Distributive law:

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

De Morgan's law:

- $(A\cup B)'=A'\cap B'$ • $(A\cap B)'=A'\cup B'$

permutation: ${}_nP_r=n(n-1)\cdots(n-r+1)=\frac{n!}{(n-r)!}$

combination: ${}_nC_r=\binom{n}{r}=\frac{n!}{r!(n-r)!}={}_nC_{n-r}=\frac{{}_nP_r}{r!}$

Conditional Probability: $P(A|B)=\frac{P(A\cap B)}{P(B)}$

$P(A\cap B\cap C)=P(A\cap B)P(C|A\cap B)=P(A)P(B|A)P(C|A\cap B)$

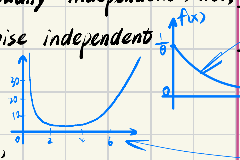
Independent Events: $P(A\cap B)=P(A)P(B)$

$\Rightarrow P(A|B)=P(A),\ P(B|A)=P(B)$

$\Rightarrow$ A and B', A' and B, A' and B' ✓

Events A, B, C are mutually independent, i.e.,

1. A, B, C are pairwise independent
$$\begin{cases} P(A\cap B)=P(A)P(B)\\ P(A\cap C)=P(A)P(C)\\ P(B\cap C)=P(B)P(C)\end{cases}$$
2. $P(A\cap B\cap C)=P(A)P(B)P(C)$



total probability

Bayes' Theorem:

(a) For any event A, $P(A)=\sum_{i=1}^{m}P(A\cap B_i)=\sum_{i=1}^{m}P(B_i)P(A|B_i)$

(b) If $P(A)>0$, then $P(B_k|A)=\frac{P(B_k\cap A)}{P(A)}$, $k=1,\dots,m$

$P(B_k|A)=\frac{P(B_k)P(A|B_k)}{P(A)=\sum_{i=1}^{m}P(B_i)P(A|B_i)}$ $\longrightarrow$ Bayes Theorem

$P(B_k)$: prior, $P(B_k|A)$: posterior probability 为后验概率

$P(A|B_k)$: likelihood of B_k, A is a data

$f(x): \bar{S}\to(0,1]$

Probability Mass Function (pmf):

$$\begin{cases} 1.\ f(x)>0,\ x\in\bar{S}\\ 2.\ \sum_{x\in\bar{S}}f(x)=1\\ 3.\ P(X\in A)=\sum_{x\in A}f(x),\ A\subseteq\bar{S}\end{cases}$$

We often extend the domain of $f(x)$ from $\bar{S}$ to R and let $f(x)=0$, $x\notin\bar{S}$. In this case, $\bar{S}$ is called the support of $f(x)$.

$f(x)=c$ for $x\in\bar{S}$ $\Rightarrow$ uniform distribution

Cumulative Distribution Function (cdf):

$F(x): R\to[0,1]$: $F(x)=P(X\le x)$

1. nondecreasing, $P(X\le x)=\sum_{x'\le x,\,x'\in\bar{S}}f(x')$
2. $P(a<X\le b)=F(b)-F(a)$

Mathematical Expectation: $E[g(x)]=\sum_{x\in\bar{S}}g(x)f(x)$

(1) $E(c)=c$ (2) $E[cg(x)]=cE[g(x)]$

(3) $E[c_1g_1(x)+c_2g_2(x)]=c_1E[g_1(x)]+c_2E[g_2(x)]$ (linear operator)

Mean of a RV $[g(x)=x]$:

the average value of X $E(x)=\sum_{x\in\bar{S}}xf(x)\overset{\bar{S}=\{x_1,\dots,x_k\}}{=}\sum_{i=1}^{k}x_if(x_i)$

$=\int_{\bar{S}}xf(x)dx$

Variance of a RV $[g(x)=(X-E[x])^2]$: second, first moment

$Var(X)=E[(X-E[x])^2]=\sum_{x\in\bar{S}}(x-E[x])^2f(x)=E(x^2)-(E[x])^2$

$=\int_{\bar{S}}(x-E[x])^2f(x)dx$

Standard deviation of a RV: $\sqrt{Var(X)}$

(1) $Var(c)=0$ (2) $Var(cX)=c^2Var(X)$

The r th Moment [$g(x)=x^r$ with r a positive integer]:

$E(X^r)=\sum_{x\in\bar{S}}x^rf(x)\Longrightarrow$ the r th moment of X $=\int_{\bar{S}}x^rf(x)dx$

$E[(X-b)^r]=\sum_{x\in\bar{S}}(x-b)^rf(x)\Longrightarrow$ the r th moment of X about b

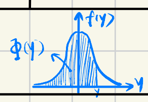
$E[(X)_r]=E[X(X-1)\cdots(X-r+1)]\Longrightarrow$ the r th factorial moment

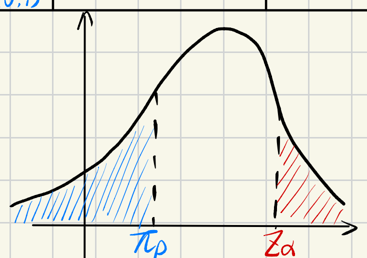
Moment Generating Function (mgf): (for $-h<t<h$)

$M(t)=E[e^{tX}]=\sum_{x\in\bar{S}}e^{tx}f(x)=\int_{\bar{S}}e^{tx}f(x)dx$

1. $M(0)=1$
2. two RVs, same mgf, same probability distribution, same pmf
3. $M'(t)=\sum_{x\in\bar{S}}xe^{tx}f(x)$ $M''(t)=\sum_{x\in\bar{S}}x^2e^{tx}f(x)$ $M^{(r)}(t)=\sum_{x\in\bar{S}}x^re^{tx}f(x)$

$M'(0)=E(x)$ $M''(0)=E(x^2)$ $M^{(r)}(0)=E[x^r]$

	pmf/pdf	mgf $M(t)$	mean	variance	cdf	解释
Bernoulli	$f(x):\bar{S}\to[0,1]$ $f(x)=p^x(1-p)^{1-x},\ x\in\bar{S}$	$p\cdot e^t+(1-p)$ $t\in(-\infty,\infty)$	p	$(1-p)p$		RV: $X: S\to\bar{S}$, $S=\{$success, failure$\}$ $X(✓)=1$ $X(✗)=0$ $\bar{S}=[0,1]$
Binomial $X\sim b(n,p)$	$f(x)=\binom{n}{x}p^x(1-p)^{n-x}$ $x=0,1,\dots,n$	$[(1-p)+pe^t]^n$ $t\in(-\infty,\infty)$	np	$np(1-p)$	$F(x)=P(X\le x)=\sum_{y\in\{X\le x\}}f(y)=\sum_{y=0}^{[x]}\binom{n}{y}p^y(1-p)^{n-y}$ $x\in(-\infty,\infty)$ $[x]:\le x$	n Bernoulli trials 中成功总次数
Negative Binomial	$f(x)=\binom{x-1}{r-1}p^r(1-p)^{x-r}$ $x=r,r+1,\dots$ $(1-w)^{-r}=\sum_{x=r}^{\infty}\binom{x-1}{r-1}w^{x-r}$	$\frac{(pe^t)^r}{[1-(1-p)e^t]^r}$ for $(1-p)e^t<1$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$		对于给定的 r，第 r 次 Benoulli 成功的次数
Geometric	$f(x)=p(1-p)^{x-1}$ $r=1$	$P(X>k)=\sum_{x=k+1}^{\infty}p(1-p)^{x-1}=(1-p)^k$; $P(X\le k)=\sum_{x=1}^{k}p(1-p)^{x-1}=1-P(X>k)=1-(1-p)^k$				
Poisson	$f(x)=\frac{\lambda^xe^{-\lambda}}{x!}$ $x=0,1,\dots$ $X\sim Poisson(\lambda)$	$e^{\lambda(e^t-1)}$	λ	λ		特定事件在给定时间间隔内发生的次数
Uniform $X\sim U(a,b)$	$f(x)=\begin{cases}\frac{1}{b-a}, & a\le x\le b\\ 0, & \text{otherwise}\end{cases}$	略 $\begin{cases}\frac{e^{tb}-e^{ta}}{t(b-a)}, & t\ne0\\ 1, & t=0\end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$F(x)=\begin{cases}0, & x<a\\ \frac{x-a}{b-a}, & a\le x\le b\\ 1, & x>b\end{cases}$	
Exponential	$f(x)=\frac{1}{\theta}e^{-\frac{x}{\theta}}$, $x\ge0,\theta>0$	$\frac{1}{1-t\theta}$, $t<\frac{1}{\theta}$	θ	θ^2		
Gamma $\alpha=1$ $w\Leftrightarrow x$ $\theta\Leftrightarrow\frac{1}{x}$ $\theta=2,\alpha=\frac{r}{2}$	$f(x)=\frac{1}{\Gamma(\alpha)\theta^\alpha}x^{\alpha-1}e^{-\frac{x}{\theta}}$ $\Gamma(n)=(n-1)!$ $x\ge0,\alpha>0,\theta>0$	$\frac{1}{(1-\theta t)^\alpha}$ $t<\frac{1}{\theta}$	$\alpha\theta$	$\alpha\theta^2$		某特定事件第α次发生 APP等待时间
Chi-square $X\sim\chi^2(r)$	$f(x)=\frac{1}{\Gamma(\frac{r}{2})2^{\frac{r}{2}}}x^{\frac{r}{2}-1}e^{-\frac{x}{2}}$ $x>0$	$(1-2t)^{-\frac{r}{2}}$ $t<\frac{1}{2}$	r	$2r$		
Normal $X\sim N(\mu,\sigma^2)$	$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}exp(-\frac{1}{2}\cdot\frac{(x-\mu)^2}{\sigma^2})$ $-\infty<x<\infty$	$exp(\mu t+\frac{1}{2}\sigma^2t^2)$	μ	σ^2		(Gaussian)
Standard Normal $Y\sim N(0,1)$	$f(y)=\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}$	$\Phi(-y)=1-\Phi(y)$			$\Phi(y)=P(Y\le y)$ $=\int_{-\infty}^{y}f(z)dz=\int_{-\infty}^{y}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz$	



$P(X\le\pi_p)=p$

$(100p)$th percentile

$P(X\ge z_\alpha)=\alpha$

The upper 100α percent point

the $100(1-\alpha)$th percentile

$p=\int_{-\infty}^{\pi_p}f(x)dx=F(\pi_p)$

25th / first

75th / third

50th / median, 2nd quantile

Normal Distribution 与其他 Distribution 的转化

1. If Y is $N(\mu,\sigma^2)$, then $X=\frac{Y-\mu}{\sigma}$ is $N(0,1)$

X has the same cdf as $N(0,1)$

证: $P(X\le x)=P(\frac{Y-\mu}{\sigma}\le x)=P(Y\le\sigma x+\mu)$

$=\int_{-\infty}^{\sigma x+\mu}f(y)dy=\int_{-\infty}^{\sigma x+\mu}\frac{1}{\sqrt{2\pi\sigma^2}}exp(-\frac{1}{2}(\frac{(y-\mu)^2}{\sigma^2}))dy$

$\overset{w=\frac{y-\mu}{\sigma}}{=}\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}exp(-\frac{1}{2}w^2)dw\Rightarrow$ cdf of $N(0,1)$

$\Rightarrow P(a\le X\le b)=P(\frac{a-\mu}{\sigma}\le\frac{X-\mu}{\sigma}\le\frac{b-\mu}{\sigma})$

$=\Phi(\frac{b-\mu}{\sigma})-\Phi(\frac{a-\mu}{\sigma})$

2. If X is $N(\mu,\sigma^2)$ with $\sigma^2>0$, then $\frac{(X-\mu)^2}{\sigma^2}\sim\chi^2(1)$

证: Let $V=\frac{(X-\mu)^2}{\sigma^2}$. Then consider the cdf of V,

$G(v)=P(V\le v)=P(-\sqrt{v}\le Z\le\sqrt{v})$,

where $Z=\frac{X-\mu}{\sigma}$, with $v\ge0$

$G(v)=\int_{-\sqrt{v}}^{\sqrt{v}}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz=2\int_0^{\sqrt{v}}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz$

$\overset{z=\sqrt{y},\ \frac{dz}{dy}=\frac{1}{2\sqrt{y}}}{=}2\int_0^{v}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y}\frac{1}{2\sqrt{y}}dy=\int_0^{v}\frac{1}{\sqrt{2\pi y}}e^{-\frac{y}{2}}dy$, $v\ge0$

$g(v)=G'(v)=\frac{1}{\sqrt{2\pi}}v^{\frac{1}{2}-1}e^{-\frac{1}{2}v}$, $v\ge0$

Now recall the pdf of $\chi^2(1)$:

$f(x)=\frac{1}{\Gamma(\frac{1}{2})2^{\frac{1}{2}}}x^{\frac{1}{2}-1}e^{-\frac{1}{2}x}$, $x\ge0$

Since $g(v)$ is a pdf, then $\int_0^\infty g(v)dv=1$

$1=\int_0^\infty\frac{1}{\sqrt{2\pi}}v^{\frac{1}{2}-1}e^{-\frac{1}{2}v}dv\overset{x=\frac{1}{2}v}{=}\frac{1}{\sqrt{\pi}}\int_0^\infty x^{\frac{1}{2}-1}e^{-x}dx$

$=\frac{1}{\sqrt{\pi}}\Gamma(\frac{1}{2})\Longrightarrow\Gamma(\frac{1}{2})=\sqrt{\pi}$

$\Longrightarrow g(v)=\frac{1}{\Gamma(\frac{1}{2})2^{\frac{1}{2}}}v^{\frac{1}{2}-1}e^{-\frac{v}{2}}$, $v>0$

Normal Distribution 的详细内容

$M(t)=E[e^{tX}]=\int_{-\infty}^{\infty}e^{tx}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}dx$

$e^{tx}e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}=exp\{-\frac{1}{2\sigma^2}[x^2-2(\mu+\sigma^2t)x+\mu^2]\}$

Consider $x^2-2(\mu+\sigma^2t)x+\mu^2=[x-(\mu+\sigma^2t)]^2-2\mu\sigma^2t-\sigma^4t^2$

$\therefore M(t)=exp(\frac{-2\mu\sigma^2t-\sigma^4t^2}{-2\sigma^2})\cdot\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma^2}}exp(-\frac{1}{2\sigma^2}[x-(\mu+\sigma^2t)]^2)dx$

Recall that $I=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma^2}}exp[-\frac{(x-\mu)^2}{2\sigma^2}]dx=1$, independent of μ

Therefore, $\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma^2}}exp(-\frac{1}{2\sigma^2}[x-(\mu+\sigma^2t)]^2)dx=1$

$\therefore M(t)=\cdots$ ✓ $M(0)=1$

$M'(t)=(\mu+\sigma^2t)exp(\mu t+\frac{1}{2}\sigma^2t^2)\Longrightarrow M'(0)=\mu$

$M''(t)=\sigma^2exp(\mu t+\frac{1}{2}\sigma^2t^2)+(\mu+\sigma^2t)^2exp(\mu t+\frac{1}{2}\sigma^2t^2)$

$\Longrightarrow M''(0)=\mu^2+\sigma^2$

Recall that $E[x]=M'(0)=\mu$,

$Var[x]=E[x^2]-(E[x])^2=M''(0)-M'(0)^2=\sigma^2$

For $X\sim N(\mu,\sigma^2)$, $E[x]=\mu$, $Var[x]=\sigma^2$

$\therefore$ The two parameters μ, σ^2 are the mean and variance, respectively.

证 Gamma Distribution $\int pdf=1$

Let $t=\frac{x}{\theta}$ $\Longrightarrow x=\theta t$, $dx=\theta dt$

(原式) $=\int_0^\infty\frac{1}{\Gamma(\alpha)\theta^\alpha}(\theta t)^{\alpha-1}e^{-t}\theta dt$

$=\frac{1}{\Gamma(\alpha)\theta^\alpha}\cdot\theta^{\alpha-1}\cdot\theta\int_0^\infty t^{\alpha-1}e^{-t}dt$

$=\frac{1}{\Gamma(\alpha)}\cdot\Gamma(\alpha)=1$

Gamma Function (generalized factorial):

$\Gamma(t)=\int_0^\infty y^{t-1}e^{-y}dy=-y^{t-1}e^{-y}\Big|_0^\infty+\int_0^\infty(t-1)y^{t-2}e^{-y}dy$

$=(t-1)\Gamma(t-1)$

排列组合

1. 抽样（摸球）：从 n 张卡片中抽取 r 张

可能结果总数	放回	不放回
有序	n^r	P_n^r
无序	$\binom{n+r-1}{r}$	$\binom{n}{r}$

2. 分配（占位）：把 r 只球投入 n 个房间

可能结果总数	不限制	有限制（只限一球）
可分辨（编号）	n^r	P_n^r
不可分辨	$\binom{n+r-1}{r}$	$\binom{n}{r}$

3. 从装有编号为 $1\sim n$ 的球的袋子中有放回地摸出 r 只球，依次记下其号码，可能总数为 n^r

4. n 个人的生日共有 365^n 种

5. 公共汽车上的 r 位乘客在以后的 n 个站下车的方式有 n^r 种

6. n 个可分辨粒子在相空间的 N 个相格中的可能分配有 N^n 种

全排列 $n\cdot(n-1)\cdots 3\cdot 2\cdot 1$

7. n 卷文集在书架上的排列方式有 $n!$ 种

8. 把 n 个客人分到 n 个房间共有 $n!$ 种安排法

9. 把编号为 $1\sim N$ 的球从袋中一只接一只摸出依序排成一列，共有 $N!$ 种不同结果

10. 把 n 封信分别装入 n 只信封共有 $n!$ 种结果

选排列 $P_n^r=n(n-1)\cdots(n-r+1)=\frac{n!}{(n-r)!}$，$1\le r\le n$

11. 从 n 种商品中选 r 种依序上货架，共有 P_n^r 种不同排列

12. 把 r 个客人分到 $n(n\ge r)$ 个房间，共有 P_n^r 种安排

重复组合数（r 个元素中可以有相同的）：$\binom{n+r-1}{r}$

总结：

可能结果总数	重复	不重复
排列	n^r	P_n^r
组合	$\binom{n+r-1}{r}$	$\binom{n}{r}$

n 只球在 n 个位置的全排列

(1) 1 号球出现在 1 号位的概率为 $\frac{1}{n}$，由对称性，1 号球出现在第 k 号位的概率也为 $\frac{1}{n}$。更进一步，第 i 号球出现在第 k 号位的概率也为 $\frac{1}{n}$。这是不放回开锁问题的答案。

(2) k 号位由预先指定的某号球占据的概率：$\frac{1}{n}$

(3) 袋中有 a 只黑球和 b 只白球，从中一只只摸出球，则第 k 次摸得黑球的概率为 $\frac{a}{a+b}$，

与顺序无关　　这是抽签问题的答案

(4) 在 n 封信与 n 只信封的匹配中，1 号信装入 1 号信封的概率为 $\frac{1}{n}$，由对称性，i 号信装入 i 号信封的概率也为 $\frac{1}{n}$

1，2 号信与信封符合的概率为 $\frac{1}{n}\cdot\frac{1}{n-1}$

由对称性，第 i，j 号信与信封符合的概率也为 $\frac{1}{n}\cdot\frac{1}{n-1}$

1，2，3 号信与信封符合的概率为 $\frac{1}{n}\cdot\frac{1}{n-1}\cdot\frac{1}{n-2}$

由对称性，第 i，j，k 号信与信封符合的概率也为 $\frac{1}{n}\cdot\frac{1}{n-1}\cdot\frac{1}{n-2}$

抽样调查

(1) 从 N 张卡片中任取一张，1 号或 k 号卡片被抽到的概率均为 $\frac{1}{N}$

(2) 从 N 张卡片中任取两张，1，2 号卡片被抽到的概率为 $\frac{1}{N}\cdot\frac{1}{N-1}\times 2$，由对称性，预先指定的 k 号与 j 号卡片被抽到的概率亦为 $\frac{1}{\binom{N}{2}}$

(3) 从 N 张卡片中任取 ℓ 张，预先指定的 $i_1, i_2, \cdots, i_\ell$ 号卡片被抽到的概率为

$$\frac{1}{\binom{N}{\ell}}=\frac{1}{N(N-1)\cdots(N-\ell+1)},$$

这也表明，在抽样中，ℓ 张卡片被一张张抽出与一次性抽出 ℓ 张卡片是等价的

(4) 袋中有 a 只黑球 b 只白球，从中摸出 n 只有 k 只黑球的概率为 $\frac{\binom{a}{k}\binom{b}{n-k}}{\binom{a+b}{n}}$

超几何分布

(5) 甲袋中有 a 只黑球 b 只白球，乙袋和丙袋都是空袋，先从甲袋中随机摸 $n\ (1\le n\le a+b)$ 只球放入乙袋，再从乙袋中随机摸 $m\ (1\le m\le n)$ 只球放入丙袋，最后从丙袋中任取一球，求该球为黑球的概率

$$\frac{\binom{a}{k}\binom{b}{n-k}}{\binom{a+b}{n}}\quad\frac{\binom{k}{i}\binom{n-k}{m-i}}{\binom{n}{m}}$$

多项系数 $\binom{n}{r_1, r_2, \cdots, r_k}=\frac{n!}{r_1!r_2!\cdots r_k!}$

$r_1+r_2+\cdots+r_k=n$　　当 $k=2$ 时化为二项系数

1、（抽样）从 n 个不同的元素中取 r_1 个构成第 1 组，r_2 个构成第 2 组，……，剩下的 r_k 个构成第 k 组，计有 $\frac{n!}{r_1!r_2!\cdots r_k!}$ 种

2、（分割）把 n 个物体分为 k 类，第一类 r_1 个，第二类 r_2 个，……，第 k 类 r_k 个，这里 $r_1+r_2+\cdots+r_k=n$，则分法共有 $\frac{n!}{r_1!r_2!\cdots r_k!}$ 种

k 类元素的抽样数

1、袋中有 M 只黑球，$(N-M)$ 只白球，从中摸出 n 只，则有 k 只黑球、$(n-k)$ 只白球的构成方法计有 $\binom{M}{k}\binom{N-M}{n-k}$ 种

2、袋中装有 i 号球 N_i 只，$i=1,2,\cdots,k$，$N_1+N_2+\cdots+N_k=N$，从中摸出 n 只，则有 i 号球 n_i 只，$i=1,2,\cdots,k$，$n_1+n_2+\cdots+n_k=n$ 的构成方法有 $\binom{N_1}{n_1}\binom{N_2}{n_2}\cdots\binom{N_k}{n_k}$ 种

3. 若 $r_1+r_2+\cdots+r_k=n$，则

$$\binom{n}{r_1}\binom{n-r_1}{r_2}\binom{n-r_1-r_2}{r_3}\cdots\binom{n-r_1-r_2-\cdots-r_{k-1}}{r_k}=\frac{n!}{r_1!r_2!\cdots r_k!}$$