Part I General Theory

Basic Law

样本空间 Ω (outcome space)，事件 A ，σ代数 Σ (if $A\in\Sigma \wedge B\in\Sigma$ then $A\cup B, A\cap B\in\Sigma$)

Kolmogorov Anxioms :

For given Ω and Σ，如果Pr有如下性质：

1、$\forall$ event $A\in\Sigma$，$Pr(A)$ 有定义 —— 分配概率

2、$Pr(\phi)=0 \wedge Pr(\Omega)=1$

3、$\{A_i\}$有限或可数，其中集合互不相交，为Σ中元素。Then $Pr(\cup_i A_i)=\Sigma_i Pr(A_i)$ significance

$\Rightarrow (\Omega, \Sigma, Prob)$ 为一概率空间

More details :

Banach-Tarski 悖论 (分球) — 取决于 Axiom of Choice . Then 体积难以定义

Advanced : σ代数及其度量

不可数 Ω 下如何定义 event ? Ω 的一个 σ-Algebra Σ 应满足：

properties : (1) $A\in\Sigma \to A^c\in\Sigma$ (2) Σ子集的可数并仍属于Σ，$\cup_{i=1}^{\infty} A_i\in\Sigma$

Thm. $\phi, \Omega\in\Sigma$，Σ对可数交封闭

Σ中的元素为 event significant ☆

拓展：

记 $x\sim y$ 为 $x-y\in\mathbb{Q}$，$[x]=\{y\in[0,1]: y\sim x\}$. By AC，从$[0,1]$等价类中选取 $\to N$

记$N_r=\{x+r: x\in N\}\cap[0,1]$，$r\in\mathbb{Q}$. Then $[0,1]=\cup_{r\in\mathbb{Q}} N_r$

which means : 无法对 N 分配概率

Cond Prob, Indep and Bayes' Thm

用条件概率 or 用已知信息、$Pr(A|B)$ 独立 Note that : 两两独立并不意味着 all events 相互独立

Bayes' Thm : $Pr(A|B)=\frac{Pr(B|A)Pr(A)}{Pr(B)}$ significant

Defination : 划分，全概率法则 $Pr(A)=\Sigma_i Pr(A|B_i)Pr(B_i)$

Thm. $Pr(A|B)=\frac{Pr(B|A)Pr(A)}{\Sigma_{i=1}^{n} Pr(B|A_i)Pr(A_i)}$，where $\{A_i\}_{i=1}^{n}$ 为Ω的划分

Some methods in Combinatorics : proof by story

Example 解的个数：$2x_1+2x_2+3x_3+3x_4=1996$ 非负整数 $\sum_{y=0}^{\lfloor 1996/6\rfloor}(2y+1)(999-3y)$

Part II Random Variables

Discrete Random Variables

Defination : Discrete random variable $X:\Omega\to\mathbb{R}$。要求Ω有限 or 可数

PDF (Probability density function) $f_X(x)=Prob(w\in\Omega: X(w)=x)$

CDF (cumulative distribution function) $F_X(x)=Prob(w\in\Omega: X(w)\le x)$

Continuous (how to define PDF & CDF)

Defination : PDF $f_X(x)$ 满足 (1)分段连续 or可积 (2) ≥ 0 (3) $\int_{-\infty}^{\infty} f_X(x)dx=1$ "标准化"

CDF $F_X(x)=\int_{-\infty}^{x} f_X(t)dt$

Expectation

$E[g(X)]=\int_{-\infty}^{+\infty} g(x)f_X(x)dx$，当$g(x)=x^r$时，$E[X^r]$称为r阶矩；$E[(X-E[X])^r]$称为X的r阶中心矩

$\mu_X=E[X]=\int_{-\infty}^{+\infty} x f_X(x)dx$，$\sigma^2=Var(X)=E[(X-E[X])^2]=E[X^2]-E[X]^2=\int_{-\infty}^{+\infty}(x-\mu)f_X(x)dx$

$\mu - f(0)$，$\sigma^2 - f'(0)$

Joint Distribution :

Def. 联合概率密度函数

Random variables $X_1, X_2, \cdots, X_n$ with f_{X_i}。Then $(X_1,\cdots,X_n)$ 联合密度函数 $f_{X_1\cdots X_n}$ 满足：

$Prob((X_1,\cdots,X_n)\subset S)=\int\cdots\int_S f_{X_1\cdots X_n}(x_1,\cdots,x_n)dx_1\cdots dx_n$ ($S\subset\mathbb{R}^n$)

且

$f_{X_i}(x_i)=\underbrace{\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}}_{\{x_1\cdots x_n\}\setminus\{x_i\}} f(x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_n)\prod_{j\ne i} dx_j$ 边缘概率密度函数 (marginal) 对其余变量积分

相互独立 iff $f_{X_1\cdots X_n}(x_1,\cdots,x_n)=\prod_{i=1}^{n} f_{X_i}(x_i)$

Linearity of Expectation (The process of proof is lengthy but obvious) $\forall X, Y$，$E[aX+bY]=aE[X]+bE[Y]$

Thm. X、Y独立，then $E[XY]=E[X]E[Y]$，$\sigma_{X+Y}^2=\sigma_X^2+\sigma_Y^2$ $\longrightarrow$ investment portfolios & risk

Covariance (协方差) $\sigma_{XY}=E[(X-\mu_X)(Y-\mu_Y)]$ Then $Var(\Sigma X_i)=\Sigma Var(X_i)+2\sum_{i<j}Cov(X_i, X_j)=\sum_{1\le i,j\le n}Cov(X_i, X_j)$

Higher moment : 3阶中心矩 skewness (偏斜度)，4阶 kurtosis 峰度

相关系数：$\rho=\frac{Cov(X,Y)}{\sigma_X\sigma_Y}\in[-1,1]$ (Cauchy-S)

$Cov(X,Y)=E[XY]-\mu_X\mu_Y$

Convolutions

$(f*g)(z)\triangleq\int_{-\infty}^{+\infty} f(t)g(z-t)dt$ or $\Sigma f(x_n)g(z-x_n)$，注：除非$z-x_n=y_m$，否则$f(z-x_n)=0$

Thm X、Y独立，$Z=X+Y$，then $f_Z(z)=(f_X * f_Y)(z)$ (using CDF)

Changing Variables

$y=g(x)$，$h=g^{-1}$ Then $f_Y(y)=f_X(x)\left|\frac{dx}{dy}\right|=f_X(h(y))\,|h'(y)|$ $h'(y)=\frac{1}{g'(h(y))}$ 要求g'不变号(有限个0)

concise proof : $F_Y(y)=Pr(Y\le y)=Pr(g(X)\le y)=Pr(X\le h(y))$ then 求导 直观：$f_Y(y)dy=f_X(x)dx$

$Z=XY$，X、Y独立且非负。取log / 直接证明 $\to f_Z(z)=\int_0^{+\infty} f_X(t)f_Y(\frac{z}{t})\frac{dt}{t}$

$Z=X/Y$ $f_Z(z)=z^{-2}\int_0^{+\infty} f_X(t)f_Y(t/z)t\,dt$

Tricks : 用级数可交换次序求 μ 和 σ^2

Part III Distribution

Discrete Distribution

Bernoulli $X\sim Bern(p)$: $P(X=1)=p$ $P(X=0)=1-p$，then $\mu=p$ $\sigma^2=p(1-p)$

Binomial $X\sim Bin(n,p)$: $P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$，then $\mu=np$，$\sigma^2=np(1-p)$

$X\sim Mutinomial(n,k,p_1,\cdots,p_k)$: $\binom{n}{x_1,x_2,\cdots,x_k}p_1^{x_1}\cdots p_k^{x_k}$ 其中$\Sigma x_i=n$

Geometric $X\sim Geom(p)$: $P(X=n)=p(1-p)^{n-1}$，then $\mu=\frac{1}{p}$，$\sigma^2=\frac{1-p}{p^2}$

Negative Binomial $X\sim NegBin(r,p)$: $P(X=k)=\binom{k+r-1}{k}p^k(1-p)^r$，then $\mu=\frac{pr}{1-p}$，$\sigma^2=\frac{pr}{(1-p)^2}$ (r次失败时已成功k次)

Poisson $X\sim Pois(\lambda)$: $P(X=n)=e^{-\lambda}\frac{\lambda^n}{n!}$，then $\mu=\lambda$，$\sigma^2=\lambda$

Uniform 均匀分布 "德国坦克"

Continuous

Uniform $X\sim Unif(a,b)$，$\sigma^2=\frac{(b-a)^2}{12}$ 卷积结果 $Z=X+Y$ $f_Z(z)=\begin{cases} z, & 0\sim1 \\ 1-z, & 1\sim2 \\ 0, & \text{other}\end{cases}$

$[0,1]$随机生成：$x=\Sigma a_i/2^i$ where $a_i\in\{0,1\}$

Exponential $X\sim Exp(\lambda)$，$f_X(x)=\lambda^{-1}e^{-\frac{x}{\lambda}}$，$\mu=\lambda$，$\sigma^2=\lambda^2$

$X=X_1+\cdots+X_n$，then $f_X(x)=\frac{x^{n-1}e^{-x/\lambda}}{\lambda^n(n-1)!}$ with $n\lambda$ and $n\lambda^2$，which is Erlang Distribution

Cumulative distribution method for generating random variables

$X: f_X(x), F_X(x)$，$Y\sim Unif(0,1)$ then we could get X by $X=F_X^{-1}(Y)$

proof : let $Z=F_X^{-1}(Y)$，$Prob(Z\le z)=Prob(-\infty\le F_X^{-1}(Y)\le z)=Prob(0\le Y\le F_X(z))=F_X(z)$

The Normal Distribution

$f_X(x)=\frac{1}{\sqrt{2\pi\sigma^2}}exp(-\frac{(x-\mu)^2}{2\sigma^2})$，$X\sim N(\mu,\sigma^2)$ also called Gauss Distribution

sum of Normal random variables $X_i\sim N(\mu_i,\sigma_i^2)$ then $\Sigma X_i\sim N(\Sigma\mu_i, \Sigma\sigma_i^2)$

$Erf(x)=\frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt$ (一般用$Erf(\frac{x}{\sqrt{2}})/2$)

Gamma Function $\Gamma(k)=\int_0^{+\infty}x^{k-1}e^{-x}dx$

μ_{2m} in Normal distribution 2m阶矩 $\mu_{2m}=\frac{2^m}{\sqrt{\pi}}\Gamma(m+\frac{1}{2})=(2m-1)!!$

Weibull distribution $f_{k,\sigma}(x)=(k/\sigma)(x/\sigma)^{k-1}e^{-(x/\sigma)^k}$ $x\ge0$ for 生存分析，胜率预测

Cauchy $f_X(x)=\frac{1}{\pi}\frac{1}{1+x^2}$ 无均值、方差

广义 Cauchy $f_X(x)=\frac{1}{b\pi}\frac{1}{1+(x-a)^2/b^2}$

$X\sim Cauchy(a_1,b_1)$，$Y\sim Cauchy(a_2,b_2)$ then $X+Y\sim Cauchy(a_1+a_2, b_1+b_2)$ 留数定理

Chi-square Distribution

Def $f(x)=\frac{1}{2^{\nu/2}\Gamma(\nu/2)}x^{(\nu/2-1)}e^{-\frac{x}{2}}$，$x\ge0$ 记作 $X\sim\chi^2(\nu)$，ν为自由度

Thm. if $X\sim N(0,1)$ then $X^2\sim\chi^2(1)$

Thm. if $X_1, X_2\sim N(0,1)$，$Y=X_1^2+X_2^2$，then $Y\sim\chi^2(2)$ ‖ 稳定分布：$Y_1\sim\chi^2(\nu_1)$，$Y_2\sim\chi^2(\nu_2)$，then $Y_1+Y_2\sim\chi^2(\nu_1+\nu_2)$

Proof : $Prob(Y\le y)=\iint_{x_1^2+x_2^2\le y}\frac{1}{\sqrt{2\pi}}e^{-\frac{x_1^2}{2}}\frac{1}{\sqrt{2\pi}}e^{-\frac{x_2^2}{2}}dx_1dx_2$

Part IV Limit Theorem

Inequalities and Laws of Large Number

Markov's Inequality X 有均值且非负，then $Prob(X \geq a) \leq \frac{E[X]}{a}$ $(\forall a>0)$

Chebyshev's Inequality X, μ_X, σ_X, then $Prob(|X-\mu_X| \geq k\sigma_X) \leq \frac{1}{k^2}$

Proof: $Y=(X-\mu_X)^2$, then using Markov

依分布收敛，依概率收敛

几乎必然收敛 $Prob(w \in \Omega : \lim_{n\to\infty} X_n(w) = X(w)) = 1$，必然收敛 $\forall w \in \Omega$

Thm. Laws of Large Numbers

$\{X_i\}$ 为独立同分布的随机变量，均值为 μ。记 $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$，则 $\bar{X}_n \xrightarrow{P} \mu$

Generating Functions and Convolutions

Def $G_a(s) = \sum_{n=0}^{\infty} a_n s^n$ which is 生成函数　　Def $c = a * b$: $c_k = \sum_{i=1}^k a_i b_{k-i}$

Thm. $G_{X_1+\cdots+X_n}(s) = \prod_{i=1}^n G_{X_i}(s)$ where $G_X(s) = E[s^X] = \sum_m s^{x_m} Prob(X = x_m)$ 独立 ☆

连续：$G_X(s) = \int_{-\infty}^{+\infty} s^x f(x)\,dx$　　离散型是唯一的

Def moment generating function $M_X(t) = \sum_{m=-\infty}^{+\infty} e^{tx_m} f(x_m)$ or $\int_{-\infty}^{+\infty} e^{tx} f(x)\,dx$ 即 $E[e^{tX}]$

Properties: $M_X(t) = 1 + \mu_1' t + \frac{\mu_2'}{2!} t^2 + \cdots$，即 $\left.\frac{d^k M_X(t)}{dt^k}\right|_{t=0} = \mu_k'$　　用于求 μ_k'

$M_{\alpha X+\beta}(t) = e^{\beta t} M_X(\alpha t)$

$M_{X_1+X_2}(t) = M_{X_1}(t) M_{X_2}(t)$ 要求 X_1 X_2 独立

随机变量取非负值时由矩母函数唯一确定

☆ 已知所有矩无法唯一地确定分布 e.g. $f(x) = \frac{1}{\sqrt{2\pi x^2}} e^{-(\log^2 x)/2}$，$g(x) = f(x)[1+\sin(2\pi\log x)]$　　$(e^{-\frac{1}{x^2}})$

Central Limit Theorem

Thm $X_1, \cdots, X_N$ 独立同分布 $\wedge \exists \delta>0$，当 $|t|<\delta$ 时矩母函数收敛：Then, $N \to +\infty$，$\frac{\bar{X}_N - \mu}{\sigma/\sqrt{N}} \longrightarrow$ normal distribution

Def Standardization $Z = \frac{X-\mu}{\sigma}$

标准正态分布 Z 的矩母函数 $M_Z(t) = e^{t^2/2}$

Laplace

定理 20.5.3 设矩母函数 $M_X(t)$ 和 $M_Y(t)$ 在 0 附近的一个邻域内存在 (即存在一个 δ, 使得当 $|t|<\delta$ 时这两个函数都存在). 如果在这个邻域内有 $M_X(t) = M_Y(t)$, 那么对于所有的 u 均有 $F_X(u) = F_Y(u)$. 因为概率密度函数是累积分布函数的导数, 所以 $f = g$.

Fourier

定理 20.5.4 设 $\{X_i\}_{i\in I}$ 是一个随机变量序列, 它们的矩母函数是 $M_{X_i}(t)$. 假设存在一个 $\delta>0$, 使得当 $|t|<\delta$ 时有 $\lim_{i\to\infty} M_{X_i}(t) = M_X(t)$, 其中 $M_X(t)$ 是一个矩母函数. 另外, 当 $|t|<\delta$ 时, 所有矩母函数均收敛. 那么存在唯一的累积分布函数 F, 使得 F 的矩均由 $M_X(t)$ 给出, 而且对于 $F_X(x)$ 的任意一个连续点 x, 均有 $\lim_{n\to\infty} F_{X_i}(x) = F_X(x)$.