}}}\sum(x_i-\bar x)y_i=\frac{S_{xy}}{\sqrt{S_{xx}}}=\sqrt{S_{xx}}\hat\beta$，$|z|_2^2=|y|_2^2$

$\sum(y_i-\bar y)^2=\sum(y_i-\hat y)^2+\sum(\bar y-\hat y_i)^2\Rightarrow\sum(y_i-\hat y)^2=\sum_{i=3}^nz_i^2\Rightarrow S^2$与$\hat\alpha,\hat\beta$独立

$\underbrace{\sum y_i^2-n\bar y^2}=\sum y_i^2-z_1^2$. $\quad\underbrace{\hat\beta(x_i-\bar x)}=z_2^2$ $\quad\begin{cases}q_i^Tq_2=0\Rightarrow q_i^Tx_j=0\\q_i^Tq_1=0\Rightarrow q_i^Tc=0\end{cases}$

$E(z_i)=E(q_i^T(\alpha+\beta x_j+\varepsilon_j))$

$=E(q_i^T\varepsilon_j)=0\quad(i\ge3)$.

$E(z)=[\sqrt n\bar y,\sqrt{S_{xx}}\hat\beta,0,\cdots,0]^T$

$Var(z_i)=Var(\sum_jq_{ij}(\alpha+\beta x+\varepsilon))=\sigma^2\cdot|q_i|_2^2=\sigma^2$. $\Rightarrow z_i\sim N(0,\sigma^2)$

∴ $\frac{(n-2)S^2}{\sigma^2}\sim\chi^2(n-2)$.

⑧ $X_{1\cdots2023},Y_{1\cdots2024}\sim N(0,1)$. $Q=\sum^{2023}(X_i^2+Y_i^2)-2023\bar X^2-2024\bar Y^2+Y_{2024}^2$

$\sum X_i^2=\sum(X_i-\bar X+\bar X)^2=\sum(X_i-\bar X)^2+2023\bar X^2$. $Q=\underbrace{\sum^{2023}(X_i-\bar X)^2}_{\sim\chi^2(2022)}+\underbrace{\sum^{2024}(Y_i-\bar Y)^2}_{\sim\chi^2(2023)}\sim\chi(4045)$

⑨ $X\sim U(0,\theta)$. $\hat\theta=\max\{X_i\}$ $\frac{n+1}{n}$ MSE?

$\theta_1=\max\{\cdots\}$，$F_{\theta_1}(x)=[F_X(x)]^n=(\frac x\theta)^n$，$E(\theta_1)=\int_0^\theta1-(\frac x\theta)^ndx=\frac{n}{n+1}\theta$

$F_{\theta_1^2}(x)=F_{\theta_1}(\sqrt x)$，$0\le x\le\theta^2$，$E(\theta_1^2)=\int_0^{\theta^2}1-(\frac{\sqrt x}{\theta})^ndx=\frac{n}{n+2}\theta^2$

$\Rightarrow Var(\theta_1)=[\frac{n}{n+2}-(\frac{n}{n+1})^2]\theta^2$. $MSE=(\frac{n+1}{n})^2Var(\theta_1)=\frac{\theta^2}{n(n+2)}$