

Probability (unfinished)

Yuxin Wu
(ppwwyyxxc@gmail.com)

Contents

1	Basic Concepts	2
1.1	Events	2
1.2	Probability Space	2
1.3	Properties of Probability	3
1.4	Classical/Geometrical Definitions of Probability	3
1.5	Conditional Probability	4
1.6	Independence of Events	5
2	Random Variables	6
2.1	Definition	6
2.2	Expectation & Variance	6
2.3	Probability Generating Function	7
2.4	Characteristic Function	8
3	Distributions	8
3.1	Common Discrete Distributions	8
3.2	Common Continuous Distributions	10
4	Multivariate Distributions	14
4.1	Joint Distribution	14
4.2	Common Multivariate Distributions	14
4.3	Convolution	16
4.4	Other Functions on Random Variables	17
4.5	Correlation	18
4.6	Gauss Distribution	19
4.7	Conditional Distribution	20
5	Analytic Topics	21
5.1	Estimation	21
5.2	Convergence	21
5.3	Law of Large Numbers	21
5.4	Central Limit Theorem	22
6	Random Walk	22
6.1	Combinatorial Perspectives	22
6.2	Generating Function Perspectives	24
6.3	Ruin Problem	25
6.4	Higher Dimensions	26

7	Branching Process	26
7.1	Extinction Probability	26
7.2	Total Progeny	27
7.3	Geometric Distribution	27
7.4	Busy Periods	27
8	Recurrent Event	27
8.1	Definition	27
8.2	A Theorem of Particular Importance	28
8.3	Number of Occurences	29
8.4	Runs of Specific Patterns	29

1 Basic Concepts

1.1 Events

一次随机试验中每一种可能的结果称为一个**基本事件**或**样本点** ω ,所有基本事件的全体为该试验的样本空间 Ω

同一试验的样本空间可能不唯一,因为观察结果的角度不同.对扔两次色子, $\Omega_1 = \{++,+-,--, -+\}$, $\Omega_2 = \{\text{两正,两负,一正一负}\}$

至多可数的样本空间称为离散样本空间,不可数称为连续样本空间.

Ω 的可测子集 A 称为事件.对结果 $\omega \in A$,则称事件 A 发生了.

$A \subset B \Rightarrow A$ 发生了 B 必发生.

Morgan 律: $(\cup A_i)^c = \cap A_i^c$

1.2 Probability Space

概率空间 (Ω, \mathcal{F}, P) :

Ω 是全体可能结果组成的集合. \mathcal{F} 是全体可观测事件组成的事件族. $P: \mathcal{F} \rightarrow [0, 1]$ 是求事件的概率的运算.

当 \mathcal{F} 满足以下条件时,称其为 σ -代数:

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
3. 可数并: $A_1 \cdots \in \mathcal{F} \Rightarrow \cup_{i=1}^{\infty} A_i \in \mathcal{F}$

事实上,由可数并,可推出有限并,可数交,有限交 $\in \mathcal{F}$.

当 Ω 为至多可数集时,总可取 Ω 的所有子集族作为 \mathcal{F} . 当 Ω 不可数时,取这样的 \mathcal{F} 会造成数学上的困难,因此只取感兴趣的,可以知道概率的事件的最小 σ -代数.

概率的定义:对每个事件 $A \in \mathcal{F}$, 定义实数 $P(A)$,满足以下条件:

1. 非负性: $P(A) \geq 0$
2. 正则性: $P(\Omega) = 1$
3. 可数可加性:

对两两互不相容的事件 $A_1 \cdots \in \mathcal{F}$, $P(\cup A_n) = \sum P(A_n)$

试验的样本空间,事件域(σ 代数)及定义在其上的概率构成的三元组 (Ω, \mathcal{F}, P) 称为描述一个随机试验的概率空间.

1.3 Properties of Probability

事件序列的极限定义: $\overline{\lim}_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$ (当且仅当有无穷个 A_n 发生)

$\lim_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$ (当且仅当至多有有限个 A_n 不发生)

当上下极限相等时(如对于单调事件序列),称为序列 A_n 的极限.

利用可数可加,可得到如下结论:

1. $P(\emptyset) = 0 : P(\Omega) = P(\Omega \cup \emptyset \cup \emptyset \dots) = P(\Omega) + P(\emptyset) + \dots$
2. 有限可加
3. 求逆: $P(A) + P(A^c) = 1$
4. Jordan 公式(容斥),归纳证明
5. $P(A - B) = P(A) - P(A \cap B)$,特别地, $B \subset A \Rightarrow P(B) \leq P(A)$
6. 下连续性: 设 A_i 单调增($A_1 \subset A_2 \subset \dots$), 则 $P(\lim_{n \rightarrow \infty} A_n) = \lim_{n \rightarrow \infty} P(A_n)$.

$$P(\cup A_n) = P(A_1) + \sum_{i=1}^{\infty} P(A_{i+1} - A_i) = P(A_1) + \lim_{n \rightarrow \infty} \sum_{i=1}^{n-1} [P(A_{i+1} - P(A_i))] = \lim_{n \rightarrow \infty} P(A_n)$$

7. 上连续性: 设 A_i 单调减, 则 $P(\lim_{n \rightarrow \infty} A_n) = \lim_{n \rightarrow \infty} P(A_n)$

$$1 - \lim_{n \rightarrow \infty} P(A_n) = \lim_{n \rightarrow \infty} P(A_n^c) = P(\bigcup_{n=1}^{\infty} A_n^c) = P((\bigcap_{n=1}^{\infty} A_n)^c) = 1 - P(\bigcap_{n=1}^{\infty} A_n)$$

概率的上下连续性等价,统称为连续性.

8. 有限可加 + 下连续 \Leftrightarrow 可数可加.

由下连续性,

$$P(\bigcup_{n=1}^{\infty} A_n) = P(\bigcup_{n=1}^{\infty} F_n) \underset{\text{下连续}}{=} \lim_{n \rightarrow \infty} P(F_n) = \lim_{n \rightarrow \infty} P(\bigcup_{i=1}^n A_i) \underset{\text{有限可加}}{=} \lim_{n \rightarrow \infty} \sum_{i=1}^n P(A_i) = \sum_{i=1}^{\infty} P(A_i) \underset{\text{收敛}}{=}$$

9. 推广可数可加: $A_1, A_2 \dots$ 满足 $P(A_i A_j) = 0$ (弱于互斥), 则 $P(A) = \sum P(A_n)$

1.4 Classical/Geometrical Definitions of Probability

古典概型: 基本事件只有有限个且概率相同.

掷硬币 n 次,取每种排列为基本事件,即为古典概型:

$$P(\text{首次正面出现在 } k \text{ 次}) = \frac{1}{2^k}$$

掷硬币直到出现正面为止,基本事件 ω_k 为“首次正面出现在第 k 次”,则有无穷个基本事件,且概率不同. 利用可数可加性, $\omega_\infty = 0$,但不是不可能事件.

一般地,对于至多可数集合 Ω ,每个基本事件的概率都可求出时, $\forall A \subset \Omega, P(A) = \sum_{\omega \in A} P(\omega)$

若无限抛掷硬币,将排列作为基本事件,则有不可数个基本事件,此时若考虑等可能分析,则每个基本事件概率为 0.无法求出某个事件的概率(因为不可数个实数的和没有意义).

几何概型:随机现象的样本空间充满某个可测区域,且任一点落在度量相同的子区域内是等可能的.

古典/几何概型的另一个问题:Bertrand Paradox-圆内一弦长度超过正三角形边长的概率由三种解释.

原因:当可能结果有无穷个时,难以规定“等可能”这一概念,因此概率空间被模糊定义了.

Buffon's Needle:

分析做法:设针中点与最近平行线距 $x \in [0, \frac{d}{2}]$,与直线成角 $\varphi \in [0, \pi]$,在上区域中求 $x \leq \frac{l}{2} \sin \varphi$ 部分的概率. $P(A) = \frac{\int_0^\pi \frac{l}{2} \sin \varphi d\varphi}{\frac{d}{2} \pi} = \frac{2l}{d\pi}$

期望做法:相不相交是二点分布,只需求其期望. 由对称性及可加性,期望与长度成正比.而考虑直径为 d 的圆周扔下必有两交点,即期望为 2,由此可求出比例系数.

推广: 闭半圆扔下有几个交点? 考虑互补半圆,利用容斥原理.

1.5 Conditional Probability

概率空间 (Ω, \mathcal{F}, P) 中 $P(B) > 0$. 定义 $P_B(A) = P(A|B) = \frac{P(A \cap B)}{P(B)}$, 则可证 $(\Omega, \mathcal{F}, P_B)$ 也是概率空间.

乘法公式: $P(\bigcap_{i=1}^{n-1} A_i) > 0$ (使得条件概率有意义)时,由定义归纳可得

$$\Rightarrow P(\bigcap_{i=1}^n A_i) = \prod_{i=1}^n P(A_i | \bigcap_{j=1}^{i-1} A_j)$$

可靠性函数与风险率:设前 t 时刻正常, $[t, t + \Delta t]$ 时段失效的概率为 $\lambda(t)\Delta t + o(\Delta t)$,求设备在 $(0, t)$ 上无故障的概率.

设 A_t 表示设备在 $(0, t)$ 内正常, $P(\overline{A_{t+\Delta t}} | A_t) = \lambda(t)\Delta t + o(\Delta t)$.

$$p(t + \Delta t) = P(A_t)P(\overline{A_{t+\Delta t}} | A_t) = p(t)[1 - \lambda(t)\Delta t + o(\Delta t)] \Rightarrow \frac{dp(t)}{dt} = -\lambda(t)p(t)$$

注意到 $p(0) = 1$, 有 $p(t) = e^{-\int_0^t \lambda(s)ds}$

全概率公式:设 $B_1, B_2 \dots$ 为样本空间 Ω 的一个正划分,则

$$P(A) = \sum_{i=1}^{\infty} P(AB_i) = \sum_{i=1}^{\infty} P(B_i)P(A|B_i)$$

赌徒输光:两人各有赌资 $i, n-i$, 每次赌博胜者拿走对方 1 元, 胜率分别为 $p, 1-p$.

设 A_i 表示甲有 i 元, 最终破产. B 表示某次甲胜, $P(B) = p$, 则有 $P(A_i|B) = P(A_{i+1}), P(A_i|\bar{B}) = P(A_{i-1})$

于是 $P(A_i) = pP(A_{i+1}) + (1-p)P(A_{i-1})$, 边界 $P(A_0) = 1, P(A_n) = 0$.

$$P(A_i) = \begin{cases} 1 - \frac{1-r^i}{1-r^n} & p \neq \frac{1}{2} \\ 1 - \frac{i}{n} & p = \frac{1}{2} \end{cases} \quad r = \frac{1-p}{p}$$

对赌场 ($n \rightarrow \infty$), 甲最终会输光的概率为 $P(A_i) = \min\{1, r^i\}$

Polya 模型:从黑球, 红球中任取若干次. 取出的红球与黑球个数确定的情形下, 概率是否与次序相关.

若放回抽样, 结果不影响下次, 故概率相等.

若不放回抽样, 前次结果影响后次, 但概率仍与次序无关.

若放回若干同色球(传染病模型), 每次取出会增加下次取出同色球的概率. 但结果与次序无关.

若放回若干异色球(安全模型), 结果才与次序有关.

敏感问题问卷调查:在问卷上要求每个人准备一枚硬币, 对于指定的隐私题目, 请填写人投掷一次硬币: 如果正面朝上, 则如实填写个人的真实情况; 如果反面朝上, 那么就再投掷一次硬币, 正面就填“是”, 反面就填“否”. 当然, 若第一次投掷硬币为正的话, 填写人完全可以假装再投一次硬币来掩人耳目.

假设回收后有效问卷有 M 份, 其中该问题答“是”的有 N 个人. 如实填写了该问题的人平均有 $\frac{M}{2}$ 个; 在另外 $\frac{M}{2}$ 人中, 平均有 $\frac{M}{4}$ 人答的“是”. 因此, 我们所需要的最终结果应该为 $\frac{(N - M/4)}{M/2}$

$$\text{Bayes: } P(B_i|A) = \frac{P(B_i)P(A|B_i)}{P(A)} = \frac{P(B_i)P(A|B_i)}{\sum_{t=1}^n P(B_t)P(A|B_t)}$$

1.6 Independence of Events

定义: $P(AB) = P(A)P(B)$. 实际中以经验判断.

A, B 独立 $\Rightarrow A$ 与 \mathcal{F}_B 中任一事件独立.

多个事件相互独立: 直观想法— A 与 $\mathcal{F}_{B,C,\dots}$ 中任一事件独立.

定义: 其中任意 k 个事件的交的概率等于概率的乘积.

无穷个事件相互独立: 任意有限个事件相互独立.

$A_1 \cdots A_n$ 相互独立, 则任意对其分组, 各组事件分别产生的事件域相互独立.

$$\begin{aligned} \text{相关系数: } r(A, B) &= \frac{P(AB) - P(A)P(B)}{\sqrt{P(A)[1 - P(A)]P(B)[1 - P(B)]}} \\ \begin{cases} -P(A)P(B) \\ -[1 - P(A)][1 - P(B)] \end{cases} &\leq P(AB) - P(A)P(B) \leq \begin{cases} P(A)[1 - P(B)] \\ P(B)[1 - P(A)] \end{cases} \Rightarrow |r(A, B)| \leq 1 \\ r(A, B) = 1 &\Leftrightarrow P(A) = P(AB) = P(B) \\ r(A, B) > 0 &\Leftrightarrow P(A|B) > P(A) \Leftrightarrow P(B|A) > P(B) \end{aligned}$$

2 Random Variables

2.1 Definition

样本空间 $\Omega \rightarrow \mathbb{R}$ 的函数 $X = X(\omega)$ 称为**随机变量**. 值域有限或可列称为离散随机变量, 值域充满数轴上的某个区间, 称为连续随机变量. 记 $F(x) = P(X \leq x)$ 为 X 的**分布函数**.

显然, F 是 $(-\infty, \infty)$ 的单调不减函数, 有界, 于是各点有左右极限, 且无穷处有极限.

$$\begin{aligned} F(b) - F(c+0) &= F(b) - \lim_{a \rightarrow c^+} F(a) = \lim_{a \rightarrow c^+} P(a < X \leq b) \\ &= P(\bigcup_{a \rightarrow c^+} \{a < X \leq b\}) = P(c < X \leq b) \end{aligned}$$

$$F(d-0) - F(a) = P(a < X < d)$$

即 $F(c+0) = F(c)$ (右连续), $F(d-0) = P(X < d)$, 且可得 $F(-\infty) = 0, F(\infty) = 1$.

满足以上性质的函数 F 必定为某随机变量的分布函数.

任意 Borel 集 $B \subset \mathbb{R}$, $P(X \in B)$ 可由 F 计算得到.

特别的, $P(X = x) = F(x) - F(x-0)$, 若 F 在 x 连续, 则 $P(X = x) = 0$

p -分位数: 满足 $P(X \leq x) \geq p, P(X < x) \leq p$ 的 x .

对连续随机变量, 等价定义 $F(x) = p$ 的点为 p - (下侧) 分位数.

$$p = \frac{1}{2} \text{ 时称为中位数. } \Leftrightarrow \frac{1}{2} \leq F(x) \leq \frac{1}{2} + P(X = x)$$

中位数的统计意义: 使得 $E|X - a|$ 最小的 a

存在非负可积函数, 使得 $F(x) = \int_{-\infty}^x p(t)dt$, 则称 $p(x)$ 为 X 的**概率密度函数**.

在 $F(x)$ 导数存在的点有 $p(x) = F'(x)$, 其余点处 $p(x)$ 可任意取值.

密度公式: 函数 $g(x)$ 单调连续, $h(x) = g^{-1}(x)$ 连续可微, 则 $Y = g(X)$ 的密度函数为 $p_Y(y) = p_X(h(y))|h'(y)|$

$$\text{Proof: } F_Y(y) = P(g(X) \leq y) = \int_{h(-\infty, y]} p_X(x)dx \stackrel{x=h(z)}{=} \int_{-\infty}^y \text{RHS}$$

$$\text{应用: } X \sim \Gamma(\alpha, \lambda) \Rightarrow kX \sim \Gamma(\alpha, \frac{\lambda}{k}), 2\lambda X \sim \chi^2(2\alpha)$$

Diagonal: $F_X(x)$ 若为严格增的连续函数, $F_X(X) \sim U(0, 1)$

$$\text{Proof: } F_{F_X(X)}(x) = P(F_X(X) \leq x) = P(X \leq F_X^{-1}(x)) = F_X(F_X^{-1}(x)) = x$$

2.2 Expectation & Variance

$$E[X] = \int_{\mathbb{R}} x dF(x), \text{ 其中积分为 Lebesgue 积分.}$$

$$= \begin{cases} \sum x_i p(x_i) = \sum_{n=0}^{\infty} [P(X > n) - P(X < -n)], \text{ 离散, } \sum |x_i| P(X = x_i) < \infty \\ \int_{\mathbb{R}} x p(x) dx = \int_{\mathbb{R}^+} P(X > x) dx - \int_{\mathbb{R}^-} P(X < x) dx, \text{ 连续, } \int_{\mathbb{R}} |x| p(x) dx < \infty \end{cases}$$

- 绝对收敛保证了和的存在且与顺序无关.

特别地,当 $X > 0$ 时,有期望计算公式 $E[X] = \sum_{n=0}^{\infty} P(X > n)$. 本质是换一种切分方式求面积.

当期望存在时: $nP(X > n) = n \int_n^{\infty} dF(x) \leq \int_n^{\infty} x dF(x)$

上式取极限 $n \rightarrow \infty$, 得 $nP(X > n) \rightarrow 0$, 即 $\lim_{x \rightarrow \infty} x(1 - F(x)) = 0$

同理有 $\lim_{x \rightarrow -\infty} xF(x) = \lim_{x \rightarrow \infty} x(1 - F(x)) = 0$

由此极限可推出**期望的几何意义**: $\int_{-\infty}^{EX} F(x) dx = \int_{EX}^{\infty} (1 - F(x)) dx$

即: $y = F(x)$, $x = EX$, 将 $0 \leq y \leq 1$ 分成面积相等的两部分.

Proof: 对两边进行分部积分即可.

Cauchy-Schwartz:

$E[X^2], E[Y^2] < \infty$, 则 $(E[XY])^2 \leq E[X^2] E[Y^2]$

Proof: 考虑 $f(u) = E[Xu + Y]^2 = (E[X^2])u^2 + 2E[XY]u + E[Y^2]$ 的判别式即可.

期望的统计意义:

$E[(X - a)^2] = E[(X - EX)^2] + 2E[(X - EX)(EX - a)] + (EX - a)^2 = E[(X - EX)^2] + (EX - a)^2 \geq E[(X - EX)^2]$.

若 $E[X^2]$ 存在, 则定义 $\text{Var}[X] = E[X - EX]^2 = \begin{cases} \sum (x_i - EX)^2 p(x_i) \\ \int_{\mathbb{R}} (x - EX)^2 p(x) dx \end{cases}$

$\text{Var}[X] = E[X^2 - 2XEX + (EX)^2] = EX^2 - (EX)^2$

其他统计量:

变异系数 $C_v(X) = \frac{\sigma(X)}{EX} = \frac{\sqrt{\text{Var}[X]}}{EX}$. 消去了量纲的影响.

偏度系数 $\beta_s = \frac{E[X - EX]^3}{(\text{Var}[X])^{\frac{3}{2}}}$. 描述偏离对称性的程度.

峰度系数 $\beta_k = \frac{E[X - EX]^4}{(\text{Var}[X])^2} - 3$. 描述相比于正态分布的尖峭程度(尾部粗细). 正态分布峰度为 0.

2.3 Probability Generating Function

对一取非负整数的随机变量 X , 设 $p_k = P(X = k)$, 定义 $G_X(t) = E[t^X] = \sum_{k=0}^{\infty} p_k t^k$

注意到 $p_k = \frac{G^{(k)}(0)}{k!}$, $G(1) = 1$, $G(t)$ 对 $|t| \leq 1$ 绝对收敛.

$G(t)$ 在 $[0, 1]$ 上连续, 此时由 $G'(t) = \sum k p_k t^{k-1}$, $G''(t) = \sum k(k-1) p_k t^{k-2}$ 知, G 是递增的下凸函数.

$G^{(r)}(1) = E[X^r]$, 其中 x^r 为 x 的 r 次降阶乘.

$\Rightarrow \text{Var}[X] = G''(1) - G'(1)^2 + G'(1)$

X, Y 独立 $\Rightarrow G_{X+Y}(t) = G_X(t)G_Y(t)$, $G_{X-Y}(t) = G_X(t)G_Y(t^{-1})$

Compound Distribution: X_i i.i.d., N 是随机变量, $S_N = \sum_{k=1}^N X_k$

$$G_{S_N}(t) = \sum_{n=0}^{\infty} E[t^{S_N} | N=n] P(N=n) = \sum_{n=0}^{\infty} E[t^{S_n}] P(N=n) = \sum_{n=0}^{\infty} E[t^X]^n P(N=n) = G_N(G_X(t))$$

$$\Rightarrow E[S_N] = E[N] E[X], \text{Var}[S_N] = E[N] \text{Var}[X] + \text{Var}[N] E[X]^2$$

2.4 Characteristic Function

$$\varphi_X(t) = E[e^{itX}]$$

$$\varphi_{\vec{X}}(\vec{t}) = E[e^{i\vec{t}^T \vec{X}}]$$

界: $|\varphi_X(t)| = |E[e^{itX}]| \leq E[|e^{itX}|] = 1$

对称性: $\varphi_X(t)$ 是实偶函数 $\Leftrightarrow \varphi_X(t) = \overline{\varphi_X(t)} = E[\overline{e^{itX}}] = E[e^{-itX}] = \varphi_X(-t) = \varphi_{-X}(t) \Leftrightarrow F_X(x) = F_{-X}(x)$ 为对称分布 $\Leftrightarrow F_X(x)$ 关于 $(0, \frac{1}{2})$ 对称

线性变换: $\varphi_{a+bX}(t) = e^{iat} \varphi_X(bt)$

卷积变乘积: X, Y 独立, $\varphi_{X+Y}(t) = E[e^{it(X+Y)}] = \varphi_X(t) \varphi_Y(t)$

矩的计算: $\varphi^{(k)}(t) = \int_{\mathbb{R}} i^k x^k e^{itx} p(x) dx = i^k E[X^k e^{itX}] \Rightarrow \varphi^{(k)}(0) = i^k E X^k$

独立性判定: X_1, \dots, X_n 独立 $\Leftrightarrow \varphi_{\vec{X}}(\vec{t}) = \prod_{k=1}^n \varphi_{X_k}(t_k)$

分析性质:

$\varphi(x)$ 在 \mathbb{R} 上一致连续

对连续型随机变量 X , $\varphi_X(t) = \int_{\mathbb{R}} e^{itx} p(x) dx$ 为 $p(x)$ 的 Fourier 变换.

于是有逆变换: $F(x) = \lim_{y \rightarrow -\infty} \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-ity} - e^{-itx}}{it} \varphi(t) dt$

$$p(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \varphi_X(t) dt$$

即 $F(x)$ 与 $\varphi(t)$ 有一一对应的关系. 且连续性定理:

$$\{F_n(x)\} \text{弱收敛到 } F(x) \Leftrightarrow \{\varphi_n(x)\} \text{收敛到 } \varphi(x)$$

表明这种对应关系也是连续的

3 Distributions

3.1 Common Discrete Distributions

1. Bernoulli (Binomial): $b(n, p)$

$$P(X=k) = C_n^k p^k (1-p)^{n-k}, k=0, \dots, n$$

$$EX = np, \text{Var}[X] = E[X(X-1)] + EX - (EX)^2 = np(1-p)$$

$$b(n, p) = n * b(1, p) \Rightarrow \varphi(t) = (1 - p + pe^{it})^n$$

概率最大值发生在 $k = \begin{cases} (n+1)p, (n+1)p - 1 & , (n+1)p \in \mathbb{N} \\ \lfloor (n+1)p \rfloor & , (n+1)p \notin \mathbb{N} \end{cases}$

2. **Poisson:** $P(\lambda), (\lambda > 0)$

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

概率最大值发生在 $k = \lfloor \lambda \rfloor$

$$EX = \text{Var}[X] = \lambda, \varphi(t) = \sum_{k=0}^{\infty} e^{itk} \frac{\lambda^k}{k!} e^{-\lambda} = e^{\lambda(e^{it}-1)}$$

Poisson 定理, 对二项分布 $b(n, p_n), \lim_{n \rightarrow \infty} np_n = \lambda > 0$,

$$\begin{aligned} P(X = k) &= C_n^k p_n^k (1 - p_n)^{n-k} \\ &= \frac{n(n-1) \cdots (n-k+1)}{k! n^k} (np_n)^k \left(1 - \frac{np_n}{n}\right)^{n-k} \\ &= \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) \frac{[\lambda + o(1)]^k}{k!} \left[1 - \frac{\lambda + o(1)}{n}\right]^{n-k} \\ &\rightarrow \frac{\lambda^k}{k!} e^{-\lambda}, n \rightarrow \infty \end{aligned}$$

3. **Geometry:** $G(p)$

$$P(X = k) = p(1-p)^{k-1}, k \in \mathbb{N}^+$$

$$EX = \sum_{n=1}^{\infty} \sum_{k=1}^n p(1-p)^{n-1} = \sum_{k=1}^{\infty} \sum_{n=k}^{\infty} p(1-p)^{n-1} = \frac{1}{p}$$

$$\text{类似方法使用两次求出 } E[X(X-1)] \Rightarrow \text{Var}[X] = \frac{1-p}{p^2}$$

$$\varphi(t) = \frac{pe^{it}}{1 - (1-p)e^{it}}$$

$$\text{尾概率 } P(X > m) = (1-p)^m$$

$$\text{无记忆性} \Leftrightarrow P(X > m+n) = P(X > m)P(X > n) \Leftrightarrow X \sim G(P(X \leq 1))$$

(即解 Cauchy 方程)

4. **HyperGeometry:** $h(n, N, M), (n, M \leq N)$

意义: N 件物品含有 M 件次品, 不放回抽取 n 次得到的次品数.

$$P(X = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, k \in [\max\{0, n-N+M\}, \min\{M, n\}]$$

$$EX = \sum_k k \frac{C_M^k C_{N-M}^{n-k}}{C_N^n} = \frac{Mn}{N} \sum_k \frac{C_{M-1}^{k-1} C_{N-M}^{n-k}}{C_{N-1}^{n-1}} = \frac{Mn}{N}$$

(注意到每次的期望都是 $\frac{M}{N}$ 即可)

$$\text{类似地使用 Vandermonde Convolution, 有 } E[X(X-1)] = \frac{M(M-1)n(n-1)}{N(N-1)}$$

$$\Rightarrow \text{Var}[x] = \frac{nM(N-M)(N-n)}{N^2(N-1)}$$

二项逼近: 当 $N \rightarrow \infty, \frac{M}{N} \rightarrow p$ 时, $h(n, N, M) \rightarrow b(N, p)$

5. Pascal (Negative Binomial): $Nb(r, p)$

意义:事件发生第 r 次时的实验次数. $Nb(r, p) = r * G(p)$

$$P(X = k) = C_{k-1}^{r-1} p^r (1-p)^{k-r}, k = r, r+1, \dots$$

$$EX = \frac{r}{p}, \text{Var}[X] = \frac{r(1-p)}{p^2}$$

$$\varphi(t) = \left(\frac{pe^{it}}{1 - (1-p)e^{it}} \right)^r$$

Banach's matchbox problem:

左右各 n 根火柴随机用,发现一边用完时另一边还还有 k 根的概率 $p_{n,k}$:

事件(的一半)相当于:成功 $n+1$ 次时恰已有 $n-k$ 次失败,设 $Y \sim Nb(n+1, \frac{1}{2})$

$$\text{立即有: } p_{n,k} = 2P(Y = 2n - k + 1) = C_{2n-k}^n 2^{-2n+k}$$

6. Fixed Points

See my paper at http://learn.tsinghua.edu.cn:8080/2011011271/fixed_points.pdf

3.2 Common Continuous Distributions

1. Uniformed: $U(a, b)$

$$p(x) = \frac{1}{b-a} I_{[a,b]}(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{else} \end{cases} \text{ 其中 } I \text{ 为区间示性函数.}$$

$$EX = \frac{a+b}{2}, \text{Var}[X] = \frac{(b-a)^2}{12}$$

$$U(-1, 1) \text{ 的特征函数 } \varphi_{X'}(t) = \int_0^1 \cos(tx) dx = \frac{\sin t}{t} \Rightarrow \varphi(t) = e^{\frac{(a+b)it}{2}} \frac{\sin(\frac{(b-a)t}{2})}{\frac{(b-a)t}{2}}$$

$$\text{或: } \varphi(t) = \int_a^b \frac{e^{itx}}{b-a} dx = \frac{e^{ibt} - e^{iat}}{it(b-a)}, \text{ 与上式其实是相等的.}$$

2. Exponential: $Exp(\lambda), (\lambda > 0)$

$$p(x) = \lambda e^{-\lambda x} I_{\mathbb{R}^+}(x)$$

$$F(x) = 1 - e^{-\lambda x}$$

设备在时刻 t 的失效率 $\lambda(t) = \lambda$ 为常数,则寿命 $X \sim Exp(\lambda)$

$$EX = \int_{\mathbb{R}^+} P(X > x) dx = \int_{\mathbb{R}^+} e^{-\lambda x} dx = \frac{1}{\lambda}$$

$$EX^2 = \int_{\mathbb{R}^+} P(X^2 > x) dx \stackrel{x=u^2}{=} \frac{2}{\lambda} \int_{\mathbb{R}^+} u \lambda e^{-\lambda u} du = \frac{2EX}{\lambda} = \frac{2}{\lambda^2}$$

$$\text{Var}[X] = EX^2 - (EX)^2 = \frac{1}{\lambda^2}$$

$$\varphi(t) = \lambda \int_{\mathbb{R}^+} e^{itx - \lambda x} dx = \frac{\lambda}{\lambda - it}$$

无记忆性 $\Leftrightarrow P(X > m+n) = P(X > m)P(X > n) \Leftrightarrow X \sim Exp(\lambda)$

$(0, t]$ 发生的次数 $\sim P(\lambda t)$, 则第一次发生的时间 $\sim Exp(\lambda)$

$$X \sim U(0, 1) \Rightarrow \frac{-\ln X}{\lambda} \sim \text{Exp}(\lambda)$$

$$X, Y \sim \text{Exp}(\frac{1}{\lambda}) \Rightarrow X - Y \sim L(0, \lambda), |X - Y| \sim \text{Exp}(\frac{1}{\lambda})$$

3. Gauss (Normal): $N(\mu, \sigma^2)$

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\text{对 } Y \sim N(\mu, \sigma^2), \text{ 标准化: } X = \frac{Y - \mu}{\sigma}, \text{ 则 } X \sim N(0, 1)$$

$$P(X \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \stackrel{\text{def}}{=} \Phi(x)$$

利用 Poisson 积分, 可验证 $\Phi(\infty) = 1$

$$P(|Y - \mu| < k\sigma) = \Phi(k) - \Phi(-k) = 2\Phi(k) - 1$$

$$EX = \int_{\mathbb{R}} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0$$

$$\text{Var}[X] = EX^2 = 1. \text{ 分部积分}$$

$$EX^{2k+1} = 0, EX^{2k} = (2k-1)!!$$

$$\varphi_X(t) = 2 \int_{\mathbb{R}^+} \cos(tx) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx,$$

$$\text{求导有 } \frac{d\varphi_X(t)}{dt} = 2 \int_{\mathbb{R}^+} \sin(tx) d\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\right) = -t\varphi_X(t) \Rightarrow \varphi_X(t) = e^{-\frac{t^2}{2}}.$$

$$\text{于是, } \varphi_Y(t) = e^{i\mu t - \frac{\sigma^2 t^2}{2}}$$

$$\text{误差函数: } \text{Erf}(x) = 2\Phi(x\sqrt{2}) - 1$$

4. Gamma: $\Gamma(\alpha, \lambda)$, ($\alpha > 0$ 为形状参数, $\lambda > 0$ 为尺度参数)

$$\Gamma \text{ 函数: } \Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$$

$$p(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} I_{[0, \infty)}(x)$$

$$EX = \frac{\lambda^\alpha}{\Gamma(\alpha)} \Gamma(\alpha+1) = \frac{\alpha}{\lambda}$$

$$EX^2 = \frac{\alpha(\alpha+1)}{\lambda^2}, \text{Var}[X] = \frac{\alpha}{\lambda^2}$$

$$\Gamma(1, \lambda) = \text{Exp}(\lambda), \Gamma\left(\frac{n}{2}, \frac{1}{2}\right) = \chi^2(n)$$

$$\varphi(t) = \left(\frac{\lambda - it}{\lambda}\right)^\alpha$$

5. Beta: $Be(a, b)$, ($a, b > 0$)

$$\text{Beta 函数: } B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

$$p(x) = \frac{1}{B(a, b)} x^{a-1} (1-x)^{b-1} I_{(0,1)}(x)$$

$$EX = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+b+1)} = \frac{a}{a+b}$$

$$EX^2 = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a+2)\Gamma(b)}{\Gamma(a+b+2)} = \frac{a(a+1)}{(a+b)(a+b+1)}$$

$$\text{Var}[X] = \frac{ab}{(a+b)^2(a+b+1)}$$

$$\text{特例: } p(x) = nx^{n-1} \Leftrightarrow X \sim Be(n, 1); p(x) = \frac{x^{n-1}(1-x)}{n+n^2} \Leftrightarrow X \sim Be(n, 2)$$

6. **Chi Square:** $\chi^2(k)$. (自由度 $k > 0$)

$$Y_1 \stackrel{i.i.d}{=} \dots \stackrel{i.i.d}{=} Y_k \sim N(0, 1) \Rightarrow X = \sum Y^2 \sim \chi^2(k)$$

$$p(x) = \frac{x^{\frac{k}{2}-1} e^{-\frac{x}{2}}}{2^{k/2} \Gamma(\frac{k}{2})} I_{[0, +\infty)}(x)$$

$$EX = k, \text{Var}[X] = 2k$$

$$\varphi(t) = (1 - 2it)^{-\frac{n}{2}}$$

7. **Logarithmic Normal:** $LN(\mu, \sigma^2)$

$$X \sim N(\mu, \sigma^2) \Rightarrow Y = e^X \sim LN(\mu, \sigma^2)$$

$$p(x) = \frac{1}{\sqrt{2\pi x \sigma}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} I_{\mathbb{R}^+}(x)$$

$$EX = e^{\mu + \frac{\sigma^2}{2}}, \text{Var}[X] = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1). \text{中位数 } e^\mu$$

8. **Cauchy:** $Cauchy(\mu, \lambda)$.

$$Y \stackrel{i.i.d}{=} Z \sim N(0, 1) \Rightarrow \frac{Y}{Z} \sim Cauchy(0, 1)$$

$$p(x) = \frac{\lambda}{\pi(\lambda^2 + (x - \mu)^2)}$$

期望与方差不存在.

$$\varphi(t) = e^{i\mu t - \lambda|t|}$$

9. **Fisher-Snedecor:** $F(m, n)$

$$\text{独立的 } X \sim \chi^2(m), Y \sim \chi^2(n). F = \frac{X/m}{Y/n} \sim F(m, n).$$

$$p_F(x) = \frac{1}{B(\frac{m}{2}, \frac{n}{2})} \left(\frac{m}{n}\right)^{\frac{m}{2}} x^{\frac{m}{2}-1} \left(1 + \frac{mx}{n}\right)^{-\frac{m+n}{2}} I_{[0, +\infty)}$$

$$EX = \frac{n}{n-2} (n > 2), \text{Var}[X] = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)} (n > 4)$$

$$\frac{mF}{n+mF} \sim B\left(\frac{m}{2}, \frac{n}{2}\right)$$

$$X \sim F(k_1, k_2) \Leftrightarrow \frac{1}{X} \sim F(k_2, k_1), F_{1-a}(k_1, k_2) = \frac{1}{F_a(k_2, k_1)}$$

注意到 $\frac{1}{m}\chi^2(m)$ 是 $m+1$ 个样本的样本方差.

10. **Student t:** $t(n)$. Student 是其论文笔名(1908)

$$\text{独立的 } X \sim N(0, 1), Y \sim \chi^2(n). t = \frac{X}{\sqrt{\frac{Y}{n}}} \sim t(n)$$

$$\text{由对称性, } F_t(x) = F_t(0) + \frac{1}{2}P(t^2 \leq x^2) = \frac{1}{2} + \frac{1}{2}F_{t^2}(x^2)$$

$$\text{而 } t^2 \sim F(1, n). \text{ 于是 } p_t(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})}(1 + \frac{x^2}{n})^{-\frac{n+1}{2}}$$

$$t(1) = \text{Cauchy}(0, 1), t(\infty) = N(0, 1)$$

$$EX = 0(n > 1), \text{Var}[X] = \frac{n}{n-2}(n > 2)$$

$$\text{注意到对于正态样本, 有 } \frac{\sqrt{n}(\bar{x} - \mu)}{\sigma} \sim N(0, 1)$$

$$\Rightarrow \frac{\sqrt{n}(\bar{x} - \mu)}{s} = \frac{N(0, 1)}{\sqrt{\frac{(n-1)s^2}{\sigma^2(n-1)}}} = \frac{N(0, 1)}{\sqrt{\frac{\chi^2(n-1)}{n-1}}} \sim t(n-1)$$

11. **Weibull:** $W(\lambda, k)$

$$\text{设备失效率} = \lambda t^{k-1} \Rightarrow \text{寿命服从 } W(\lambda, k).$$

$$W(\lambda, 1) = \text{Exp}(\lambda)$$

$$p(x) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k} I_{[0, +\infty)}(x)$$

$$F(x) = 1 - e^{-\left(\frac{x}{\lambda}\right)^k}$$

$$EX = \lambda \Gamma(1 + \frac{1}{k}), \text{Var}[X] = \lambda^2 \Gamma(1 + \frac{2}{k}) - (EX)^2$$

12. **Rayleigh:** $R(\sigma)$

$$X \stackrel{\text{i.i.d.}}{=} Y \sim N(0, \sigma^2) \Rightarrow \sqrt{X^2 + Y^2} \sim R(\sigma)$$

$$p(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} I_{[0, +\infty)}(x)$$

$$F(x) = 1 - e^{-\frac{x^2}{2\sigma^2}}$$

$$EX = \sigma \sqrt{\frac{\pi}{2}}, \text{Var}[X] = \frac{4-\pi}{2} \sigma^2$$

13. **Laplace:** $L(\mu, b)$

$$p(x) = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}}, F(x) = \frac{1 + \text{sgn}(x-\mu)}{2} (1 - e^{-\frac{|x-\mu|}{b}})$$

$$EX = \mu, \text{Var}[X] = 2b^2$$

$$\varphi(x) = \frac{e^{\mu it}}{1 + b^2 t^2}$$

14. **Pareto** $p(x) = (\alpha - 1)x_0^{\alpha-1}x^{-\alpha}I_{[x_0, +\infty)}, \alpha > 2, x_0 > 0$

$$EX = \frac{\alpha-1}{\alpha-2}x_0, \text{Var}[X] = \frac{(\alpha-1)x_0^2}{(\alpha-3)(\alpha-2)^2} (\alpha > 3)$$

4 Multivariate Distributions

4.1 Joint Distribution

联合分布函数: $F(x_1, x_2, \dots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n)$

二维联合分布函数 $F(x, y) \Leftrightarrow$

1. 对每个变元单调非减, 右连续;
2. $F(-\infty, y) = F(x, -\infty) = 0 \leq F(x, y) \leq 1 = F(\infty, \infty)$;
3. $P(a < X \leq b, c < Y \leq d) = F(b, d) - F(a, d) - F(b, c) + F(a, c) \geq 0$.

边际分布函数: $F_X(x) = F(x, \infty), F_Y(y) = F(\infty, y)$

联合密度函数: 存在非负函数 $p(x, y), F(x, y) = \int_{-\infty}^x \int_{-\infty}^y p(u, v) du dv$

边际密度函数: 当联合密度函数存在时, 每个边际分布函数对应的密度

$$p_X(x) = \int_{\mathbb{R}} p(x, y) dy, p_Y(y) = \int_{\mathbb{R}} p(x, y) dx$$

在 $F(x, y)$ 偏导数存在处有 $p(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y)$

独立: $F(x_1, \dots, x_n) = \prod F_i(x_i) \Leftrightarrow \begin{cases} P(X_1 = x_1, \dots, X_n = x_n) = \prod P(X_i = x_i) \\ p(x_1, \dots, x_n) = \prod p_i(x_i) \end{cases}$

4.2 Common Multivariate Distributions

1. 多项分布

n 次实验, 每次有 r 种可能结果, 概率分别为 $p_1 \cdots p_r$. 以各种结果出现的次数作为随机变量, 有

$$P(X_1 = n_1, \dots, X_r = n_r) = \frac{n!}{\prod n_i!} \prod p_i^{n_i}$$

概率是多项式 $(\sum p_i x_i)^n$ 展开式中的系数.

$$\begin{aligned} P(X_1 = n_1) &= \frac{n!}{n_1!(n - n_1!)} p_1^{n_1} \sum_{n_2 + \dots + n_r = n - n_1} \frac{(n - n_1)!}{\prod_{i=2}^r n_i!} \prod_{i=2}^r p_i^{n_i} \\ &= \frac{n!}{n_1!(n - n_1!)} p_1^{n_1} \left(\sum_{i=2}^r p_i \right)^{n - n_1} \\ &= \frac{n!}{n_1!(n - n_1!)} p_1^{n_1} (1 - p_1)^{n - n_1}. \text{边缘分布为二项分布.} \end{aligned}$$

2. 多维超几何分布

N 个球中, i 号球有 N_i 个, 任取 n 个, 其中各号球的个数作为随机变量, 有 $P(X_1 = n_1, \dots, X_r =$

$$n_r) = \frac{\prod C_{N_i}^{n_i}}{C_N^n}$$

3. 多维均匀分布 $U(D)$

D 为 \mathbb{R}^n 的有界可测子集, 测度为 $S_D, p(x_1 \cdots x_n) = \frac{1}{S_D} I_D(x_1 \cdots x_n)$

4. 二维指数分布

$$F(x, y) = (1 - e^{-x} - e^{-y} + e^{-x-y-\lambda xy})I_{\mathbb{R}^+ \times \mathbb{R}^+}(x, y)$$

边际分布为 $Exp(1)$

5. 二维正态分布 $N(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$, ($|\rho| < 1$)

$$p(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{Q}{2(1-\rho^2)}}$$

$$\text{其中 } Q = \frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}$$

由 $|\rho| < 1$ 可知 Q 为正定二次型.

$$\begin{aligned} \text{边际分布: } p_X(x) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{\mathbb{R}} e^{-\frac{Q}{2(1-\rho^2)}} dy \\ &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \int_{\mathbb{R}} e^{-\frac{T^2}{2}} dy \quad (T = \rho\frac{x-\mu_1}{\sigma_1\sqrt{1-\rho^2}} - \frac{y-\mu_2}{\sigma_2\sqrt{1-\rho^2}}) \\ &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \sigma_2\sqrt{1-\rho^2} \int_{\mathbb{R}} e^{-\frac{t^2}{2}} dt \\ &= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \text{ 为一维正态分布.} \end{aligned}$$

设 $(X, Y) \sim N(0, 0, 1, 1, \rho)$

$$\text{则 } p(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} g(x, y)$$

其中 $z(x) = g(x, y)$ 为 $Z \sim N(\rho y, 1-\rho^2)$ 的密度函数.

\Rightarrow 条件正态分布: (直观解释: 考虑 $\rho = 0, 1$ 时的情形)

$$\begin{aligned} \text{对 } (X, Y) \sim N(0, 0, 1, 1, \rho) &\Rightarrow X|Y = y \sim N(\rho y, 1-\rho^2) \\ &\Rightarrow E[X|Y] = \rho Y, E[X|Y = y] = \rho y \end{aligned}$$

对 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

$$\Rightarrow X|Y = y \sim N(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y - \mu_2), \sigma_1^2(1-\rho^2))$$

相关系数:

$$\begin{aligned} r(X, Y) &= \text{Cov}[X, Y] = E[XY] - EXEY = \iint_{\mathbb{R}^2} xy p(x, y) dx dy \\ &= \int_{\mathbb{R}} \frac{y}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \int_{\mathbb{R}} x g(x, y) dx dy \\ &= \int_{\mathbb{R}} EZ \frac{y}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\ &= \rho EY^2 = \rho \text{Var}[Y] = \rho \end{aligned}$$

做线性变换可知, 任意二维正态分布的相关系数为 ρ .

$\rho = 0$ 时易证 $p(x, y) = p_X(x)p_Y(y)$, 于是, 对于联合正态分布, 不相关与独立等价.

最大值期望: $E[\max\{X, Y\}] = E\left[\frac{X+Y+|X-Y|}{2}\right] = \sqrt{\frac{1-\rho}{\pi}}$

椭圆域内概率: $D = \{(x, y) : Q \leq t\}$, 则 $P((X, Y) \in D) = 1 - e^{-\frac{t}{2(1-\rho^2)}}$

证: 先做仿射变换 $u = \frac{x-\mu_1}{\sigma_1} - \rho \frac{y-\mu_2}{\sigma_2}, v = \frac{y-\mu_2}{\sigma_2} \sqrt{1-\rho^2}, u^2 + v^2 = Q$

$$J = \begin{vmatrix} \frac{1}{\sigma_1} & 0 \\ -\frac{\rho}{\sigma_2} & \frac{\sqrt{1-\rho^2}}{\sigma_2} \end{vmatrix} = \frac{\sqrt{1-\rho^2}}{\sigma_1 \sigma_2}$$

于是, LHS = $\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \iint_D e^{-\frac{Q}{2(1-\rho^2)}} dx dy = \frac{1}{2\pi(1-\rho^2)} \iint_{u^2+v^2 \leq t} e^{-\frac{u^2+v^2}{2(1-\rho^2)}} du dv$

再做变换 $u = r \sin \theta, v = r \cos \theta, |J^{-1}| = r$, 有:

$$\text{LHS} = \frac{1}{2\pi(1-\rho^2)} \int_0^{2\pi} d\theta \int_0^{\sqrt{t}} r e^{-\frac{r^2}{2(1-\rho^2)}} dr = 1 - e^{-\frac{t}{2(1-\rho^2)}}$$

4.3 Convolution

和的分布(卷积):

$$P(X+Y=k) = \sum_{i \in \mathbb{Z}} P(X=i)P(y=k-i)$$

(X, Y) 的联合密度函数为 $p(x, y)$, 有:

$$F_{X+Y}(z) = \iint_{x+y \leq z} p(x, y) dx dy \Rightarrow p_{X+Y}(z) = \int_{\mathbb{R}} p(z-t, t) dt$$

$$X, Y \text{ 独立时, 有: } F_{X+Y}(z) = \iint_{x+y \leq z} p_X(x)p_Y(y) dx dy$$

$$\Rightarrow p_{X+Y}(z) = \int_{\mathbb{R}} p_X(z-y)p_Y(y) dy = \int_{\mathbb{R}} p_X(x)p_Y(z-x) dx$$

1. **Poisson 分布** $P(\lambda_1) * P(\lambda_2) = P(\lambda_1 + \lambda_2)$

$$\Leftarrow \sum_{i=0}^k \frac{\lambda_1^i e^{-\lambda_1}}{i!} \frac{\lambda_2^{k-i} e^{-\lambda_2}}{(k-i)!} = \frac{e^{-\lambda_1-\lambda_2}}{k!} \sum_{i=0}^k C_k^i \lambda_1^i \lambda_2^{k-i} = \frac{e^{-\lambda_1-\lambda_2}}{k!} (\lambda_1 + \lambda_2)^k$$

(Raikov) 独立变量的和服从 Poisson 分布, 则每个都服从 Poisson 分布.

2. **二项分布** $b(n, p) * b(m, p) = b(m+n, p)$

$$\begin{aligned} \Leftarrow \sum_{i=0}^k P(X=i)P(Y=k-i) &= \sum_{i=\max\{0, k-m\}}^{\min\{n, k\}} C_n^i p^i (1-p)^{n-i} C_m^{k-i} p^{k-i} (1-p)^{m-(k-i)} \\ &= p^k (1-p)^{m+n-k} \sum C_n^i C_m^{k-i} \\ &= C_{m+n}^k p^k (1-p)^{m+n-k} \end{aligned}$$

3. **Gamma 分布** $\Gamma(\alpha_1, \lambda) * \Gamma(\alpha_2, \lambda) = \Gamma(\alpha_1 + \alpha_2, \lambda)$

$$\begin{aligned}\Leftarrow p_{X+Y}(z) &= \frac{\lambda^{\alpha_1+\alpha_2} e^{-\lambda(z-y)} e^{-\lambda y}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \int_0^z (z-y)^{\alpha_1-1} y^{\alpha_2-1} dy \\ &\stackrel{y=zt}{=} \frac{\lambda^{\alpha_1+\alpha_2} e^{-\lambda z}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} z^{\alpha_1+\alpha_2-1} \int_0^1 (1-t)^{\alpha_1-1} t^{\alpha_2-1} dt \\ &= \frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1 + \alpha_2)} z^{\alpha_2+\alpha_2-1} e^{\lambda z}\end{aligned}$$

\Rightarrow 卡方分布 $\chi^2(m) * \chi^2(n) = \chi^2(m+n)$

4. **正态分布** $N(\mu_1, \sigma_1^2) * N(\mu_2, \sigma_2^2) = N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

$$\begin{aligned}\Leftarrow p_{X+Y}(z) &= \frac{1}{2\pi\sigma_1\sigma_2} \int_{\mathbb{R}} e^{-\frac{Q}{2}} dy \quad (Q = \frac{(z-y-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2}) \\ &\stackrel{\text{对}Q\text{配方}}{=} \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{(z-\mu_1-\mu_2)^2}{2(\sigma_1^2+\sigma_2^2)}} \int_{\mathbb{R}} e^{-\frac{A}{2}(y-T)^2} dy, (A = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}) \\ &= \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{(z-\mu_1-\mu_2)^2}{2(\sigma_1^2+\sigma_2^2)}} \sqrt{\frac{2\pi}{A}} \\ &= \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} e^{-\frac{(z-\mu_1-\mu_2)^2}{2(\sigma_1^2+\sigma_2^2)}}\end{aligned}$$

(Cramér)独立变量的和服从正态分布,则每个都服从正态分布.

4.4 Other Functions on Random Variables

期望: $E[g(X, Y)] = \begin{cases} \sum_{i,j} g(x_i, y_j) P(X = x_i, Y = y_j) \\ \iint_{\mathbb{R}^2} g(x, y) p(x, y) dx dy \end{cases}$

期望的可加性对任意随机变量成立:

$$E[X + Y] = \int_{\mathbb{R}} x \left(\int_{\mathbb{R}} p(x, y) dy \right) dx + \int_{\mathbb{R}} y \left(\int_{\mathbb{R}} p(x, y) dx \right) dy = EX + EY$$

最值: $Y = \max\{X_1, \dots, X_n\}, F_Y(x) = \prod F_i(x)$

$$Z = \min\{X_1, \dots, X_n\}, F_Z(x) = 1 - \prod (1 - F_i(x))$$

当 X_1, \dots, X_n i.i.d. 时, $p_Y(x) = nF(x)^{n-1}p(x), p_Z(x) = n(1 - F(x))^{n-1}p(x)$

第 k 小值(次序统计量): 设 X_1, \dots, X_n i.i.d., 设“第 k 小值”这个随机变量为 M_k

$$p_{M_k}(x) = n! \frac{F(x)^{k-1}}{(k-1)!} \frac{[1 - F(x)]^{n-k}}{(n-k)!} p(x)$$

$$i < j, p_{M_i, M_j}(x, y) = n! \frac{F(x)^{i-1}}{(i-1)!} \frac{[F(y) - F(x)]^{j-i-1}}{(j-i-1)!} \frac{[1 - F(y)]^{n-j}}{(n-j)!} p(x)p(y)$$

$$p_{M_1, \dots, M_n}(x_1, \dots, x_n) = \begin{cases} n! \prod p(x_i), & \text{if } x_1 < \dots < x_n \\ 0, & \text{else} \end{cases}$$

设极差 $R_n = M_n - M_1$, 由 p_{M_1, M_n} 做 Jacobi 可得 p_{R_n, M_1}

$$\Rightarrow F_{R_n}(x) = \int_{\mathbb{R}} np(t)(F(x+t) - F(t))^{n-1} dt$$

双射: $\begin{cases} U = g(X, Y) \\ V = h(X, Y) \end{cases}, p_{U,V}(u, v) = p_{X,Y}(x(u, v), y(u, v)) \left| \frac{\partial(x, y)}{\partial(u, v)} \right|$

独立积: $Z = XY$

设 $T = Y$, 利用二维双射, 有 $p_{Z,T}(z, t) = \frac{p_X(\frac{z}{t})p_Y(t)}{|t|}$. 对 t 积分即得 $p_Z(z)$

独立商: $Z = \frac{X}{Y}$

设 $T = Y$, $p_{Z,T}(z, t) = p_X(zt)p_Y(t)|t|$. 对 t 积分.

线性变换: $\vec{Y} = \mathbf{A}\vec{X} + \vec{B}$

$$p_{\vec{Y}}(\vec{x}) = p_{\vec{X}}(\mathbf{A}^{-1}\vec{x} - \mathbf{A}^{-1}\vec{B})|\det \mathbf{A}^{-1}|$$

正态分布的标准化: 设 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

$$\text{做变换} \begin{pmatrix} X' \\ Y' \end{pmatrix} = \begin{pmatrix} \frac{1}{\sigma_1} & 0 \\ 0 & \frac{1}{\sigma_2} \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} - \begin{pmatrix} \frac{\mu_1}{\sigma_1} \\ \frac{\mu_2}{\sigma_2} \end{pmatrix}$$

则 $(X', Y') \sim N(0, 0, 1, 1, \rho)$

4.5 Correlation

对独立的 X_1, \dots, X_n , 显然有 $E[\prod X_i] = \prod EX_i$, $\text{Var}[\sum a_i X_i] = \sum a_i^2 \text{Var}[X_i]$

对任意的 X, Y , 有 $\text{Var}[X \pm Y] = \text{Var}[X] + \text{Var}[Y] \pm 2E[(X - EX)(Y - EY)]$.

记协方差 $\text{Cov}[X, Y] = E[(X - EX)(Y - EY)] = E[XY] - EXEY$ (若 $E[XY]$ 存在)

$$\text{记相关系数 } r(X, Y) = \frac{\text{Cov}[X, Y]}{\sigma(X)\sigma(Y)}$$

独立 $\Rightarrow E[XY] = EXEY \Leftrightarrow r(X, Y) = 0 \Leftrightarrow$ 不相关

$r(X, Y) = \pm 1 \Leftrightarrow X, Y$ 几乎处处线性相关 $\Leftrightarrow P(Y = aX + b) = 1$

$E[XY]$ 是对角线上非负的对称双线性函数, 可作为随机变量的内积的定义.

因此, 由 Cauchy's Inequality:

$$\text{Cov}[X, Y]^2 = |\langle X - EX, Y - EY \rangle|^2 \leq |X - EX||Y - EY| = \text{Var}[X] \text{Var}[Y]$$

$$\Leftrightarrow |r(X, Y)| \leq 1.$$

r 可看做 $X - EX, Y - EY$ 的夹角余弦, EX 看做 X 在常数子空间上的投影.

将对角线上非负的对称双线性函数 $\text{Cov}[X, Y]$ 看做内积, 同样可得 $|r(X, Y)| \leq 1$.

线性回归: 求使 $E[X - (aY + b)]^2$ 最小的 (a, b) :

$$\text{考虑内积空间中垂直最小, } E[(X - (\hat{a}Y + \hat{b}))(aY + b)] = 0$$

$$\Rightarrow E[X - (\hat{a}Y + \hat{b})] = 0, E[Y(X - (\hat{a}Y + \hat{b}))] = 0.$$

$$\text{解得 } \hat{a} = \frac{\text{Cov}[X, Y]}{\text{Var}[Y]}, \hat{b} = EX - \hat{a}EY. \quad \hat{X} = \frac{\text{Cov}[X, Y]}{\text{Var}[Y]}(Y - EY) + EX$$

$$\begin{aligned} \text{此时误差 } E[X - \hat{X}]^2 &= E\left[(X - EX) - \frac{\text{Cov}[X, Y]}{\text{Var}[Y]}(Y - EY)\right]^2 \\ &= \text{Var}[X] - \frac{\text{Cov}[X, Y]^2}{\text{Var}[Y]} \\ &= \text{Var}[X][1 - r(X, Y)^2] \end{aligned}$$

协方差矩阵 $\vec{X} = (X_1, \dots, X_n)'$, 每个分量期望都存在, 则随机向量 \vec{X} 有期望.
 定义协方差矩阵为 $\text{Cov}[\vec{X}] = (\text{Cov}[X_i, X_j])_{n \times n} = \text{E}[(\vec{X} - \text{E}\vec{X})(\vec{X} - \text{E}\vec{X})']$
 $\text{Var}[\sum X_i] = \sum \text{Var}[X_i] + \sum_{1 \leq i < j \leq n} \text{Cov}[X_i, X_j]$ 为矩阵中所有元素和.
 $\text{Cov}[\vec{X}]$ 为半正定对称阵: 任取向量 $\vec{\alpha}$,

$$\begin{aligned} \vec{\alpha}' \text{Cov}[\vec{X}] \vec{\alpha} &= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \text{Cov}[X_i, X_j] \\ &= \sum_{i=1}^n \sum_{j=1}^n \text{E}[\alpha_i (X_i - \text{E}X_i) \alpha_j (X_j - \text{E}X_j)] \\ &= \text{E} \left[\sum_{i=1}^n \sum_{j=1}^n \alpha_i (X_i - \text{E}X_i) \alpha_j (X_j - \text{E}X_j) \right] \\ &= \text{E} \left[\sum_{i=1}^n \alpha_i (X_i - \text{E}X_i) \sum_{j=1}^n \alpha_j (X_j - \text{E}X_j) \right] \\ &= \text{E} \left[\sum_{i=1}^n \alpha_i (X_i - \text{E}X_i) \right]^2 \geq 0 \end{aligned}$$

4.6 Gauss Distribution

$\vec{Z} = (Z_1, \dots, Z_n)^T$, $Z_i \sim N(0, 1)$ 相互独立, $A = (a_{ij})_{m \times n}$, $\vec{\mu} = (\mu_1, \dots, \mu_m)^T$, 则 $\vec{X} = A\vec{Z} + \vec{\mu}$ 服从 m 维 Gauss 分布, 记 $\vec{X} \sim N(\text{E}\vec{X}, \text{Cov}[\vec{X}]) = N(\vec{\mu}, \Sigma)$

$$\text{E}\vec{X} = \vec{\mu}, \text{Var}[X_i] = \sum_k a_{ik}^2$$

$$\begin{aligned} \text{Cov}[Z_k, Z_l] = \delta_{kl} \Rightarrow \text{Cov}[X_i, X_j] &= \sum_{k,l} a_{ik} a_{jl} \text{Cov}[Z_k, Z_l] = \sum_k a_{ik} a_{jk} \\ \Rightarrow \Sigma = \text{Cov}[\vec{X}] &= AA^T \end{aligned}$$

\vec{X} 服从 m 维 正态分布 $\Leftrightarrow r(A) = r(AA^T) = r(\text{Cov}[\vec{X}]) = m \Leftrightarrow \text{Cov}[\vec{X}]$ 正定

$\varphi(\vec{t}) = e^{i\vec{t}^T \vec{\mu} - \frac{1}{2} \vec{t}^T \Sigma \vec{t}}$, 且对于每个非负实对称阵 Σ , 此函数都是 Gauss 分布.

- 由特征函数容易证明, Gauss 分布做线性变换后仍为 Gauss 分布.

$$p(\vec{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{|\Sigma|}} e^{-\frac{(\vec{x} - \vec{\mu})^T \Sigma^{-1} (\vec{x} - \vec{\mu})}{2}}$$

Matrix Tricks:

设 n 维随机向量 $\vec{X} \sim N(0, I)$, 取正交阵 A , 使其首行为 $(\frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}})$, 则显然 $\vec{Y} = A\vec{X} \sim N(0, I)$

$$\begin{aligned}\bar{X} &= \frac{\sum X_i}{n} = \frac{Y_1}{\sqrt{n}} \\ (n-1)S^2 &= \sum (X_i - \bar{X})^2 = \sum X_i^2 - n\bar{X}^2 = X^T X - Y_1^2 \\ &= Y^T A A^{-1} Y - Y_1^2 = Y_2^2 + \dots + Y_n^2 \sim \chi^2(n-1) \\ &\Rightarrow \bar{X} \text{ 与 } (n-1)S^2 \text{ 独立}\end{aligned}$$

设 n 个独立同分布正态样本 $X_i \sim N(\mu, \sigma^2)$, 同上设 A , 并特别地, 构造 A 的第 $k \geq 2$ 行为 $(\frac{-1}{\sqrt{(k-1)k}}, \frac{-1}{\sqrt{(k-1)k}}, \dots, \frac{k}{\sqrt{(k-1)k}}, 0, \dots, 0)$, 使得行和为 0.
则有 $EY_1 = \sqrt{n}\mu, EY_k = 0 (k \geq 2), \text{Var}[Y_k] = \sigma^2 (k \geq 1)$

4.7 Conditional Distribution

设 A 为事件, $F_{X|A}(x) = P(X \leq x|A) = \int_{-\infty}^x p_{X|A}(u) du$

当 Y 为随机变量时, 根据新的下标重载函数 F :

$$F_{X|Y}(x|y) = P(X \leq x|Y = y) = \int_{-\infty}^x \frac{p_{X,Y}(u, y)}{p_Y(y)} du = \int_{-\infty}^x p_{X|Y}(u|y) du$$

逆乘法公式:

若 $p_{X,Y}(x, y) = g(y)h(x, y)$, 且 $\forall y, \int_R h(x, y) dx = 1$

则 $g(y) = p_Y(y), h(x, y) = p_{X|Y}(x|y)$

连续全概率公式: $p_X(x) = \int_{\mathbb{R}} p_Y(y) p_{X|Y}(x|y) dy$

连续 Bayes 公式: $p_{X|Y}(x|y) = \frac{p_X(x) p_{Y|X}(y|x)}{\int_{\mathbb{R}} p_X(x) p_{Y|X}(y|x) dx}$

条件期望:

$E[X|Y = y] = \int_{\mathbb{R}} x p_{X|Y}(x|y) dx$, 其值是关于 Y 的随机变量, 记做 $E[X|Y]$

$E[g(X)h(Y)|Y = y] = E[g(X)h(y)|Y = y] = h(y)E[g(X)|Y = y]$

$\Rightarrow E[g(X)h(Y)|Y] = h(Y)E[g(X)|Y]$

重期望公式:

$$\begin{aligned}EX &= \int_{\mathbb{R}} \int_{\mathbb{R}} x p(x, y) dx dy = \int_{\mathbb{R}} \int_{\mathbb{R}} x p(x|y) p_Y(y) dx dy \\ &= \int_{\mathbb{R}} E[X|Y = y] p_Y(y) dy = E[E[X|Y]] \\ EX &= \sum_j E[X|Y = y_j] P(Y = y_j) \text{ 用于很多趣题.}\end{aligned}$$

推论: 随机个随机变量和. 设 $X_1 \dots$ i.i.d, 则有 $E\left[\sum_{i=1}^N X_i\right] = EXEN$

条件期望是最佳平方逼近: $EX^2 < \infty$ 时, $E[X - E[X|Y]]^2 = \min E[X - \varphi(Y)]^2$

Proof: 由 $E[E[X|Y]]^2 \leq E[E[X^2|Y]] = EX^2$ 知存在性.

考虑期望在线性空间中的点积意义, 需证 $E[(X - E[X|Y])\varphi(Y)] = 0$.

$$\text{LHS} = E[E[(X - E[X|Y])\varphi(Y)|Y]] = E[\varphi(Y)E[X - E[X|Y]|Y]] = E[\varphi(Y)(E[X|Y] - E[E[X|Y]|Y])] = 0$$

5 Analytic Topics

5.1 Estimation

1. Chebyshev

$$\begin{aligned} \forall \varepsilon > 0, P(|X - EX| \geq \varepsilon) \\ &= \int_{|x - EX| \geq \varepsilon} p(x) dx \leq \int_{|x - EX| \geq \varepsilon} \frac{(x - EX)^2}{\varepsilon^2} p(x) dx \\ &\leq \frac{\int_{\mathbb{R}} (x - \mu)^2 p(x) dx}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2} \end{aligned}$$

统计意义: 与均值的距离远近对概率的限定.

2. Kolmogorov

设 X_1, \dots, X_n 相互独立, 期望方差为 μ_k, σ_k^2 , $S_k = \sum_{i=1}^k S_i$

$$\forall t > 0, P\left(\max_{k \in [1, n]} \frac{|S_k - ES_k|}{\sqrt{\text{Var}[S_n]}} < t\right) \geq 1 - t^{-2}$$

$n = 1$ 的情形即 Chebyshev.

5.2 Convergence

依概率收敛: $\forall \varepsilon > 0, P(|X_n - X| \geq \varepsilon) \rightarrow 0 (n \rightarrow \infty), \Leftrightarrow X_n \xrightarrow{P} X$

弱收敛: $\forall F$ 的连续点 x , $\lim_{n \rightarrow \infty} F_n(x) = F(x) \Leftrightarrow F_n(x) \xrightarrow{W} F(x)$

依分布收敛: $X_n \xrightarrow{L} X$

5.3 Law of Large Numbers

1. 弱大数定律: $\frac{\sum X_i}{n} \xrightarrow{P} \mu$

意义: 用平均值作为期望是合理的, 即使不知道其分布

Chebyshev: X_n 两两不相关, 方差一致有界

Markov: X_n 两两不相关, $\lim_{n \rightarrow \infty} \frac{1}{n^2} \text{Var} \left[\sum X_i \right] = 0$ (Markov 条件)

Bernstein: 只需 X_n 渐进不相关 ($\lim_{|k-l| \rightarrow \infty} \text{Cov}[X_k, X_l] = 0$), 方差一致有界

Khinchine: 只需 X_n i.i.d., 期望存在

推论: **J. Bernoulli:** n 次试验中 A 发生的次数 S_n , $\frac{S_n}{n} \xrightarrow{P} p$

意义: 概率是频率的极限

2. 强大数定律: $P(\lim_{n \rightarrow \infty} \frac{\sum X_i}{n} = \frac{\sum EX_i}{n}) = 1$

Borel: X_n i.i.d., $EX^4 < +\infty$

Kolmogorov: X_n 独立, $\sum \frac{\text{Var}[X_i]}{i^2} < +\infty$

5.4 Central Limit Theorem

1. **Lindeberg-Levy:** X_n i.i.d., $Y_n = \frac{\sum X_i - n\mu}{\sigma\sqrt{n}} \xrightarrow{L} N(0, 1)$

2. **De Moivre-Laplace:** n 次实验中 A 发生了 S_n 次, $Y_n = \frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{L} N(0, 1)$

3. **Lindeberg/Lyapunov:** X_n 独立, 设 $B_n = \sqrt{\sum \sigma_i^2}$, 若满足 Lindeberg 条件:

$$\forall t > 0, \lim_{n \rightarrow \infty} \frac{1}{B_n^2} \sum_{i=1}^n \int_{|x-\mu_i| > tB_n} (x - \mu_i)^2 p_i(x) dx = 0$$

或 Lyapunov 条件(弱于 Lindeberg):

$$\exists t > 0, \lim_{n \rightarrow \infty} (\sum_{i=1}^n \text{Var}[X_i])^{-t} \sum_{i=1}^n E[|X_i - \mu_i|^{2+t}] = 0$$

则 $\frac{1}{B_n} \sum (X_i - \mu_i) \xrightarrow{L} N(0, 1)$.

6 Random Walk

6.1 Combinatorial Perspectives

Counting Paths: 从 $(0, 0)$ 走到 (n, x) , $(|x| \leq n, n+x \equiv 0 \pmod{2})$ 的折线条数:

设有 p 个 $+1$, q 个 -1 , $\begin{cases} n = p+q \\ x = p-q \end{cases}$, 条数为 $N_{n,x} = C_{p+q}^p = C_n^{\frac{n+x}{2}}$

Reflection Principle:

设 $b_0, b_n > 0$, 从 $(0, b_0)$ 到 (n, b_n) 的经过 x 轴的折线与 $(0, -b_0)$ 到 (n, b_n) 的折线一一对应.

Ballot Theorem: 甲 $p > q$ 票, 甲始终领先:

等价于 $(1, 1)$ 到 $(p+q, p-q)$ 的不过 x 轴的折线, 有 $N_{n-1, x-1} - N_{n-1, x+1} = \frac{x}{n} N_{n, x}$

下面只考虑从原点出发的折线, 采用如下记号: $X_k = \pm 1, S_k = \sum_{i=1}^k X_i, S_0 = 0$

Reach: $p_{n,r} = \Pr\{S_n = r\} = N_{n,r} 2^{-n} = C_n^{\frac{n+r}{2}} 2^{-n}$

Return: $u_{2n} = p_{2n,0} = \frac{(2n)!}{(n!)^2} 2^{-2n} \approx \frac{1}{\sqrt{\pi n}}$ (Stirling's Approximation)

Important Lemma:

$$\Pr\{S_1, \dots, S_{2n} \neq 0\} = \Pr\{S_{2n} = 0\} = 2\Pr\{S_1, \dots, S_{2n} > 0\} = \Pr\{S_1, \dots, S_{2n} \geq 0\}$$

$$\begin{aligned} \text{Proof: } \Pr\{S_1, \dots, S_{2n} > 0\} &= \sum_{r=1}^n \Pr\{S_1, \dots, S_{2n-1} > 0, S_{2n} = 2r\} \\ &= \sum_{r=1}^n \frac{N_{2n-1, 2r-1} - N_{2n-1, 2r+1}}{2^{2n}} \quad (\text{by ballot theorem}) \\ &= \sum_{r=1}^n \frac{1}{2} (p_{2n-1, 2r-1} - p_{2n-1, 2r+1}) \\ &= \frac{1}{2} p_{2n-1, 1} - 0 = \frac{1}{2} u_{2n} \end{aligned}$$

又由 S 的整数连续性知 $\Pr\{S_1, \dots, S_{2n} \neq 0\} = 2\Pr\{S_1, \dots, S_{2n} \geq 0\}$

考虑每一条恒正路径的概率, 第一步走到 $(1, 1)$ 的概率为 $\frac{1}{2}$, 之后对应于一条长为 $2n-1$ 的恒非负路径, 即 $\Pr\{S_1, \dots, S_{2n} > 0\} = \frac{1}{2} \Pr\{S_1, \dots, S_{2n-1} \geq 0\} = \frac{1}{2} \Pr\{S_1, \dots, S_{2n} \geq 0\}$

最后一步是因为 $S_{2n-1} \geq 0$ 蕴含 $S_{2n} \geq 0$. \square

几何证明: 考虑到 $S_{2n} = 0$ 的每一条路径, 设其所有最小值点中最左边的一个为 $M(k, m)$, 将 $(0, 0) \sim M$ 的部分沿 $x = k$ 翻转后续接到 $(2n, 0)$ 后, 以 M 为新原点, 便得到一条恒非负路径. \square

First Return: $f_{2n} = \Pr\{S_1, \dots, S_{2n-1} \neq 0, S_{2n} = 0\}$

由引理立刻得, $f_{2n} = u_{2n-2} - u_{2n} = \frac{u_{2n}}{2n-1}$

有递推关系: $u_{2n} = \sum_{i=1}^n f_{2i} u_{2n-2i}$

Last Return: 长为 $2n$ 的路径, 最后一次经过 x 轴在 $2k$ 处的概率为 $a_{2n, 2k} = u_{2k} u_{2n-2k}$.

Proof: 由引理, 前 $2k$ 次结束于 x 轴的概率为 u_{2k} , 后 $2n-2k$ 次不经过 x 轴的概率为 u_{2n-2k} .

ArcSin Distribution: 由 u_{2k} 的逼近, 有 $a_{2n, 2k} \approx \frac{f(\frac{k}{n})}{n}$, $f(x) = \frac{1}{\pi \sqrt{x(1-x)}}$

积分: $\sum_{k < xn} a_{2n, 2k} \approx \frac{2}{\pi} \arcsin \sqrt{x}$

Sojourn Time: 有 $2k$ 时间在一象限, $2n-2k$ 时间在四象限的概率为 $b_{2n, 2k} = a_{2n, 2k}$

Proof: 设 First Return 发生在 $2r$ 处, 此前恒正有 $\frac{2^{2r} f_{2r}}{2} \times 2^{2n-2r} b_{2n-2r, 2k-2r}$ 种可能, 恒负有 $\frac{2^{2r} f_{2r}}{2} \times 2^{2n-2r} b_{2n-2r, 2k}$ 种可能.

$$\Rightarrow b_{2n, 2k} = \frac{f_{2r}}{2} \left[\sum_{r=1}^k b_{2n-2r, 2k-2r} + \sum_{r=1}^{n-k} b_{2n-2r, 2k} \right], (1 \leq k \leq n-1)$$

再由 $b_{2n, 0} = b_{2n, 2n} = u_{2n}$, 归纳即证.

Changes of Sign: 显然 $S_{2n+1} \neq 0$, 在前 $2n+1$ 次中符号变化的次数为 r 的概率为 $\xi_{2n+1, r}$

对 $r > 1$, 只考虑 $X_1 = 1$ 的情形, 将路径在变号点分两段计数归纳, 可证 $\xi_{2n+1, r} = 2p_{2n+1, 2r+1}$.

领先不易改变: $\xi_{n, 0} > \xi_{n, 1} > \dots$

Maximum: 由反射原理, $(0, 0) \sim (n, k)$ 的路径, 最大值为 r 的概率为 $p_{n, 2r-k} - p_{n, 2r+2-k}$

上式对 k 求和,得长为 n 的路径,最大值为 r 的概率为 $p_{n,r}$ 或 $p_{n,r+1}$ (只有一者有意义).
取 $r = 0$ 即是引理的证明.

First Passage: n 时刻初过 r 的概率为 $\varphi_{n,r}$

$(0, 0) \sim (n-1, r-1)$ 的最大值为 $r-1$ 的路径概率为 $p_{n-1,r-1} - p_{n-1,r+1}$
 $\Rightarrow \varphi_{n,r} = \frac{p_{n-1,r-1} - p_{n-1,r+1}}{2} = \frac{rp_{n,r}}{n+1}$

rth Return: 第 r 次返回发生在 n 时刻的概率为 $\varphi_{n-r,r}$

Proof: 设第 r 次返回发生在时刻 n 的长为 n 的路径有 A 条,则其中全在 x 轴下方的路径有 $\frac{A}{2^r}$ 条.
对每一条这种路径,将它所有从 x 轴出发的那 r 步删去,对应到一条 $n-r$ 时刻初过 r 的路径.

Sojourn Time with Return: 设 $S_{2n} = 0$, 前 $2n$ 时间中, $2k$ 时间在上方的概率与 k 无关.

Proof: 设 $2r$ 时刻 First Return, 讨论 $1 \sim 2r$ 恒正或恒负, 和式变换消去 k . 概率为 $\frac{u_{2n}}{n+1} = 2f_{2n+2}$

Duality: 设 $X_i^* = X_{n-i}, S_i^* = S_n - S_{n-i}$, 可得到一系列对偶命题. 起点-终点, 初-末对应.

Asymptotic:

访问次数 V_n , 部分和 S_n 的阶为 \sqrt{n} .

$P(S_n < t\sqrt{n}) \rightarrow \Phi(t), P(V_n < t\sqrt{n}) \rightarrow 2\Phi(2t) - 1$

初过 r 时间与第 r 次返回时间的阶为 r^2 .

初过 r 发生在 tr^2 之前的概率 $\rightarrow 2 - 2\Phi(\frac{1}{\sqrt{t}})$

第 r 次返回发生在 tr^2 之前的概率 $\rightarrow 2 - 2\Phi(\frac{1}{\sqrt{t}})$. (反直觉)

直觉认为返回后“无记忆”, 因此累积时间应当与 r 同阶.

6.2 Generating Function Perspectives

以下的 Random Walk 中, $+1, -1$ 分别具有概率 p, q .

Waiting Time for a Gain: 设 $\varphi_n = \varphi_{n,1}$ 为初过 1 的概率, $\Phi(x) = \sum_{n=0}^{\infty} \varphi_n x^n$

由定义有 $\varphi_0 = 0, \varphi_1 = p$.

对 $n > 1$, 必须以 q 的概率走到 $S_1 = -1$, φ_{n-1} 的概率走到 First Return 点 $S_v = 0, \varphi_{n-v}$ 概率走到 $S_n = 1$. 于是, $\varphi_n = q \sum_{k=1}^{n-2} \varphi_k \varphi_{n-1-k}$

注意到右边的 Convolution 形式, 可得 $\Phi(x) - px = qx\Phi^2(x) \Rightarrow \Phi(x) = \frac{1 - \sqrt{1 - 4pqx^2}}{2qx}$

解法二: $\Phi(x) = E[x^n] = pE[x^n | X_1 = 1] + qE[x^n | X_1 = -1]$.

$X_1 = 1 \Rightarrow n = 1 \Rightarrow E[x^n | X_1 = 1] = x$

$X_1 = -1 \Rightarrow n = 1 + n_1 + n_2$ (此处仍按照第一步和初返点来分段)

$\Rightarrow E[x^n | X_1 = -1] = E[x^{1+n_1+n_2}] = x\Phi^2(x)$. 代入即得 $\Phi(x)$

总会赢的概率: $\sum \varphi_n = \min\{\frac{p}{q}, 1\}$

赢的期望时间: $\Phi'(1) = \frac{1}{p-q}$. 公平赌博期望无穷次才能获利.

Return: $u_{2n} = C_{2n}^n p^n q^n, U(x) = \sum_{n=0}^{\infty} u_{2n} x^{2n} = \frac{1}{\sqrt{1-4pqx^2}}$

First Return: 设 $f_n = f_n^- + f_n^+, X_1 = -1$ 时的 $f_{2n}^- = q\varphi_{2n-1}$

$$\Rightarrow F^-(x) = \sum_{n=1}^{\infty} f_{2n}^- x^{2n} = qx\Phi(x) = \frac{1 - \sqrt{1-4pqx^2}}{2}$$

将 p, q 互换得到 $F^+(x) = F^-(x) \Rightarrow F(x) = \sum_{n=1}^{\infty} f_n x^n = 1 - \sqrt{1-4pqx^2}$

总会返回的概率: $F(1) = 1 - |p - q|$; 永不返回的概率: $|p - q|$
 $p = \frac{1}{2}$ 时一定会返回, 期望时间为 $F'(1) = \infty$.

6.3 Ruin Problem

考虑一次赌博, 若总资产为 a , 对于从 z 出发的赌徒, 走到 a 或 0 (破产) 代表赌博结束. 也即在 0 和 a 处有 absorbing barrier 的 random walk.

破产概率:

设从 z 出发在 0 和 a 处被吸收的概率分别为 q_z, p_z , 对于 q_z , 有如下差分方程:

$$q_z = pq_{z+1} + qq_{z-1}, 1 \leq z \leq a-1; q_0 = 1, q_a = 0$$

$p \neq q$ 时, 通解为 $q_z = A + B(\frac{q}{p})^z$; $p = q = \frac{1}{2}$ 时, 通解为 $q_z = A + Bz$. 代入边界条件, 可得

$$\text{破产概率 } q_z = \begin{cases} \frac{(q/p)^a - (q/p)^z}{(q/p)^a - 1}, & q \neq p \\ 1 - \frac{z}{a}, & q = p \end{cases} \rightarrow \begin{cases} 1, & p \leq q \\ (q/p)^z, & p > q \end{cases} \quad (a \rightarrow \infty)$$

将 $p, q; a - z, z$ 交换即得 p_z 的表达式, 可见 $p_z + q_z = 1$ (概率 1 会终止).

持续时间:

假设持续时间的期望存在有限, 为 D_z , 则可得差分方程:

$$D_z = p(D_{z+1} + 1) + q(D_{z-1} + 1) = pD_{z+1} + qD_{z-1} + 1, D_0 = D_a = 0$$

$p \neq q$ 时, 有特解 $D_z = \frac{z}{q-p} \Rightarrow$ 通解 $D_z = \frac{z}{q-p} + A + B(\frac{q}{p})^z$

$p = q$ 时, 有特解 $D_z = -z^2 \Rightarrow$ 通解 $D_z = -z^2 + A + Bz$, 代入边界, 得

$$D_z = \begin{cases} \frac{z}{q-p} - \frac{a}{q-p} \frac{1 - (q/p)^z}{1 - (q/p)^a}, & p \neq q \\ z(a-z), & p = q \end{cases} \rightarrow \begin{cases} \infty, & p \geq q \\ \frac{z}{q-p}, & p < q \end{cases} \quad (a \rightarrow \infty)$$

生成函数:

设 $u_{z,n}$ 表示 n 时刻破产 (走到 0), $U_z(t) = \sum_{n=0}^{\infty} u_{z,n} t^n$.

有 $u_{0,n} = u_{a,n} = u_{z,0} = 0 (0 < z \leq a, n \geq 1), u_{0,0} = 1$ 及 $u_{z,n+1} = pu_{z+1,n} + qu_{z-1,n}$

$$\Leftrightarrow U_z(t) = ptU_{z+1}(t) + qtU_{z-1}(t), U_0(t) = 1, U_a(t) = 0$$

由特征根法可解出 $U_z(t) = A(t)\lambda_1^z(t) + B(t)\lambda_2^z(t), \lambda_i(t) = \frac{1 \pm \sqrt{1-4pqt^2}}{2pt}$

代入边界,得破产时刻的生成函数 $U_z(t) = (\frac{q}{p})^z \frac{\lambda_1^{a-z}(t) - \lambda_2^{a-z}(t)}{\lambda_1^a(t) - \lambda_2^a(t)}$

类似可得获胜时刻的生成函数 $V_z(t) = \frac{\lambda_1^z(t) - \lambda_2^z(t)}{\lambda_1^a(t) - \lambda_2^a(t)}$

$a \rightarrow \infty$ 时,注意 Vieta 定理可得 $U_z(t) = \lambda_2^z(t)$, λ 取小根.

半开随机徘徊的破产对应 First Passage. 上式表明初过 $-z$ 的等待时间是 z 个独立时间和.

$U_z(t)$ 的显式展开: $u_{z,n} = \frac{2^n}{a} p^{\frac{n-z}{2}} q^{\frac{n+z}{2}} \sum_{k=1}^{a-1} [\cos^{n-1} \frac{k\pi}{a} \sin \frac{k\pi}{a} \sin \frac{kz\pi}{a}]$

6.4 Higher Dimensions

Return:在一维/二维 Random Walk 中,质点将以概率 1 返回初始位置,而三维中概率约为 0.35.

一维时, $u_{2n} \approx \frac{1}{\sqrt{\pi n}} = O(n^{-\frac{1}{2}})$, $\sum u_{2n}$ 发散,于是 $\sum f_{2n} = 1$,必定会返回.

二维时, $u_{2n} = \frac{1}{4^{2n}} \sum_{k=0}^n \frac{(2n)!}{k!k!(n-k)!(n-k)!} = \frac{1}{4^{2n}} C_{2n}^n \sum_{k=0}^n (C_n^k)^2 = \frac{(C_{2n}^n)^2}{4^{2n}} = O(n^{-1})$.

三维时, $u_{2n} = \frac{1}{6^{2n}} \sum_{j+k \leq n} \frac{(2n)!}{j!j!k!k!(n-j-k)!(n-j-k)!} = \frac{1}{2^{2n}} C_{2n}^n \sum_{j+k \leq n} [\frac{n!}{3^n j!k!(n-j-k)!}]^2$

当 $j = k = \frac{n}{3}$ 时和式内取到最大项,只分析此项有 $u_{2n} = O(n^{-\frac{3}{2}})$, $\sum u_{2n}$ 收敛 $\Rightarrow \sum f_{2n} < 1$.

三维空间中可以以概率 1 到达任意一条与坐标轴平行的直线.(投影成二维 Random Walk)

期望距离:在 d 维空间中,设 n 时刻到达 $(x_{1,n} \cdots x_{d,n})$,考虑距原点距离平方的增量:

$$E[D_{n+1}^2 - D_n^2] = E\left[-2 \sum_d x_{d,n} X_d + \sum_d X_d^2\right] = E\left[\sum_d X_d^2\right] = 1 \Rightarrow E[D_n] = \sqrt{n}$$

7 Branching Process

设 X_k 表示第 k 代的个数, $X_0 = 1$. 每个个体以相同的概率分布独立产生一定数量的后代.

设 $G_n(t)$ 是 X_n 的 PGF, $G_1(t) = G(t) = \sum_{k=0}^{\infty} p_k t^k$.

则显然有 $G_n(t) = G(G_{n-1}(t))$

$$E[X_n] = G'_n(1) = G'_{n-1}(G(1))G'(1) = G'_{n-1}(1)G'(1) = \cdots = [G'(1)]^n \stackrel{def}{=} \mu^n$$

$$\text{Var}[X_n] = \begin{cases} n\sigma^2, \mu = 1 \\ \sigma^2 \mu^{n-1} \frac{\mu^n - 1}{\mu - 1}, \mu \neq 1 \end{cases}$$

7.1 Extinction Probability

在 n 代或以前灭绝的概率 $e_n = P(X_n = 0) = G_n(0) = G(e_{n-1})$

由于 $G(t)$ 在 $[0, 1]$ 单调增, $e_2 = G(e_1) > G(0) = p_0 = e_1$, 归纳可得 e_n 单调增, 设其极限为 e .

显然 e 是 $G(t)$ 在 $[0, 1]$ 上的不动点, 且归纳易知其最小不动点.

由于 $G(t)$ 过 $(0, p_0)$ 和 $(1, 1)$ 且递增下凸, 除 1 外可能还有一个或零个不动点.

当 $\mu = G'(1) \leq 1$ 时, $G(t)$ 的唯一不动点为 1, 即 $e = 1$, 过程必然灭绝.

当 $\mu > 1$ 时, 灭绝概率趋于 $G(t)$ 的较小不动点.

7.2 Total Progeny

设 $Y_n = \sum_{k=0}^n X_k$, 其 PGF 为 $R_n(t)$, $R_1(t) = tG(t)$

类似地, 考虑所有人分别是 X_1 中哪些人的后裔, 可得 $R_n(t) = tG(R_{n-1}(t))$

$\forall t < 1, R_2(t) = tG(tG(t)) < tG(t) = R_1(t)$, 归纳可得 $R_n(t) < R_{n-1}(t)$, 所以 $R_n(t) \rightarrow \rho(t)$

那么无穷代后的总后代数的 PGF $\rho(t)$ 是 $x = tG(x)$ 的根.

由凸性可知根至多有 2 个. 考查 $x = 0, x = e, x = 1$ 时的情形知根只有一个, 在 $(0, e]$ 之间.

$t = 1$ 时显然 $\rho(1) = e$. ($e \neq 1$ 时 $\rho(t)$ 不是一个 PGF)

$$E[Y_n] \rightarrow \begin{cases} \infty, \mu \geq 1 \\ \frac{1}{1-\mu}, \mu < 1 \end{cases}$$

7.3 Geometric Distribution

设个体繁衍的概率分布为几何分布, 生成函数 $G(t) = \frac{q}{1-pt}$

可得 $G_n(t) = q \frac{p^n - q^n - (p^{n-1} - q^{n-1})pt}{p^{n+1} - q^{n+1} - (p^n - q^n)pt}$

灭绝概率 $e = \min\{\frac{q}{p}, 1\}$

总后代数分布 $\rho(t) = \frac{1 - \sqrt{1 - 4pqt}}{2p}$

7.4 Busy Periods

顾客在整数时间独立到来接受单线程服务, 时长为整数符合 PGF: $\beta(t)$, 相互独立.

将每个顾客服务时间内到来的顾客都视为其后代, 则后代个数为随机和 $X_1 + \dots + X_N$, X_i 表示 i 时刻有没有顾客, 为两点分布.

\Rightarrow 个体繁衍的 PGF: $G(t) = \beta(pt + q)$. 设 $\sigma = \beta'(1)$ 为平均时长.

根据灭绝判据, $\mu = G'(1) = p\sigma \leq 1$ 时, 忙期一定会结束. $\mu < 1$ 时, 忙期内到达的顾客数期望有限.

Uninterrupted Period:

以时刻作为 Branching Process 的元素, 将每个顾客服务时间内的所有时刻视做服务开始时刻的后代.

对于一个时刻, 其繁衍 PGF 为 $G(t) = q + p\beta(t)$. 其总后代数为相应的 $\rho(t)$, 这里包含了开始时刻无顾客的情形.

Uninterrupted Period 要求开始时刻来了顾客, 且每个顾客独立繁衍的总后代数即其总时间.

其 PGF 为 $\beta(\rho(t))$

8 Recurrent Event

8.1 Definition

一个(未必独立的)重复实验序列, 实验结果可能为 $E_j (j = 1, \dots)$

设 \mathcal{E} 是有限结果序列的一个可判定属性.

语句“ \mathcal{E} 在 E_{k_1}, \dots 的第 n 个位置出现”的意义是 E_{k_1}, \dots, E_{k_n} 具有属性 \mathcal{E}

若属性满足以下两个条件, 则称其定义了一个 Recurrent Event:

1. 在结尾处判定出现

\mathcal{E} 在 $E_{k_1}, \dots, E_{k_{n+m}}$ 的第 n 和第 $n+m$ 个位置出现, 等价于 \mathcal{E} 在 E_{k_1}, \dots, E_{k_n} 和 $E_{k_{n+1}}, \dots, E_{k_{n+m}}$ 的最后出现

2. 出现后立刻丧失记忆

若 \mathcal{E} 在序列第 n 位出现, 则恒有 $P(E_{k_1}, \dots, E_{k_{n+m}}) = P(E_{k_1}, \dots, E_{k_n})P(E_{k_{n+1}}, \dots, E_{k_{n+m}})$

例子:

Random Walk 里的 Return, 从负方向 Return, 更新最大值;

排队论里的达到空闲.

考虑连续 Bernoulli 实验中的连续成功次数, 若规定不计算重叠, 则连续 r 次成功也是 Recurrent Event.

仿照 6.1 节的记号, 设 \mathcal{E} 在 n 处出现的概率为 u_n , 在 n 处首次出现的概率为 f_n . 设 $u_0 = 1, f_0 = 0$. 并设其 PGF 分别为 $U(t), F(t)$. 需要注意, u_k 不是一个概率分布, f_k 仅当事件一定会发生 ($f = \sum_{n=0}^{\infty} f_n = 1$) 时是概率分布.

设随机变量 T 表示等待时间, $P(T = n) = f_n, P(T = \infty) = 1 - f$.

设第 r 次出现位置的分布是 $f_k^{(r)}$ 则它应是 $f_k^{(r-1)}$ 与 f_k 的卷积. 于是其 PGF 满足 $F^{(r)}(t) = F^r(t)$ 至少出现 r 次的概率为 $F^r(1) = f^r$. 当 $f = 1$ 时, 称 \mathcal{E} 常返 (persistent), 否则称暂留 (transient).

若使得 $u_n \neq 0$ 的所有 n 有公因子 $\lambda > 1$, 则称 \mathcal{E} 是周期的.

由于 $u_n = f_1 u_{n-1} + \dots + f_n u_0 (n \geq 1) \Rightarrow U(s) - u_0 = F(s)U(s) \Rightarrow U(s) = \frac{1}{1 - F(s)}$

\mathcal{E} 暂留等价于 $\frac{1}{1 - f} = u = \sum_{n=0}^{\infty} u_n$ 收敛.

8.2 A Theorem of Particular Importance

设 \mathcal{E} 常返, 非周期. $\mu = E[T] = F'(1)$, 则 $u_n \rightarrow \mu^{-1}$. 特别地, $\mu = \infty$ 时, $u_n \rightarrow 0$.

Proof: (P. Erdos & W. Feller, 1949)

设 $\lambda = \sup u_n$, u_{n_v} 是趋于 λ 的子列.

首先证明, $\forall j$, s.t. $f_j > 0, u_{n_v - j} \rightarrow \lambda$.

Proof: 反设 $u_{n_v - j} \rightarrow \lambda' < \lambda$, 对足够大的 v :

$$\begin{aligned} \lambda - \varepsilon &< u_{n_v} = \sum_{i=0}^{n_v} f_i u_{n_v - i} \\ &< \left(\sum_{\substack{i=0 \\ i \neq j}}^{n_v} f_i \right) (\lambda + \varepsilon) + f_j (\lambda' + \varepsilon) \\ &< (1 - f_j) (\lambda + \varepsilon) + f_j (\lambda' + \varepsilon) = \lambda - f_j (\lambda - \lambda') + \varepsilon \Rightarrow \lambda' = \lambda \quad \square \end{aligned}$$

重复此操作, 可得命题: $u_{n_v} \rightarrow \lambda, f_j > 0 \Rightarrow u_{n_v - k_j} \rightarrow \lambda$

由 \mathcal{E} 的非周期性, 所有使得 $f_j > 0$ 的 j 无公因子.

$\Rightarrow \exists$ 足够大的 N 和 a_1, \dots, a_l , s.t. $f_{a_i} > 0$, 且 $\forall k > N, k$ 可表为 $\sum_{i=1}^l x_i a_i$

设 $N^* = n_v - N$, 则 $\forall M, u_{N^* - M} = u_{n_v - (N + M)} \rightarrow \lambda$

再令 $r_n = \sum_{k=n+1}^{\infty} f_k$, $R(t) = \sum_{k=0}^{\infty} r_k t^k$, 由和式变换可知 $R(t)U(t) = (1-t)^{-1}$

即 $\sum_{k=0}^{N^*} r_k u_{N^*-k} = 1. \Rightarrow \forall M, \sum_{k=0}^M r_k u_{N^*-k} \leq 1$

令式中 $v \rightarrow \infty$, 得 $\forall M, \sum_{k=0}^M r_k \leq \frac{1}{\lambda}$

由期望的第二种求法有 $\mu = F'(1) = R(1)$. 于是 $\lambda \leq \frac{1}{\mu}$.

重考虑 $\lambda = \inf u_n$ 的情形, 类似以上推理可证 $\lambda \geq \frac{1}{\mu}$. (需注意第一步放缩时利用常返性).

于是 $u_n \rightarrow \mu$. □

推广: \mathcal{E} 常返, 具有周期 λ 时, 有 $u_{n\lambda} \rightarrow \frac{\lambda}{\mu}$

8.3 Number of Occurrences

设 N_n 为前 n 次中 \mathcal{E} 出现的次数. $N_n = \sum_{k=1}^n Y_k, P(Y_k = 1) = u_k \Rightarrow E[Y_k] = u_k$

$\therefore E[N_n] = \sum_{k=1}^n u_k$. 当等待时间的 μ, σ 有限时 $E[N_n] \rightarrow \frac{n}{\mu}, \text{Var}[N_n] \rightarrow \frac{n\sigma^2}{\mu^3}$

当 $\mu = \infty$ 时, $E[N_n]$ 可能不与 n 同阶(反直觉).

如公平赌博中的 Return, $u_{2n} \rightarrow \frac{1}{\pi n} \Rightarrow E[N_{2n}] \rightarrow 2\sqrt{\frac{n}{\pi}}$

8.4 Runs of Specific Patterns

Success Runs:

设 \mathcal{E} 为 Bernoulli 实验中长为 r 的连续成功(不计算重叠), 则 \mathcal{E} 是 Recurrent Event.

依照定义, 任意连续的 r 次成功的实验 $n-r+1, \dots, n$ 中必定有一次出现了 \mathcal{E}

$\Rightarrow p^r = u_n + u_{n-1}p + \dots + u_{n-r+1}p^{r-1}, n \geq r$.

以及 $u_1 = \dots = u_{r-1} = 0, u_0 = 1$.

将其写为 GF 形式: $p^r(t^r + t^{r+1} \dots) = (U(t) - 1)(1 + pt + p^2t^2 + \dots + p^{r-1}t^{r-1})$

$\Leftrightarrow U(t) = \frac{1-t+qp^rt^{r+1}}{(1-t)(1-p^rt^r)} \Rightarrow F(t) = \frac{p^rt^r(1-pt)}{1-t+qp^rt^{r+1}}$

平均等待时间 $\mu = F'(1) = \frac{1-p^r}{qp^r}$

Pattern: SSFFSS.

类似上面的递推, 这里有 $p^4q^2 = u_n + p^2q^2u_{n-4} + p^3q^2u_{n-5}$

若欲求期望等待时间, 令 $n \rightarrow \infty$ 即可.

Runs of Either Kind: 设 \mathcal{E} 为出现一个长为 a 的连续成功或长为 b 的连续失败.

设 $\mathcal{E}_1, \mathcal{E}_2$ 为连续 a 次成功, 连续 b 次失败这两个循环事件.

除 0 时刻外其余的发生概率都应直接相加, 所以 $U(t) = U_1(t) + U_2(t) - 1$

连续 a 次成功发生在连续 b 次失败之前的概率 x :

解法 1: 设 u, v 分别为“第一次成功/失败”条件下此事件的概率, 列出二元线性方程组解之, $x = pu + qv$.

解法 2: 设 $\mathcal{E}_1, \mathcal{E}_2$ 对应的等待时间 PGF 为 $F(t), G(t)$, x_n 为 \mathcal{E}_1 初次发生在 n 且 \mathcal{E}_2 未发生的概率, 其 PGF 为 $X(t)$, 类似定义 y_n , 则有

$$\begin{cases} x_n = f_n - \left(\sum_{i=1}^{n-1} y_i f_{n-i} \right) \Rightarrow X(t) = F(t) - Y(t)F(t) \\ y_n = g_n - \left(\sum_{i=1}^{n-1} x_i g_{n-i} \right) \Rightarrow Y(t) = G(t) - X(t)G(t) \end{cases} \Rightarrow X(t) = \frac{F(t)[1 - G(t)]}{1 - F(t)G(t)}$$

利用 L' Hospital, $x = \lim_{t \rightarrow 1^-} X(t) = \frac{p^{a-1}(1 - q^b)}{p^{a-1} + q^{b-1} - p^{a-1}q^{b-1}}$