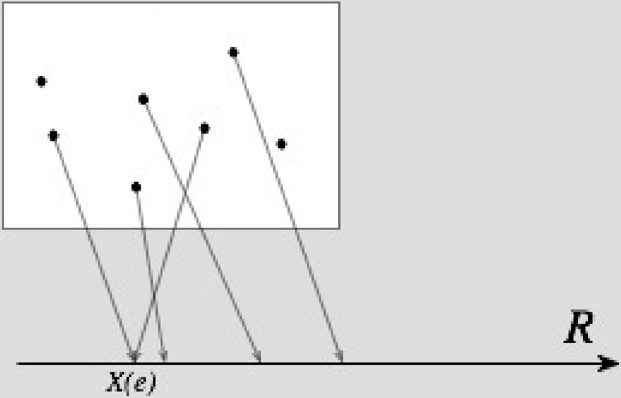
概要目录

基本术语、随机变量和分布函数性质的复习

- 概率空间(Ω, F, P)的三个要素：Ω，F，P
- 样本空间Ω：实验或随机试验所有可能结果的集合，而随机试验中的每个可能结果称为样本点。如：掷硬币、投骰子、抽扑克牌
- 事件Σ：样本空间的一个子集
- 样本空间Ω的幂集F：集合元素为事件Σ
- 概率P：从集合F到实数域R的函数，$P: F \to R$。每个事件都被赋予一个0和1之间的概率值，且$P(\Omega) = 1$。简写$Pr\{\cdot\}$

基本术语、随机变量和分布函数性质的复习

- 实随机变量X：从样本空间Ω映射到实数域的函数$X:\Omega\rightarrow R$



基本术语、随机变量和分布函数性质的复习

- 如何理解$Pr\{\omega \in \Omega : X(\omega) > 0\}$?
  - 事件$\{\omega \in \Omega : X(\omega) > 0\}$：使$X$为正的样本集合
  - 概率$Pr\{\cdot\}$：事件的概率值

- 例如，随机掷两个骰子，整个样本空间可以由36个元素组成：
  $S = \{(i,j) | i = 1, \ldots, 6; j = 1, \ldots, 6\}$

- 随机变量X(获得的两个骰子的点数和)可以有11个整数值
  $X(i,j) := i + j, x = 2, 3, \ldots, 12$

- 随机变量Y(获得的两个骰子的点数差)只有6个整数值
  $Y(i,j) := | i - j |, y = 0, 1, \ldots, 5$

基本术语、随机变量和分布函数性质的复习

- 缩写词r.v.表示"实随机变量(real random variables)"
- 离散型的r.v.X：如果存在有限或可列个不同的值$\lambda_1, \lambda_2, \ldots$使得$a_i \equiv Pr\{X = \lambda_i\} > 0, i = 1, 2, \ldots$,而且$\Sigma_i a_i = 1$
- 连续型的r.v.X：若对每个实数λ,$Pr\{X = \lambda\} = 0$
- 分布函数F：$F(\lambda) = Pr\{X \leq \lambda\}$
- 概率密度p：如果存在非负函数$p(t), -\infty < t < \infty$,使得r.v.$X$的分布函数$F(\lambda) = \int_{-\infty}^{\lambda} p(t)dt$

基本术语、随机变量和分布函数性质的复习

- 离散型r.v.的m阶矩：$E[X^m] = \Sigma_i \lambda_i^m Pr\{X = \lambda_i\}$,假定级数绝对收敛
- 连续型r.v.的m阶矩：$E[X^m] = \int_{-\infty}^{\infty} x^m p(x) dx$,假定积分绝对收敛
- X的一阶矩通常称为均值或期望，用m_X或μ_X表示
- X的m阶中心矩定义为r.v.$X - m_X$的m阶矩
- 一阶中心矩为0，二阶中心距称为X的方差，记为σ_X^2
- r.v.X的中位数ν：$Pr\{X \geq \nu\} \geq \frac{1}{2}, Pr\{X \leq \nu\} \geq \frac{1}{2}$

基本术语、随机变量和分布函数性质的复习

- 如果X是r.v.，g是函数，那么$Y = g(X)$也是随机变量
- 对离散型的r.v.X：$E[g(X)] = \Sigma_{i=1}^{\infty} g(\lambda_i) Pr\{X = \lambda_i\}$
- 对连续型的r.v.X：$E[g(X)] = \int g(t) p_X(t) dt$

联合分布函数

- 联合分布函数：$F(\lambda_1, \lambda_2) = F_{XY}(\lambda_1, \lambda_2) = Pr\{X \leq \lambda_1, Y \leq \lambda_2\}$
- 边缘分布函数：$F(\lambda, +\infty) \equiv \lim_{\lambda_2 \to \infty} F(\lambda_1 = \lambda, \lambda_2)$
  $F(+\infty, \lambda) \equiv \lim_{\lambda_1 \to \infty} F(\lambda_1, \lambda_2 = \lambda)$
- 相互独立：$F(\lambda_1, \lambda_2) = F(\lambda_1, +\infty) \cdot F(+\infty, \lambda_2)$
- 联合概率密度p_{XY}：如果存在二元实变量的非负函数$p_{XY}(s,t)$
  使得$F_{XY}(\lambda_1, \lambda_2) = \int_{-\infty}^{\lambda_2} \int_{-\infty}^{\lambda_1} p_{XY}(s,t) dsdt$
- 如果X, Y相互独立，则有$p_{XY}(s,t) = p_X(s) \cdot p_Y(t)$，
  其中p_X和p_Y分别是X和Y的边缘分布概率密度

联合分布函数

- 多个r.v.的联合分布函数： $F(\lambda_1, \lambda_2, \ldots, \lambda_n) = F_{X_1, X_2, \ldots, X_n}(\lambda_1, \lambda_2, \ldots, \lambda_n) = Pr\{X_1 \leq \lambda_1, X_2 \leq \lambda_2, \ldots, X_n \leq \lambda_n\}$
- 多个r.v.的边缘分布函数： $F_{X_{i_1}, \ldots, X_{i_k}}(\lambda_{i_1}, \ldots, \lambda_{i_k}) = \lim_{\lambda_i \to \infty, i \neq i_1, \ldots, i_k} F(\lambda_1, \ldots, \lambda_n)$
- 多个r.v.相互独立： $F(\lambda_1, \ldots, \lambda_n) = F_{X_1}(\lambda_1) \cdot \ldots \cdot F_{X_n}(\lambda_n)$
- 多个r.v.联合概率密度：如果存在实变量的非负函数$p(t_1, \ldots, t_n)$使得$F(\lambda_1, \ldots, \lambda_n) = \int_{-\infty}^{\lambda_n} \ldots \int_{-\infty}^{\lambda_1} p(t_1, \ldots, t_n) dt_1 \ldots dt_n$

联合分布函数

- 如果X和Y分别具有均值m_X和m_Y,协方差σ_{XY}定义为乘积矩: $\sigma_{XY} = E[(X - m_X)(Y - m_Y)]$
- 当X_1和X_2相互独立，$X = X_1 + X_2$的概率密度p是p_1和p_2的卷积：$p(x) = \int p_1(x - y)p_2(y)dy = \int p_2(x - y)p_1(y)dy$

条件分布和条件期望

- 已知事件B的条件下，事件A的条件概率：
  $Pr\{A|B\} = \frac{Pr\{AB\}}{Pr\{B\}}, Pr\{B\} > 0$
- 已知$Y = y$时，离散型X的条件分布函数：
  $F_{X|Y}(x|y) = \frac{Pr\{X \le x, Y=y\}}{Pr\{Y=y\}}, Pr\{Y = y\} > 0$
- 已知$Y = y$时，连续型X的条件分布函数：
  $F_{X|Y}(x|y) = \frac{\int_{\xi \le x} p_{XY}(\xi, y) d\xi}{p_Y(y)}, p_Y(y) > 0$
- 条件分布的三个本质特征：
  - 对每个固定的y, $F_{X|Y}(x|y)$是$X = x$的分布函数
  - 对每个固定的x, $F_{X|Y}(x|y)$是y的函数
  - 对任意x, y, $Pr\{X \le x, Y \le y\} = \int_{\eta \le y} F_{X|Y}(x|\eta) p_Y(\eta) d\eta$
    或$Pr\{X \le x, Y \le y\} = \Sigma_{\eta \le y} F_{X|Y}(x|\eta) Pr\{Y = \eta\}$

条件分布和条件期望

- 全概率公式：$Pr\{X \le x\} = \Sigma_y Pr\{X \le x|Y = y\}Pr\{Y = y\}$
  或$Pr\{X \le x\} = \int_{-\infty}^{+\infty} Pr\{X \le x|Y = y\}p_Y(y)dy$

- 已知$Y = y$时，条件密度函数：
  $p_{X|Y}(x|y) = \frac{d}{dx}F_{X|Y}(x|y) = \frac{p_{XY}(x,y)}{p_Y(y)}, p_Y(y) > 0$

- 已知$Y = y$时，$g(X)$的条件期望：
  $E[g(X)|Y = y] = \int g(x)p_{X|Y}(x|y)dx = \frac{\int g(x)p_{XY}(x,y)dx}{p_Y(y)}, p_Y(y) > 0$
  或$= \Sigma_{i=1}^{\infty} g(x_i)Pr\{X = x_i|Y = y\} = \frac{\Sigma_{i=1}^{\infty} g(x_i)Pr\{X=x_i,Y=y\}}{Pr\{Y=y\}}$,
  $Pr\{Y = y\} > 0$

- 对于每个函数g,若$E[|g(X)|] < +\infty$,则$E[g(X)|Y = y]$
  是y的函数

条件分布和条件期望

- 对于任意有界函数h,
  $E[g(X)h(Y)] = \Sigma_{i=1}^{\infty} E[g(X)|Y = y_i]h(y_i)Pr\{Y = y_i\}$
  或$E[g(X)h(Y)] = \int E[g(X)|Y = y]h(y)p_Y(y)dy$
- 期望的全概率公式：$E[g(X)] = \Sigma_{i=1}^{\infty} E[g(X)|Y = y_i]Pr\{Y = y_i\}$
  或$E[g(X)] = \int E[g(X)|Y = y]p_Y(y)dy$
- 条件期望的性质：$E[a_1 g_1(X) + a_2 g_2(X)|Y = y]$
  $=a_1 E[g_1(X)|Y = y] + a_2 E[g_2(X)|Y = y]$
- $E[g(X)|Y = y]$是实变量y的函数，记为随机变量$E[g(X)|Y]$
- 对于任意有界函数h, $E[g(X)h(Y)] = E\{E[g(X)|Y]h(Y)\}$
- 期望的全概率公式：$E[g(X)] = E\{E[g(X)|Y]\}$

条件分布和条件期望

- $E[a_1 g(X_1) + a_2 g(X_2)|Y] = a_1 E[g(X_1)|Y] + a_2 E[g(X_2)|Y]$
- 若X和Y独立，则$E[g(X)|Y] = E[g(X)]$
- $E[g(X)f(Y)|Y] = f(Y)E[g(X)|Y]$
- $E[g(X)f(Y)] = E\{E[g(X)|Y]f(Y)\}$
- $E[c|Y] = c$
- $E[f(Y)|Y] = f(Y)$
- $E[g(X)] = E\{E[g(X)|Y]\}$

特征函数

- 特征函数$\phi(t) = E[e^{itX}], i = \sqrt{-1}, -\infty < t < +\infty$
- $\phi(t) = \int_{-\infty}^{\infty} e^{it\lambda} p(\lambda) d\lambda$,或$\phi(t) = \Sigma_{k=0}^{\infty} e^{it\lambda_k} Pr\{X = \lambda_k\}$
- 缩写词c.f.表示“特征函数(characteristic function)”
- 特征函数的重要意义:
  - 分布函数与特征函数之间一一对应，用特征函数表达分布函数的方程式是勒维(Levy)变换公式
  - 相互独立的随机变量的和的特征函数是它们的特征函数的乘积
  - 由特征函数的微商得到矩：$E[X^k] = \frac{1}{i^k}\phi^{(k)}(0) = \frac{1}{i^k}\frac{d^k\phi(0)}{dt^k}$

母函数和拉普拉斯变换

- 对于仅取非负整数值的随机变量，与特征函数相应的母函数定义：

$$g(s) = \Sigma_{k=0}^{\infty} s^k Pr\{X = k\} = E[s^X], \quad s\text{是复变量，且}|s| \leq 1$$

- 母函数可经过变量替换$s = e^{it}$变成X的特征函数：

$$\phi(t) = E[e^{itX}] = E[(e^{it})^X] = g(e^{it})$$

- 母函数的基本性质:
  - 母函数唯一确定其分布函数
  - 独立非负整数值随机变量和的母函数是它们母函数的乘积
  - 各阶矩可通过逐次微分得到，阶乘矩由下式得到：
    $E[X(X-1)\dots(X-k)] = g^{(k+1)}(1) = \frac{d^{k+1}g(1)}{ds^{k+1}}$
    $E[X] = g^{(1)}(1), E[X^2] = g^{(2)}(1) + g^{(1)}(1)$

母函数和拉普拉斯变换

- 例子：设$N, X_1, X_2, \ldots$是独立非负整数值随机变量，N的母函数为$g_N(s)$，并假定X_i同分布，具有相同的母函数$g(s)$，请确定和$R = X_1 + X_2 + \ldots + X_N$的母函数$g_R(s)$

- $g_R(s) = E[s^R] = E[s^{X_1+\ldots+X_N}] = E\{E[s^{X_1+\ldots+X_N}|N]\}$
  $= \Sigma_{n=0}^{\infty} E[s^{X_1+\ldots+X_n}|N = n]Pr\{N = n\}$
  $= \Sigma_{n=0}^{\infty} E[s^{X_1+\ldots+X_n}]Pr\{N = n\}$
  $= \Sigma_{n=0}^{\infty} g^n(s)Pr\{N = n\}$
  $= E[g^N(s)] = g_N[g(s)]$

- 利用复合函数微分链式法则，有$g_R'(s) = g_N'[g(s)] \cdot g'(s)$
  令$s = 1$,得$E[R] = E[N] \cdot E[X]$

- 再计算二阶微分$g_R''(s) = g_N''[g(s)] \cdot (g'(s))^2 + g_N'[g(s)] \cdot g''(s)$,
  得到$g_R''(1) = g_N''(1) \cdot E[X]^2 + E[N] \cdot g''(1)$

母函数和拉普拉斯变换

- 回顾$g''(1) = E[X(X-1)] = E[X^2] - E[X] = \sigma_X^2 + E[X]^2 - E[X]$
- 类似地，$g_R''(1) = \sigma_R^2 + E[R]^2 - E[R]$，$g_N''(1) = \sigma_N^2 + E[N]^2 - E[N]$
- 代入$g_R''(1) = g_N''(1) \cdot E[X]^2 + E[N] \cdot g''(1)$，得到
- $\sigma_R^2 + E[R]^2 - E[R]$
  $= (\sigma_N^2 + E[N]^2 - E[N]) \cdot E[X]^2 + E[N] \cdot (\sigma_X^2 + E[X]^2 - E[X])$
  $= (\sigma_N^2 + E[N]^2) \cdot E[X]^2 + E[N] \cdot (\sigma_X^2 - E[X])$
  $= \sigma_N^2 \cdot E[X]^2 + (E[N] \cdot E[X])^2 + E[N] \cdot \sigma_X^2 - E[N] \cdot E[X]$
- 利用上面结论$E[R] = E[N] \cdot E[X]$，得到

$$\sigma_R^2 = E[X]^2 \sigma_N^2 + E[N] \sigma_X^2$$

- 推广到特征函数$\phi_R(t) = g_N(\phi(t))$

母函数和拉普拉斯变换

- 非负随机变量的拉普拉斯变换：$\psi_X(s) = E[e^{-sX}]$, s是复变量
- $\psi_X(s) = \int_0^\infty e^{-sx} p_X(x)dx$,或$\psi_X(s) = \Sigma_{n=0}^\infty e^{-s\lambda_n} Pr\{X = \lambda_n\}$
- 拉普拉斯变换与特征函数的关系：$s = it$时，$\psi_X(s) = \phi_X(-t)$
- 对相互独立的非负随机变量组，$\psi_{X_1+\ldots+X_n}(s) = \Pi_{k=1}^n \psi_{X_k}(s)$
- 拉普拉斯变换唯一确定其分布函数

分布函数的例子

常见的连续型概率分布

连续型分布	密度函数 $p(x)$	参数范围	特征函数 $\phi(t)$	均值	方差
正态分布	$\frac{1}{\sqrt{2\pi}\sigma}\exp-\frac{(x-m)^2}{2\sigma^2}$ $-\infty<x<\infty$	m 实数 $\sigma>0$	$\exp\left[-\frac{\sigma^2t^2}{2}+\mathrm{i}mt\right]$	m	σ^2
指数分布	$\lambda\mathrm{e}^{-\lambda x}, x>0$	$\lambda>0$	$\frac{\lambda}{\lambda-\mathrm{i}t}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Γ 分布	$\frac{\lambda}{\Gamma(\alpha)}(\lambda x)^{\alpha-1}\mathrm{e}^{-\lambda x}$, $x>0$	$\lambda>0$ $\alpha>0$	$\frac{\lambda^\alpha}{(\lambda-\mathrm{i}t)^\alpha}$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$
均匀分布	$\frac{1}{b-a}, a<x<b$	$a<b$	$\frac{\mathrm{e}^{\mathrm{i}ub}-\mathrm{e}^{\mathrm{i}ua}}{\mathrm{i}u(b-a)}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Beta 分布	$\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}x^{p-1}(1-x)^{q-1}$, $0<x<1$	$p>0$ $q>0$	$\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}\int_0^1\mathrm{e}^{\mathrm{i}tx}x^{p-1}(1-x)^{q-1}\mathrm{d}x$	$\frac{p}{p+q}$	$\frac{qp}{(p+q)^2(p+q+1)}$

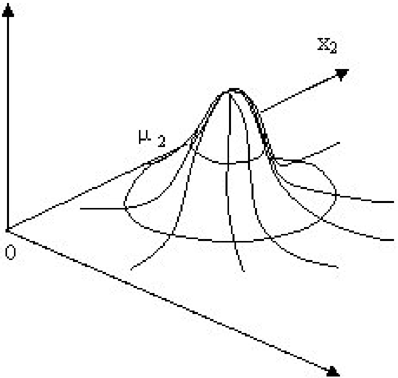
分布函数的例子

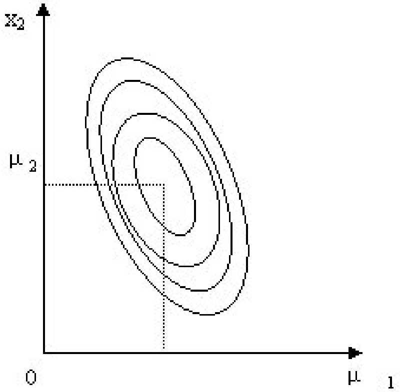
常见的离散型概率分布

离散型分布	概率分布列	参数的可能取值	母函数	均值	方差
泊松分布	$\frac{e^{-\lambda}\lambda^n}{n!},\ n=0,1,2,\cdots$	$\lambda>0$	$e^{-\lambda+\lambda s}$	λ	λ
二项分布	$\binom{N}{n}p^n q^{N-n},\ n=0,1,\cdots,N$	$N=1,2,\cdots$ $0<p<1$ $q=1-p$	$(1-p+ps)^N$	Np	Npq
负二项（巴斯卡）分布	$\binom{\alpha+n-1}{n}p^\alpha q^n,\ n=0,1,2,\cdots$	$\alpha>0$ $0<p<1$	$\left(\frac{p}{1-qs}\right)^\alpha$	$\frac{\alpha q}{p}$	$\frac{\alpha q}{p^2}$
几何分布	$p(1-p)^n,\ n=0,1,2,\cdots$	$0<p<1$	$\frac{p}{1-qs}$	$\frac{q}{p}$	$\frac{q}{p^2}$

多元正态分布

- 联合正态分布： $Pr\{X_1 \le a, X_2 \le b\}$
  $= \int_{-\infty}^{a}\int_{-\infty}^{b} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\{-\frac{1}{2}Q(x_1,x_2)\} dx_1 dx_2,$
  其中 $Q(x_1,x_2) = \frac{1}{1-\rho^2}\{(\frac{x_1-m_1}{\sigma_1})^2 - 2\rho(\frac{x_1-m_1}{\sigma_1})(\frac{x_2-m_2}{\sigma_2}) + (\frac{x_2-m_2}{\sigma_2})^2\}$





多元正态分布

- 容易验证$E[X_i]=m_i$,
  X_i的方差是σ_i^2
- 协方差为$E[(X_1-m_1)(X_2-m_2)]=\rho\sigma_1\sigma_2$,
  ρ称为相关系数
- 联合特征函数: $\phi_{X_1,X_2}(t_1,t_2)=E[e^{i(t_1X_1+t_2X_2)}]$
  $=\exp\{i(t_1m_1+t_2m_2)-\frac{1}{2}(t_1^2\sigma_1^2+2\rho t_1\sigma_1t_2\sigma_2+t_2^2\sigma_2^2)\}$
- 已知$X_1=x_1$时X_2的条件概率密度函数
  $p_{X_2|X_1}(x_2|x_1)=\frac{1}{\sqrt{2\pi}\sigma}\exp[-\frac{1}{2}(\frac{x_2-m}{\sigma})^2]$,
  其中$m=m_2+\frac{\sigma_2}{\sigma_1}\rho(x_1-m_1),\sigma=\sigma_2\sqrt{1-\rho^2}$

多项分布

- 多项分布：取非负整数值$0,\ldots,n$的r个离散型随机变量的联合分布
  $$Pr\{X_1=k_1\ldots X_r=k_r\}=\begin{cases}\frac{n!}{k_1!\ldots k_r!}p_1^{k1}\ldots p_r^{k_r}, & if\ \ k_1+\ldots+k_r=n\\ 0, & otherwise\end{cases}$$
  其中$p_i>0, i=1,\ldots,r$,且$\Sigma_{i=1}^r p_i=1$
- 举例：将n个球顺序取出随机放到r个不同的篮子里，放到第i个篮子的概率为p_i
- 第i个篮子恰好有k_i个球的组合数为$\frac{n!}{k_1!\ldots k_r!}$
- 因此，第i个篮子恰好有k_i个球的概率为$\frac{n!}{k_1!\ldots k_r!}p_1^{k1}\ldots p_r^{k_r}$
- 多项分布的联合母函数
  $g(s_1,\ldots,s_r)=E[s_1^{X_1}\ldots s_r^{X_r}]=(p_1s_1+\ldots+p_rs_r)^n$

多项分布

- 组合数学相关知识回顾:

	排列 { a,b } ≠ { b,a } (考虑顺序)	组合 { a,b } = { b,a } (不考虑顺序)
不重复出现 (不放回去) { a, b, c }	P_n^k	C_n^k
重复出现 (再放回去) { a, a, b }	n^k	H_n^k

$$P_n^k = \frac{n!}{(n-k)!}$$

$$C_n^k = \binom{n}{k} = \frac{P_n^k}{k!} = \frac{n!}{k!(n-k)!}$$

$$H_n^k = C_{n+k-1}^k = \binom{n+k-1}{k}$$

- 为什么$H_n^k = C_{n+k-1}^k$?
  - 提示:建立H_n^k计数与C_{n+k-1}^k计数的一一对应

极限定理

- 以概率1收敛：若$Pr\{\lim_{n\to\infty} Z_n = Z\} = 1$,则说$Z_n$以概率1收敛于$Z$,
  - 以概率1出现结果：$Z = z, Z_1 = z_1, \ldots,$ 且$\lim_{n\to\infty} z_n = z$

- 依概率收敛：若对每个正数ε，$\lim_{n\to\infty} Pr\{|Z_n - Z| > \varepsilon\} = 0$或者$\lim_{n\to\infty} Pr\{|Z_n - Z| \leq \varepsilon\} = 1$成立，则说$Z_n$依概率收敛于$Z$,
  - 可以取到足够大的n，使得Z_n任意接近Z的概率任意接近1.

- 利用样本函数解释概念(①,②,$\cdots$,ⓜ表示随机变量序列的一个实现)

$$
\begin{array}{c|ccccccc}
 & a_1, & a_2, & a_3, & \cdots, & a_n, & \cdots \to & a \\
\hline
 & Z_1, & Z_2, & Z_3, & \cdots, & Z_n, & \cdots \to & Z \\
① & z_1^1, & z_2^1, & z_3^1, & \cdots, & z_n^1, & \cdots \to & z^1 \\
② & z_1^2, & z_2^2, & z_3^2, & \cdots, & z_n^2, & \cdots \to & z^2 \\
\cdots & \cdots, & \cdots, & \cdots, & \cdots, & \cdots, & \cdots \to & \cdots \\
ⓜ & z_1^m, & z_2^m, & z_3^m, & \cdots, & z_n^m, & \cdots \to & z^m
\end{array}
$$

极限定理

- 二次平均收敛(均方收敛)：若$\lim_{n\to\infty} E[|Z_n - Z|^2] = 0$，则说$Z_n$二次平均收敛(均方收敛)于$Z$,
  - 可以取到足够大的n，在差的平方平均意义下Z_n任意逼近Z.

- 依分布收敛：设$F(t) = Pr\{Z \leq t\}$和$F_k(t) = Pr\{Z_k \leq t\}$, $k = 1, 2, \ldots$，如果$\lim_{n\to\infty} F_n(t) = F(t)$对所有$F$的连续点$t$成立，则说$Z_n$依分布收敛于$Z$.

- Z_n以概率1收敛于$Z \Rightarrow Z_n$依概率收敛于$Z \Rightarrow Z_n$依分布收敛于Z.
  - 举例：满足依概率收敛，但不满足以概率1收敛(超出课程要求)?
  - 考虑序列$Z_{1,1}, Z_{1,2}, Z_{2,2}, Z_{1,3}, Z_{2,3}, Z_{3,3}, \cdots$

$$Z_{i,j} = \begin{cases} 1 & \omega \in [\frac{i-1}{j}, \frac{i}{j}) \\ 0 & \text{其它} \end{cases},$$

随机变量$\Omega = \omega$在$(0, 1)$内服从均匀分布

极限定理

- 设$X_1, X_2, \ldots$是独立同分布随机变量序列，具有有限均值m，令$S_n = X_1 + \ldots + X_n$，$\overline{X}_n = \frac{S_n}{n}$是样本均值.
- 弱大数定理：$\overline{X}_n$依概率收敛于m，即，对任意$\varepsilon > 0$，有

$$\lim_{n\to\infty} Pr\{|\overline{X}_n - m| > \varepsilon\} = 0.$$

- 强大数定理：$\overline{X}_n$以概率1收敛于m，即，

$$Pr\{\lim_{n\to\infty} \overline{X}_n = m\} = 1.$$

- 中心极限定理：假设X_k有方差σ^2，设$Z_n = \frac{S_n - nm}{\sigma\sqrt{n}} = \frac{1}{\sigma}(\overline{X}_n - m)\sqrt{n}$，则$Z_n$依分布收敛于一个零均值、单位方差的正态随机变量，

$$\lim_{n\to\infty} Pr\{Z_n \le \alpha\} = \int_{-\infty}^{\alpha} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du, \forall\alpha.$$

不等式

- 切比雪夫不等式：设Z是非负随机变量，则对任何正数c，
$$Pr\{Z > c\} \leq \frac{1}{c}E[Z].$$
$$E[Z] = \int_0^\infty z p_Z(z)dz \geq \int_c^\infty z p_Z(z)dz \geq c \cdot \int_c^\infty p_Z(z)dz = c \cdot Pr\{Z > c\}.$$

- 如果X是具有均值μ和方差σ^2的随机变量，应用切比雪夫不等式于$Z = (X - \mu)^2$，得到
$$Pr\{|X - \mu| > \varepsilon\} = Pr\{Z > \varepsilon^2\} \leq \frac{\sigma^2}{\varepsilon^2}.$$

- 施瓦兹不等式：设X, Y是具有有限二阶矩的随机变量，则
$$(E[XY])^2 \leq E[X^2]E[Y^2].$$

- 对实数λ，$0 \leq E[(X + \lambda Y)^2] = E[X^2] + 2\lambda E[XY] + \lambda^2 E[Y^2]$，看作$\lambda$的二次函数，它至多有一个实数根，判别式非正
$$4(E[XY])^2 \leq 4E[X^2]E[Y^2].$$

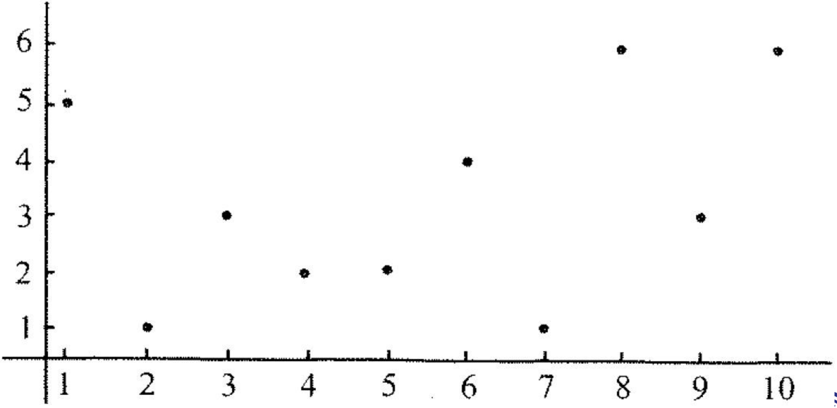
概要目录

随机过程的两个简单例子

- 随机过程理论主要研究随机变量族$\{X_t, t \in T\}$的构造,
  - T称为指标集,
  - X_t或记为$X(t)$.
- 随机过程$\{X_t, t \in T\}$的一个实现或样本函数，是指对每个$t \in T$，X_t的可能值的一个指定.
- 指标集可以是离散时间$T = \{0, 1, 2, \ldots\}$，这时$\{X_t\}$可以表示一系列试验的结果:
  - 抛掷硬币的一系列结果;
  - 智力测验时对象的一系列反应;
  - 对人口某种特征的一系列观察.

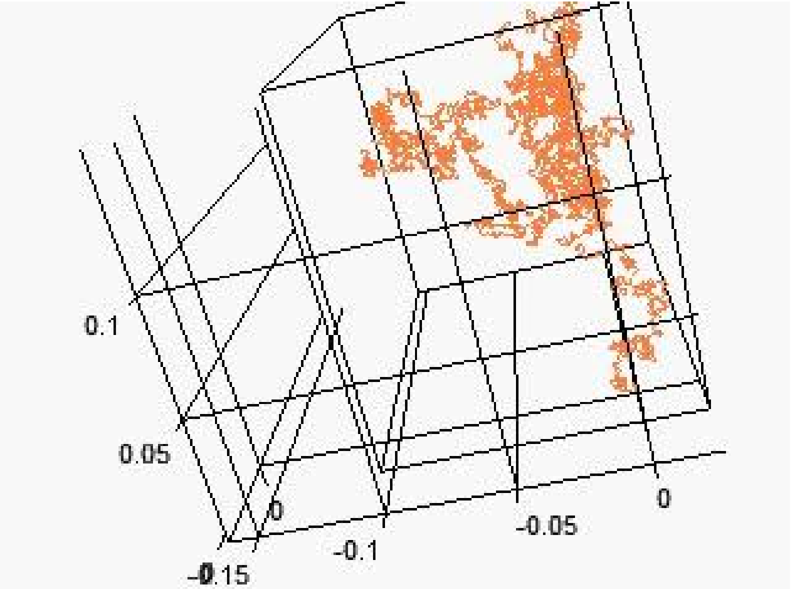
随机过程的两个简单例子

- 例如，抛掷骰子的结果$5, 1, 3, 2, 2, 4, 1, 6, 3, 6, \ldots$



- 指标集$T = [0, \infty)$的随机过程在应用中尤为重要,
    - 此时t通常解释为时间.

例1:布朗运动



例1:布朗运动

布朗运动(Wiener过程)的特征:

- 独立增量过程：若对任意n个实值$t_1, t_2, \cdots, t_n, t_1 < t_2 < \cdots < t_n$，随机变量组$X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \cdots, X_{t_n} - X_{t_{n-1}}$是独立的.
- 平稳增量过程：如果增量$X_t - X_s$的分布仅仅依赖于区间长度$t - s$而与s无关.
- 增量$X_t - X_s, t > s$的分布函数:期望为0,方差为$B(t-s)$的正态分布(将增量拆分为很小时间段上增量的和，由中心极限定理猜测增量满足正态分布)

$$Pr\{X_t - X_s \le x\} = \frac{1}{\sqrt{2\pi B(t-s)}} \int_{-\infty}^{x} e^{-\frac{u^2}{2B(t-s)}} du,$$

其中$t > s$,B是正常数.

例1:布朗运动

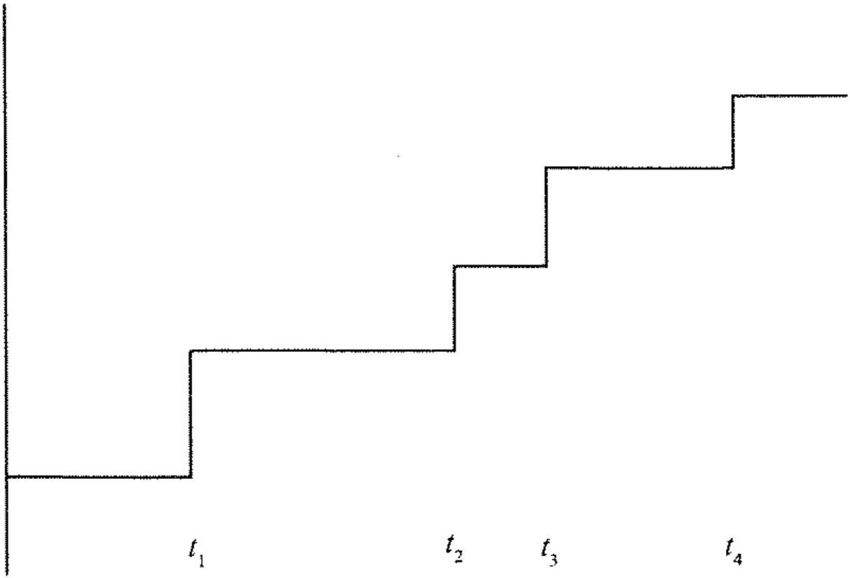
- 马尔可夫过程：对任意$t_1 < t_2 < \cdots < t_n < t$,

$$Pr\{a < X_t \le b | X_{t_1} = x_1, X_{t_2} = x_2, \cdots, X_{t_n} = x_n\}$$
$$= Pr\{a < X_t \le b | X_{t_n} = x_n\}.$$

- 布朗运动是一个马尔可夫过程:
  - 假设$X_0 = 0$，则有$E[X_t] = 0, \sigma^2(X_t) = Bt$,

$$Pr\{X_t \le x | X_{t_1} = x_1, \ldots, X_{t_n} = x_n\}$$
$$= \frac{1}{\sqrt{2\pi B(t - t_n)}} \int_{-\infty}^{x - x_n} e^{-\frac{u^2}{2B(t-t_n)}} du.$$

例2:泊松过程



例2:泊松过程

- 泊松过程的特征:
  - 平稳独立增量过程.
  - 在长度为h的时间区间中，事件至少发生一次的概率为
    $$p(h) = ah + o(h), h \to 0, a > 0.$$
  - 在长度为h的时间区间中，事件发生两次或两次以上的概率为$o(h)$ (稀有事件率、伯努利试验的二项分布导出)
- X_t的分布函数：参数为at的泊松分布,
  $$Pr\{X_t = m\} = \frac{a^m t^m}{m!} e^{-at},$$
  其中a是正常数.
- 函数$P_m(t) = Pr\{X_t = m\}$是另一个理解随机过程的视角.
- $E[X_t] = at$，$\sigma^2(X_t) = at$.

概要目录

一般随机过程的分类

- 随机过程的分类要素:状态空间的种类,指标集T以及随机变量X_t之间的相关关系.
- 状态空间S:每个X_t所有可能值所在的空间.
  - 取自然数值的过程或离散状态空间过程:$S = \{0, 1, 2, \cdots\}$.
  - 实值随机过程:S为实直线$(-\infty, +\infty)$.
  - k维向量过程:S是k维欧几里得空间.
- 指标集T:
  - 离散时间随机过程:$T = \{0, 1, 2, \cdots\}$,常用X_n代替X_t.
  - 连续时间过程:$T = [0, \infty)$.
  - T可以不是一维的:空间泊松过程,海洋波浪等.
- 随机过程的主要特征:随机变量$X_t (t \in T)$之间的相互关系,由该过程所有有限个随机变量的联合分布函数所确定.

随机过程的经典类型：平稳独立增量过程

- 增量:$X_t - X_s, t > s$.

- 独立增量过程:若对任意n个实值$t_1, t_2, \cdots, t_n, t_1 < t_2 < \cdots < t_n$, 随机变量组$X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \cdots, X_{t_n} - X_{t_{n-1}}$是独立的.

- 平稳增量过程:如果增量$X_t - X_s$的分布仅仅依赖于区间长度$t - s$, 而与s无关.

- 平稳独立增量过程的均值和方差:
  - $E[X_t] = m_0 + m_1 t$,其中$m_0 = E[X_0], m_1 = E[X_1] - m_0$.
  - $\sigma_{X_t}^2 = \sigma_0^2 + \sigma_1^2 t$,其中$\sigma_0^2 = E[(X_0 - m_0)^2], \sigma_1^2 = E[(X_1 - m_0 - m_1)^2] - \sigma_0^2$.

- 布朗运动和泊松过程都满足上述这个性质

  (注意应用中的某个随机过程可能同时满足多个经典类型，可以选择使用/共同使用不同工具来分析求解)

随机过程的经典类型：鞅(martingale)

- 鞅:对任意$t_1 < t_2 < \cdots < t_{n+1}$,以及任意实数$a_1, a_2, \cdots, a_n$,

$$E[X_{t_{n+1}} | X_{t_1} = a_1, X_{t_2} = a_2, \cdots, X_{t_n} = a_n] = a_n.$$

- 可描述公平赌博的模型(是否存在必胜策略？ Martingale？)

- 当$Z_i, i = 1, 2, \cdots$是相互独立且均值为0的随机变量时,

$$X_n = \Sigma_{i=1}^{n} Z_i, n = 1, 2, \cdots$$

  是一个离散时间鞅.

- 若$X_t, 0 \leq t < \infty$,具有独立增量且均值为0,
  则X_t是连续时间鞅.

随机过程的经典类型：马尔可夫过程

- 马尔可夫过程:对任意$t_1 < t_2 < \cdots < t_n < t$,

$$Pr\{a < X_t \le b | X_{t_1} = x_1, X_{t_2} = x_2, \cdots, X_{t_n} = x_n\}$$
$$= Pr\{a < X_t \le b | X_{t_n} = x_n\}.$$

- 马尔可夫性质蕴涵一个条件独立性，即“未来⊥过去|现在”. (思考独立性与条件独立性的区别)

- 转移概率函数：

$$P(x, s; t, A) = Pr\{X_t \in A | X_s = x\}, t > s,$$

A是实直线上一个区间.

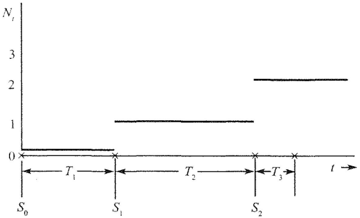
- 马尔可夫链：状态空间是有限或可数的马尔可夫过程，如泊松过程
- 扩散过程：所有样本函数是连续函数的马尔可夫过程，如布朗运动

随机过程的经典类型：平稳过程

- 严格平稳：对任意$h>0$和任意$t_1,t_2,\cdots,t_n\in T$，随机变量组$(X_{t_1+h},X_{t_2+h},\cdots,X_{t_n+h})$和$(X_{t_1},X_{t_2},\cdots,X_{t_n})$的联合分布函数相同.
- 特别地，对任意时刻t，X_t的分布都相同.
- 称马尔可夫过程有平稳转移概率：如果$P(x,s;t,A)$仅是$t-s$的函数
- 泊松过程和布朗运动都不是平稳的，但增量$Z_t=X_{t+h}-X_t$对任意固定h是平稳随机过程.
- 不存在非常数的平稳独立增量过程是平稳过程.

随机过程的经典类型：更新过程

- 更新过程：独立同分布正随机变量序列$\{T_k\}_{k=1,2,\cdots}$，T_k表示某“元件”的寿命.
- 第n次更新时间是
$$S_n = \Sigma_{i=1}^n T_i.$$
- 更新计数过程$\{N_t\}$表示在时间区间$[0,t]$更新的数目:
$$N_t = n, \text{当} S_n \le t < S_{n+1}, n = 0, 1, 2, \cdots.$$



- 带有参数λ的泊松过程是更新计数过程，其元件寿命具有参数为λ的指数分布(具有“无记忆性“).

随机过程的经典类型：点过程

- 点过程：指标集为集合$A \in \mathcal{A}$并以非负整数集$\{0, 1, \cdots, \infty\}$为状态空间的随机过程，其中S是n维空间的一个集合,$\mathcal{A}$是S的一个子集族
- 解释:“点” 以某种随机方式散布在S上,$N(A)$表示落在子集A里的点数
- $N(A_1 \bigcup A_2) = N(A_1) + N(A_2)$，若$A_1$和$A_2$不相交，$A_1 \bigcup A_2 \in \mathcal{A}$.
- $N(\emptyset) = 0$，若$\emptyset \in \mathcal{A}$.
- 每个泊松过程$\{X_t, t \in [0, \infty)\}$确定$S = [0, \infty)$上一个泊松点过程，

$$N((s, t]) = X_t - X_s.$$