概率论笔记 ©

本书为应律的个人学习笔记，以杨振明老师的《概率论（第二版)》为蓝本，记录了相关重点知识与解题技巧。

:|

概率论筆記

田欣洋[1]

2021 年 9 月 12 日

目录

第一章　随机事件与概率

- 随机事件与概率
  - doc$\Phi_{\alpha\sigma}(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} dt$
  - fonts
  - source
  - tex
    - generic
    - latex
    - plain
  - texdoc

1.1　基本概念

1.1.1　随机现象与样本空间

随机试验　，把对自然现象的一次观察称为**试验**，其特点如下

试验：E

特点

1. 相同条件下可重复进行
2. 每次结果不可预知
3. 所有可能结果已知

易错例子：~~在纸上随意写三个数字~~（结果可知，但不可重复，）

样本点　随机试验的每一个可能结果称为一个样本点，记为 ω

样本空间　全体样本点构成的集合，记为 Ω

1.1.2　事件及其运算

随机事件　：代表试验的每一个可能结果（即 Ω 中的部分样本点组成的结果），是不可再分的事件（基本事件），如若某个事件每次试验都发生则称之为**必然事件**记作 Ω，同理每次都不发生的事件称为**不可能事件**，记作 $\varnothing$。

事件的关系　：事件间的关系就是集合之间的关系。一些基本关系如下：

1. 并：$A\cup B; A+B$,　$\bigcup\limits_{i=1}^{\infty} A_i$ 表示 $A_1, A_2, \cdots$ 至少有一个发生。
2. 交：$A\cap B; A\times B$,　$\bigcap\limits_{i=1}^{\infty} A_i$ 表示 $A_1, A_2, \cdots$ 全部都发生。
3. 逆：$\bar{A}$ (A 不发生),　$A\cdot\bar{A}=\varnothing, A+\bar{A}=\Omega$
4. 差：$A-B$ (A 发生且 B 不发生)
5. 包含 $A\in B$，相等 $A=B$
6. 互不相容事件：A 与 B 不同时发生
7. 对立事件：A 与 B 不同时发生，且 $A\cup B=\Omega$，显而易见前者更宽松，范围更大。

$\overline{ABC}\neq\bar{A}\bar{B}\bar{C}$

$1-\bar{A}\bar{B}\bar{C}=A+B+C$

充分条件前推后，必要条件后推前。

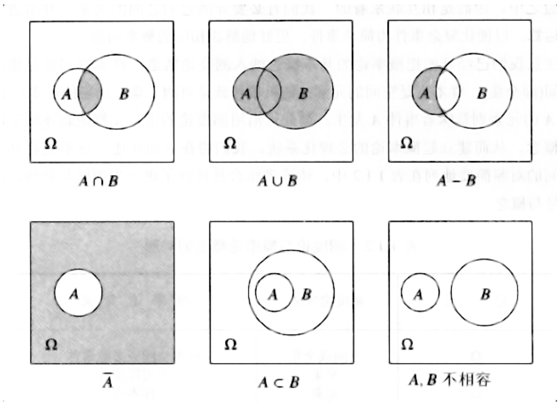


图 1.1: 事件运算 Veen 图

事件间的运算规律 设 A，B，C 为事件，则有以下规律:

1. 交换律: $A \cup B = B \cup A \qquad AB = BA$
2. 结合律: $(A \cup B) \cup C = A \cup (B \cup C) \qquad (AB)C = A(BC)$
3. 分配律: $A \cap (B - C) = AB - AC \quad A \cap (B \cup C) = AB \cup AC$
4. 对偶律: $\overline{A \cup B} = \bar{A} \cap \bar{B} \qquad \overline{A \cdot B} = \bar{A} + \bar{B}$

1.1.3 频率与概率

频率 在相同条件下进行了 n 次试验，若事件 A 在 n 次试验中发生了 m 次，则称事件 A 在这 n 此试验中发生的频率为

$$F_n(A) = \frac{m}{n}$$

性质

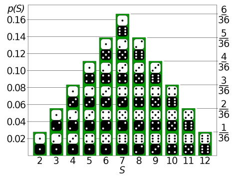
1. 非负性: $0 \leq F_n(A) \leq 1$
2. 规范性: $F_n(\Omega) = 1$
3. 有限可加性: $F_n(A \cup B) = F_n(A) + F_n(B)$
4. 稳定性: 随着 n 的增大，随机事件的频率趋于稳定

Cardano

概率 随机试验的次数充分大时，事件 A 发生的频率稳定在某数 P 附近摆动，则称 p 为事件 A 发生的统计概率，记为 P(A)=P

性质

1. 非负性：$0 \leq P(A) \leq 1$
2. 规范性：$P(\Omega) = 1$
3. 有限可加性：$P(A \cup B) = P(A) + P(B)$



掷骰子

1.2 古典概型

1.2.1 古典模型的定义

对于某一个随机试验，如果它的全体基本事件是**有限个**，而且具有**等可能性**，则称之为古典概型。则对于任意事件 A, 对应的概率（**古典概率**）为：

$$P(A) = \frac{\text{事件 A 包含的基本事件数 k}}{\text{基本事件总数 n}}$$

1.2.2 计数原理

基本计数原理

加法原理 设完成一件事有 m 种方式，每种方式分别有 $n_1, n_2, \cdots, n_m$ 种方法，则完成这件事总共有 $n_1 + n_2 + \cdots + n_m$ 种方法

乘法原理 设完成一件事有 m 个步骤，每个步骤分别有 $n_1, n_2, \cdots, n_m$ 种方法，则完成这件事总共有 $n_1 \times n_2 \times \cdots \times n_m$ 种方法

排列组合问题

排列 考虑顺序，从 n 个元素中取出 k 个元素的排列数量为：$A_n^k = \frac{n!}{(n-k)!}$

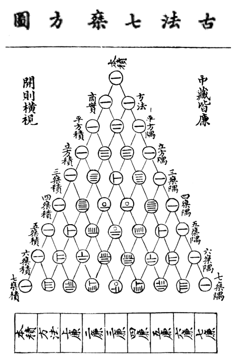
k=n 时，称为**全排列**

重复排列：从 n 个元素中有放回地取出 k 个元素（可重复）的排列数量为：$n \cdots n = n^k$

组合 不考虑顺序，从 n 个元素中取出 k 个元素的组合数量为：$C_n^k = \binom{n}{k} = \frac{P_n^k}{k!} = \frac{n!}{k!(n-k)!}$

即每个元素被多记了 $K!$ 次

组合数又称为二项式系数 $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$



杨辉三角

经典问题模型

1. 抽样问题
2. 分房问题
3. 配对问题
4. 随机取数

5. 结绳问题 将 n 个绳的两头两两相接求事件 $A=\{$恰结成 n 个圈$\}$ 的概率。

组合的性质

1. $\binom{n}{r}=\binom{n}{n-r}$
2. $\binom{n}{r}=\binom{n-1}{r-1}+\binom{n-1}{r}$ **解释：** 从 n 个元素中取 r 个，记特殊元素 a_1，即包含与不包含 a_1 两种情况，已知包含时：$\binom{n-1}{r-1}$，不包含时：$\binom{n-1}{r}$
3. $\binom{n}{0}+\binom{n}{1}+\cdots+\binom{n}{n}=2^n$ **解释：** 两系数为 1 的二项式
4. $\binom{n}{1}+2\binom{n}{2}+\cdots+n\binom{n}{n}=n2^{n-1}$
5. $\displaystyle\sum_{i=0}^{k}\binom{n}{i}\binom{m}{k-i}=\binom{n+m}{k}$
6. $\binom{n}{0}^2+\binom{n}{1}^2+\cdots+\binom{n}{n}^2=\binom{2n}{n}$

1.3 几何概型

1.3.1 定义

设 Ω 为 n 维欧氏空间中确定的集合，满足条件 $0<m(\Omega)<+\infty$。对 Ω 中的任何可测子集 A, 称 $P(A)=\frac{m(A)}{m(\Omega)}$ 为事件 A 的几何概率。

性质

1. 非负性：$P(A)\geq 0$
2. 规范性：$P(\Omega)=1$
3. 可列可加性：$P\Big(\bigcup\limits_{n=1}^{\infty}A_n\Big)=\sum\limits_{n=1}^{\infty}P(A_n)$

1. 蒲松投针
2. Bertrand 奇论

1.4 概率空间

1.4.1 事件 σ-代数

定义： 设样本空间为 Ω，$\mathscr{F}$ 是由 Ω 的一些子集构成的集合族，如果 $\mathscr{F}$ 满足以下条件：

(i) $\Omega\subset\mathscr{F}$;

(ii) 若 $A\in\mathscr{F}$，则 $\bar{A}\in\mathscr{F}$;

(iii) 若 $A_i \in \mathscr{F}, i=1,2,\cdots$,n，则 $\bigcup_{i=1}^{\infty} A_i \in \mathscr{F}$;

$\mathscr{F}$ 称为**体**，也叫**代数**

则称 $\mathscr{F}$ 为 Ω 上的事件 σ-代数，其性质如下：

性质

1. 若 $\mathscr{F},\mathscr{G}$ 都是 Ω 上的 σ-代数，则 $\mathscr{F}\cap\mathscr{G}$ 仍为 σ-代数
2. Ω 上任意多个 σ-代数的交仍为 σ-代数
3. (Borel 集) 设 $\Omega=\mathbf{R}$，取全体半直线组成的类 $H=(-\infty,x):-\infty<x<+\infty$，则 $\mathscr{B}=\sigma(H)$ 称为 Borel 集类。

花体艺术

$\mathscr{A}\mathscr{B}\mathscr{C}\mathscr{D}\mathscr{E}\mathscr{F}\mathscr{G}\mathscr{H}\mathscr{I}$

$\mathscr{J}\mathscr{K}\mathscr{L}\mathscr{M}\mathscr{N}\mathscr{O}\mathscr{P}\mathscr{Q}\mathscr{R}$

$\mathscr{S}\mathscr{T}\mathscr{U}\mathscr{V}\mathscr{W}\mathscr{X}\mathscr{Y}\mathscr{Z}$

1.4.2 概率与概率空间

概率 设 $\mathscr{F}$ 为样本空间 Ω 上的事件 σ-代数，若 $\mathscr{F}$ 上的实值函数 $P(A)$ 满足以下三个条件

1. 非负性：$P(A)\geq 0$
2. 规范性：$P(\Omega)=1$
3. 可列可加性：对于任一两两互斥事件 $A_1,A_2,\cdots,A_n$ 有 $P\left(\bigcup_{n=1}^{\infty}A_n\right)=\sum_{n=1}^{\infty}P(A_n)$

则称 P 为 $(\omega,\mathscr{F})$ 上的概率，$P(A)$ 称为事件 A 的概率。

概率空间 设 Ω 为一样本空间，设 $\mathscr{F}$ 为样本空间 Ω 上的事件 σ-代数，P 为 $\mathscr{F}$ 上的概率，称 $(\Omega,\mathscr{F},P)$ 为概率空间。

性质

1. $P(\varnothing)=0$
2. $P(\bar{A})=1-P(A)$
3. 若 $A\supset B$，则 $P(A-B)=P(A)-P(B)$，且 $P(B)\leq P(A)$
4. **次可加性** $P\Big(\bigcup\limits_{i=1}^{n}A_i\Big)\leq\sum\limits_{i=1}^{n}P(A_i)$
5. $P(\Omega)=1=P(A\cup\bar{A})=P(A)+P(\bar{A})$
6. **多除少补原理** $P\Big(\bigcup\limits_{i=1}^{n}A_i\Big)=\sum\limits_{i=1}^{n}P(A_i)+\cdots+(-1)^{n-1}P(A_1A_2\ldots A_n)$
7. **下连续性** 设 $A_1\subset A_2\subset\cdots$，且 $\bigcap\limits_{n=1}^{\infty}A_n=\varnothing$，则 $P\Big(\bigcup\limits_{n=1}^{\infty}A_n\Big)=\lim\limits_{n}P(A_n)$
8. **上连续性** 设 $A_1\supset A_2\supset\cdots$，且 $\bigcap\limits_{n=1}^{\infty}A_n=A$，则 $P\Big(\bigcap\limits_{n=1}^{\infty}A_n\Big)=\lim\limits_{n}P(A_n)$

1.5 条件概率

1.5.1 定义与乘法定理

设 $(\Omega,\mathscr{F},P)$ 为一概率空间，$B\in\mathscr{F},P(B)>0$, 则对于任何事件 $A\in\mathscr{F}$，记:

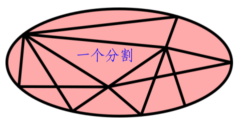
$$P(A|B)=\frac{P(AB)}{P(B)} \tag{1.1}$$

乘法定理 设 $(\Omega,\mathscr{F},P)$ 为一概率空间，$A_k\in\mathscr{F},k=1,2,\cdots n$，如果 $P(A_1\cdots A_{n-1})>0$，则有:

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1}) \tag{1.2}$$

1.5.2 全概率公式与贝叶斯公式

复杂简单化



对全集的一个分割

全概率公式 设 $(\Omega,\mathscr{F},P)$ 为一概率空间，而 $\{B_n\}$ 为 Ω 的一个分割，则对于一切 $A\in\mathscr{F}$ 有:

$$P(A)=\sum_{n}P(B_n)P(A|B_n) \tag{1.3}$$

条件全概率公式

$$P(B|D)=\sum_{i}P(A_i|D)P(B|A_i\cap D) \tag{1.4}$$

贝叶斯公式 设 $(\Omega, \mathscr{F}, P)$ 为一概率空间，而 $\{B_n\}$ 为 Ω 的一个分割，则对于 $\mathscr{F}$ 中一切有正概率的事件 A 及任意 n 有:

$$P(B_n|A) = \frac{P(B_n)P(A|B_n)}{\sum\limits_i P(B_i)P(A|B_i)}, \qquad i = 1, 2, \cdots, n, \tag{1.5}$$

1.6 事件的独立性

两个事件独立 设 $(\Omega, \mathscr{F}, P)$ 为一概率空间，若事件 A,B 满足：$P(A \cap B) = P(A)P(B)$ 则称**事件 A 与 B 相互独立。**

n 个事件独立 $\rightarrow$ n 个事件两两独立

n 个事件独立 设$(\Omega, \mathscr{F}, P)$概率空间中的 n 个事件若 $P(A_1 \cap A_2 \cap \cdots A_n) = P(A_1)P(A_2)\cdots P(A_n)$，则称 **n 个事件相互独立。**

第二章　随机变量

2.1　随机变量及其分布

2.1.1　定义与条件

定义　设 $(\Omega,\mathscr{F},P)$ 为一概率空间，$\xi=\xi(\omega)$ 是 Ω 上的实值函数，如果有：

$$\{\omega:\xi(\omega)<x\in\mathscr{F}\}\ \forall x\in\mathbb{R} \xleftrightarrow{\text{随机变量}} \xi^{-1}(B)\in\mathscr{F},\ \forall B\in\mathscr{F}$$

一般可以简写作 ξ

简单随机变量：示性函数的线性组合 $\Sigma_{i=1}^{n}a_i\mathbf{1}_{A_i}(\omega)$

则称 $\xi(\omega)$ 为随机变量，下面有几个定理：

1. **逆变换**

   1° $\xi^{-1}(\mathbb{R})=\Omega$

   2° 若 $B\subset C$，则 $\xi^{-1}(B)\subset\xi^{-1}(C)$

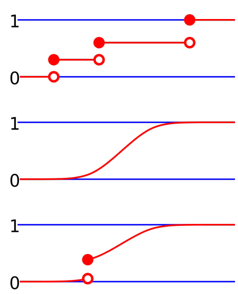
   3° $\xi^{-1}(\bar{B})=\overline{[\xi^{-1}(B)]}$

   4° $\xi^{-1}(\bigcup\limits_n B_n)=\bigcup\limits_n\xi^{-1}(B_n)$

2. **Borel 集**　　随机变量的 Borel 函数仍为随机变量。

3. **随机变量序列**　　ξ 为随机变量的充要条件是存在简单随机变量序列

$$\{\xi_n,n>1\}\ \text{使}\lim_n\xi_n(\omega)=\xi(\omega),\qquad\forall\omega\in\Omega$$



分布函数

2.1.2　分布与分布函数

$\xi^{-1}(B)$ 为随机事件，它的相应概率为：

$$\mathbf{F}(B)=P\{\xi^{-1}(B)\}=P\{\xi\in B\},\qquad\forall B\in\mathscr{B}$$

由上式定义的 $\mathscr{F}$ 成为随机变量的概率分布简称*分布或分布律*。

分布函数　**设 ξ 为概率空间 $(\Omega,\mathscr{F},P)$ 上的随机变量**，称：

$$F(x)=\mathbf{F}\{(-\infty,x)\}=P\{\xi<x\},\qquad x\in\mathbb{R}$$

为随机变量 ξ 的分布函数。

分布函数的性质

1. *单调连续性*　　任 $x_1<x_2$ 有 $F(x_1)<F_(x_2)$
2. *左连续性*　　任 $x_0\in\mathbb{R}$ 有 $\lim\limits_{x\to x_0-0}F(x)=F(x_0)$
3. *有界性*　　$F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0\quad F(+\infty)=\lim\limits_{x\to+\infty}F(x)=1$

离散型 & 连续型

1. 离散型：如果存在数列 $\{x_k\}$ 及 $\{p_k\}$，满足：$p_k \geq 0$; $\sum\limits_k p_k = 1$，使得

$$P\{\xi = x_k\} = p_k, \quad k = 1, 2, \cdots,$$

则称其为离散型, 由两个数列组成的矩阵称为 ξ 的密度阵

一个纯跳跃型函数

2. 连续型：如果存在函数 $p(x)$，满足 $p(x) \geq 0$; $\int_{-\infty}^{+\infty} p(x)dx = 1$，使得

$$\mathbf{F}(x) = \int_{-\infty}^{x} p(t)dt, \qquad x \in \mathbb{R}$$

则称其为连续型，称 $p(x)$ 为 ξ 的密度函数

不只以上两种，随机变量有很多种，分布函数可以分解为三个函数：**纯跳跃函数; 绝对连续函数; 奇异函数的和**

(即 *Lwbesgue* 分解)

2.2 伯努利概型及其分布

2.2.1 伯努利试验

定义 只有两种可能结果的试验叫做伯努利试验，其事件 σ-代数取为 $\mathscr{F} = \{\varnothing, A, \bar{A}, \Omega\}$，通常把结果 A 叫做“成功”，而把结果 $\bar{A}$ 叫做“失败”，再定义概率：

$$p = P(A) \quad q = P(\bar{A}) \quad (p, q > 0 \quad 且 \quad p + q = 1)$$

这样就定义了一重伯努利试验的概率空间 $(\Omega, \mathscr{F}, P)$

2.2.2 二项分布

即 n 重伯努利试验，记作 $B(n, p)$

$$\mathbf{b(k; n, p)} = P\{\xi = k\} = C_n^k p^k q^{n-k}, \qquad k = 0, 1, 2, ..., n$$

$$\frac{b(k; n, p)}{b(k-1; n, p)} = \frac{(n-k+1)p}{kq} = 1 + \frac{\mathbf{(n+1)p} - k}{kq}$$

$\mathbf{(n+1)p}$(不是正整数时) 即为 n 重伯努利试验的最可能成功次数，否则有两个最值：$\mathbf{(n+1)p}$ 与 $\mathbf{(n+1)p-1}$

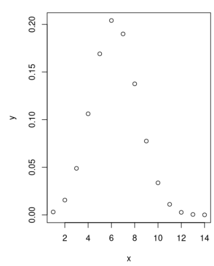
2.2.3 几何分布 *Forever-Young*

定义 考虑可列重伯努利试验中首次成功的等待次数 ξ。它则取值自然数 k 当且仅当前 k-1 次实验全失败，同时第 k 次成功。给定参数 p(0<p<1,q=1-p), 令

$$g(k; p) = q^{k-1}p > 0, \qquad k = 1, 2, \cdots$$

$$\sum_{k=1}^{\infty} g(k; p) = \sum_{k=1}^{\infty} q^{k-1}p = \frac{p}{q-1} = 1$$

于是 $g(k; p)$ 可以作为离散型分布的密度，称它为几何分布，记作 $G(p)$



二项分布

几何分布的性质

无记忆性　　取值与自然数的随机变量 ξ 有几何分布当且仅当 ξ 有**无记忆性**。

$$P\{\xi > m+n|\xi > m\} = P\{\xi > n\}, \qquad \forall m,n > 1$$

2.2.4　Pascal 分布

定义　考虑可列重伯努利试验中首次成功的等待次数 ξ。它表示第 r 次成功所需要的实验次数，ξ 的可能取值为 $k, k+1, \cdots$ 且

$P\{\xi_r = k\} = P\{$前 k-1 次试验恰好有 r-1 次成功且第 k 次成功$\} = b(r-1; k-1, p) \cdot p = C_{k-1}^{r-1} p^r q^{k-r}$

$$f(k;r,p) = C_{k-1}^{r-1} p^r q^{k-r} \qquad k = r, r+1, \cdots$$

于是称以 $f(k;r,p)$ 为密度的离散型分布为 Pascal 分布或负二项分布，记作 $G(p)$

Banach 火柴问题

试验总次数：s=n+1+n-r=2n+1-r,

P=2P(E)=2f(2n+1-r;n+1,$\frac{1}{2}$) $= \binom{2n-r}{n} 2^{r-2n}$

分赌注问题

在可列重伯努利分佈中試驗中，求事件 $E = \{n$n 次成功发生在 m 次失败之前的概率$\}$

2.2.5　Poisson 分布

泊松定理　设有一二项分布 $\{b(k;n,P_n)\}$，参数列 $\{P_n\}$ 满足 $\lim_n nP_n = \lambda > 0$ 则对于任何非负整数 k 有

需要使用不等式：
$|a^n - b^n| \le n|a-b|$

$$\begin{aligned}
\lim_n b(k;n,p_n) &= \binom{n}{k} P_n^k \cdot (1-P_n)^{n-k} \\
&= \frac{p_n^k}{k!} P_n^k \cdot (1-P_n)^{n-k} \\
&= \frac{n!}{(n-k)!k!} P_n^k \cdot (1-P_n)^{n-k} \\
&= \frac{1}{k!} n \cdot (n-1) \cdots (n-k+1)\ P_n^k \cdot (1-P_n)^{n-k} \\
&= \frac{1}{k!} 1 \cdot (1-\frac{1}{n}) \cdots (1-\frac{-k+1}{n})\ (nP_n)^k \cdot (1-P_n)^{n-k} \\
\lim_{n\to\infty} P\{\lambda = k\} &= \frac{1}{k!} \lim_{n\to\infty} (nP_n)^k (1-P_n)^{n-k} \\
&= \frac{1}{k!} \lim_{n\to\infty} (n \cdot \frac{\lambda}{n})^k (1-\frac{\lambda}{n})^{-k} (1-\frac{\lambda}{n})^n = \frac{\lambda^k}{k!} \Big[\lim_{n\to\infty} (1+\frac{-\lambda}{n})^{-\frac{n}{\lambda}}\Big]^{-\lambda} \\
&= \frac{\lambda^k}{k!} e^{-\lambda}
\end{aligned}$$

泊松分布 给定参数 $\lambda > 0$，令 $p(k;\lambda) = \frac{\lambda^k}{k!}e^{-\lambda}$ $k = 0, 1, \cdots$ 且 $\sum_{k=0}^{\infty} p(k;\lambda) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}e^{-\lambda} = 1$ 称以 $p(k;\lambda)$ 为密度的离散型分布为 Poisson 分布，记作 $P(\lambda)$

Poisson 分布的性质

1. 当 λ 是整数时，$p(k;\lambda)$ 在 $m = \lambda$ 及 $m = \lambda - 1$ 时同时取到最大值
2. 当 λ 不是整数时，$p(k;\lambda)$ 在 $m = \lambda$ 时取到最大值

Poisson 过程质点流 假定有一个于随机时刻陆续到来的质点流，质点一个一个地到达，但质点到达的时间间隔都是随机变量。对任何 $t \leq 0$，以 ξ_t 代表在 $[0, t)$ 时刻到达的质点个数，每个 ξ_t 都是非负整值随机变量。性质如下：

Poisson 过程满足的条件

1. 随机增量性 在不相交时段内到达的质点数目相互独立。
2. 平稳性 在长为 t 的时段 $[a, a+t)$ 内到达 k 个质点的概率，至于计时长度 t 有关而与计时起点 a 无关。
3. 普通性 在有限的时间区间内，只来有限个质点。

2.3 连续型分布

2.3.1 正态分布

定义 设连续型随机变量 ξ 具有概率密度：

$$\varphi_{\mu,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad -\infty < x < +\infty$$

则称 ξ 服从参数为 a, σ^2 的**正态分布**，记为 $\xi \sim N(\mu, \sigma^2)$ 标准正态分布如下：

大多数现象都近似服从正态分布

$$\varphi_{0,1}(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

图形关于 x=a 对称

分布函数

$$\Phi_{\mu,\sigma}(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dt$$

$$\Phi_{0,1}(x) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{x^2}{2}}dt$$

无法用初等形式表示

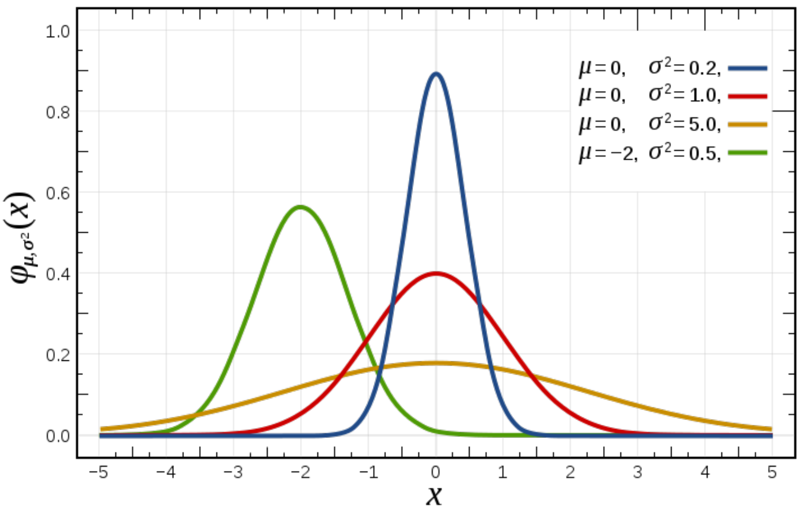


图 2.1: 四个不同参数集的概率密度函数（红色线代表标准正态分布）

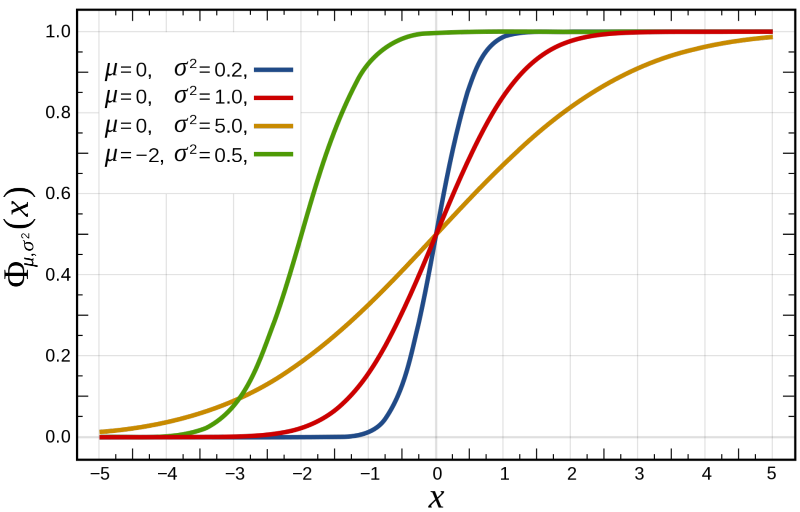


图 2.2: 概率密度函数的累积分布函数

性质

1. $P\{\xi < a\} = \Phi(a)$
2. $P\{\xi > a\} = 1 - \Phi(a)$
3. $P\{|\xi| < a\} = 2\Phi(a) - 1$
4. $\Phi(-x) = 1 - \Phi(x)$
5. 标准化： $\xi \sim N(\mu, \sigma^2) \Longrightarrow \eta = \frac{\xi-\mu}{\sigma} \sim N(0,1)$
6. 3σ 原则： 概率高达 99.74%

2.3.2 Γ 分布

定义 ：(Poisson 过程) 实质上是关于质点到达时间的分布，以 η_r 表示第 r 个质点到达时刻，于是对于 $t>0$, 事件 $\{\eta_r<t\}$ 表示第 r 个质点在时刻 t 以前到来，事件 $\{\eta_r\geq t\}$ 表示到时时刻 t 为止，到来的质点数目不少于 r。分布函数：

$$F(t)=P\{\eta_r<t\}=P\{\eta_t\geq r\}=\sum_{k=r}^{\infty}\frac{(\lambda t)^k}{k!}e^{-\lambda t}$$

关于 Γ 函数的几个性质：

$$\Gamma(r)=\int_0^{+\infty}x_{r-1}e^{-x}dx;\qquad \Gamma(n+1)=n\Gamma(n)=\cdots=n!;\qquad \Gamma(\frac{1}{2})=\sqrt{\pi}$$

密度函数：对 t 求导以及简单计算

$$p(t)=\frac{\lambda^r}{(r-1)!}x^{r-1}e^{-\lambda x}=\frac{\lambda^r}{\Gamma(r)}x^{r-1}e^{-\lambda x},\qquad x>0$$

2.3.3 指数分布

定义 ：称 $\Gamma(\lambda,1)$ 分布为指数分布，其密度函数和分布函数分别为:

$$p(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & x\leq 0,\end{cases}\qquad F(x)=\begin{cases}1-e^{-\lambda x}, & x>0,\\ 0, & x\leq 0,\end{cases}$$

指数分布也有无记忆性： $p(\xi>a)=1-p(\xi<a)=1-F(a)=e^{-\lambda a}$

2.4 多维概率分布

2.4.1 随机向量

定义 ：与一维随机变量的定义类似 $(\xi,\eta)\in(\Omega,\mathscr{F},\rho)\cdots$一维 *Borel* 集的 *Descartes* 乘积

2.4.2 联合分布

设 $(\xi,\eta)\in(\Omega,\mathscr{F},\rho)$ 称 $F(x,y)=P\{\xi<x,\eta<y\},\qquad x,y,\in\mathbf{R}$
为 (ξ,η) 的联合分布函数。

联合分布函数的性质

1. $0\leq F(x,y)\leq 1$ 而且 $F(+\infty,+\infty)=1$
   $F(x,-\infty)=f(-\infty,y)=F(-\infty,+\infty)=0$
2. $F(x,y)$ 对于每个自变量都是左连续的。
3. $F(x,y)$ 对于每个自变量都是单调非降的。
4. $P\{x_1\leq\xi\leq x_2,y_1\leq\eta\leq y_2\}=F(x_2,y_2)-F(x_1,y_2)-F(x_2,y_1)+F(x_1,y_1)$ 亦即 $F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}p(\upsilon,\nu)d\upsilon d\nu$

2.4.3 边缘分布

考虑随机变量各自的分布，各变量的边缘分布函数如下：

$$F_\xi(x) = \lim_{y\to+\infty} F(x,y) = F(x,+\infty) \qquad x \in \mathbf{R}$$
$$F_\eta(y) = \lim_{x\to+\infty} F(x,y) = F(+\infty,y) \qquad y \in \mathbf{R}$$

2.4.4 边缘密度

边缘分布由联合分布完全决定，但边缘分布不能确定联合分布。

$$p_1(x) = \int_{-\infty}^{+\infty} p(x,y)dy, \qquad x \in \mathbf{R}$$

$$p_1(y) = \int_{-\infty}^{+\infty} p(x,y)dx, \qquad y \in \mathbf{R}$$

2.4.5 二维均匀分布

如果 (ξ,η) 的联合分布密度函数为：$p(x,y) = \begin{cases} \frac{1}{m_D}, & (x,y) \in \mathbf{D}, \\ 0, & others, \end{cases}$

2.4.6 二维正态分布

略之

2.5 随机变量的独立性

2.5.1 条件分布

分布函数 ：设为一概率空间，$B \in \mathscr{F}, P(B) > 0$, 则对于任何事件 $A \in \mathscr{F}$，则此概率空间上的随机变量 ξ 关于 B 的条件分布函数为：$F(x|B) = P\{\xi < x|B\}$，再添加一个变量 η, 则当事件 $B = \{\eta = y\}$ 有正概率时，称下式为 $\eta = y$ 已知时 ξ 的条件分布函数。

$$F(x|\eta = y) = P\{\xi < x|\eta = y\} = \frac{P\{\xi < x, \eta = y\}}{P\{\eta = y\}}$$

密度函数 由条件函数的定义，把上述结果简写成下面形式：

$$P\{\xi = x_i|\eta = y_j\} = \frac{P_{ij}}{p_j} \qquad P\{\eta = y_j|\xi = x_i\} = \frac{P_{ij}}{p_i}$$

边缘分布律简写：
$P\{\xi = x_i, \eta = y_j\} = P_{ij}$
$P\{\xi = x_i\} = P_i$
$P\{\eta = y_j\} = P_j$
$i, j = 1, 2, \cdots$

根据联合密度函数与边缘分布可以得到密度函数如下：

$$p(x|\eta = y) = \frac{p(x,y)}{p_2(y)}$$

2.5.2 相互独立的随机变量

定义 设 $\xi_1, \cdots, \xi_n$ 是概率空间（$\Omega, \mathscr{F}, P$）上 n 个随机变量，如果他们的联合分布函数等于他们边缘分布函数的乘积，即：

$$F(x_1, \cdots, x_n) = F_1(x_1) \cdots F_n(x_n) = \prod_{i=1}^{n} F_i(x_i)$$

则称 $\xi_1, \cdots, \xi_n$ 相互独立。

独立性的定理

1. 如果随机变量 $\xi_1, \cdots, \xi_n$ 相互独立，则其中任意部分子向量仍独立
2. 两向量独立 $\xleftrightarrow{\text{充要条件}} P\{\xi \in \mathbf{B_1}, \eta \in \mathbf{B_2}\} = P\{\xi \in \mathbf{B}_1\}P\{\eta \in \mathbf{B}_2\} \quad \forall \mathbf{B_1}, \mathbf{B_2} \in \mathscr{B}$
3. 离散型随机变量相互独立 $\xleftrightarrow{\text{充要条件}} P_{ij} = P_i\ P_j \qquad \forall i, j = 1, 2, \cdots$
4. 连续型随机变量相互独立 $\xleftrightarrow{\text{充要条件}} p(x, y) = p(x)\ p(y) \qquad \forall x, y \in \mathbf{R}$

2.6 随机变量函数的分布

2.6.1 一维随机变量的分布

已知分布函数求分布密度即

$$P\{\eta < y\} = P\{f(\xi) < y\} = \int_{\{x: f(x) < y\}} dF_\xi(x)$$

一维离散型

对于离散型要算出所有可能结果

离散卷积公式 设 ξ 与 η 是相互独立的非负整值随机变量, 各有分布 $\{a_k\}$ 与 $\{b_k\}$, 那么其和的分布为:

$$\begin{aligned} P\{\xi + \eta = n\} &= \sum_{k=0}^{n} P\{\xi = k\}P\{\xi + \eta = n | \xi = k\} \\ &= \sum_{k=0}^{n} P\{\xi = k\}P\{\eta = n - k\} \\ &= \sum_{k=0}^{n} a_k b_{n-k}, \qquad n = 0, 1, 2, \cdots \end{aligned}$$

一维连续型

这个用两种方法求解:

1. 当随机变量 ξ 的密度函数 $p_\xi(x)$, 在区间内单调连续, 而且存在唯一的**反函数** $x=h(y)$, $h'(y)$ 存在且连续, 若 $\eta=f(\xi)$, 那么 η 的密度函数为:

$$p_\eta(y)=p_\xi[h(y)]|h'(y)|,\qquad y\in(\alpha,\beta)$$

2. 当随机变量在可分割的区间 (α,β) 内, 每一子区间上都有上述的反函数, 而且都连续, 那么 $\eta=f(\xi)$ 仍然为连续型变量, 起密度函数为:

$$p_\eta(y)=\begin{cases}\sum\limits_i p_\xi[h(y)]|h'(y)|, & \text{反函数存在}\\ 0 & \text{反函数不存在}\end{cases}$$

2.6.2 二维连续型

设二维随机变量 (ξ,η) 的联合密度为 $p(x,y)$, 求随机变量 $\zeta=f(\xi,\eta)$ 的概率密度。常用的方法为公式法: 即先求分布函数再求概率密度

$$F_\zeta(z)=P\{\zeta<z\}=P\{f(\xi,\eta)<z\}=\iint_{f(x,y)<z}p(x,y)dxdy\xRightarrow{Z\text{ 的密度函数}}P_\zeta(z)=\frac{dF_\zeta(z)}{dz}$$

和的分布 设二维随机变量 (ξ,η) 的联合分布函数为 $F(x,y)$, 则 $\zeta=\xi+\eta$ 的分布函数为:

$$\begin{aligned}F_\zeta(z)=P\{\xi+\eta<z\}&=\iint_{x+y<z}p(x,y)dxdy\\&=\int_{-\infty}^{+\infty}\int_{-\infty}^{z-x}p(x,y)dxdy\\&=\int_{-\infty}^{z}\left[\int_{-\infty}^{+\infty}p(x,u-x)dx\right]du\\P_\zeta(z)&=\int_{-\infty}^{+\infty}P(x,z-x)dx\xleftrightarrow{\text{相互独立时}}\int_{-\infty}^{+\infty}P_\xi(x)P_\eta(z-x)dx\end{aligned}$$

连续卷积公式 称 $\int_{-\infty}^{+\infty}f(u)g(x-u)du$ 为函数 $f(x)$ 与 $g(x)$ 的卷积, 记为 $f(x)*g(x)$, 所以独立随机变量和的概率密度是各概率密度的卷积

$$P_\zeta(z)=P_\xi(x)*P_\eta(y)=P_\eta(y)*P_\xi(x)=\int_{-\infty}^{+\infty}P_\xi(x)P_\eta(z-x)dx$$

商的分布 设二维随机变量 (ξ,η) 的密度函数为 $p(x,y)$ 则 $\zeta=\frac{\xi}{\eta}$ 的概率密度为:

$$\begin{aligned}F_\zeta(z)=P\{\frac{\xi}{\eta}<z\}&=\iint_{\frac{x}{y}<z}p(x,y)dxdy\\&=\iint_{x<yz,y>0}p(x,y)dxdy+\iint_{x>yz,y<0}p(x,y)dxdy\\P_\zeta(z)&=\int_{-\infty}^{+\infty}|y|p(yz,y)dy\xleftrightarrow{\text{相互独立时}}\int_{-\infty}^{+\infty}|y|p_\zeta(yz)p_\eta(y)dy\end{aligned}$$

第三章　数字特征与特征函数

3.1　数学期望

3.2　其它数字特征

3.3　母函数

3.4　特征函数

3.5　* 多元正态分布

1. 已知 x～E(1)(参数为一的指数分布)　，求 E(x+e^{-2x})

   解：由题可得

$$f(x)=\begin{cases} e^{-x}, & x>0 \\ 0, & others \end{cases}$$

$$\begin{aligned} E(x+e^{-2x}) &= E(x)+E(e^{-2x}) \\ &= \int_{-\infty}^{+\infty} xe^{-x}dx + \int_{-\infty}^{+\infty} e^{-2x}e^{x} \\ &= \int_{0}^{+\infty} xe^{-x}dx + \int_{0}^{+\infty} e^{-2x}e^{x} \\ &= 1+\frac{1}{3} \\ &= \frac{4}{3} \end{aligned}$$

2. 母函数

第四章　极限定理

大数定律　：关于随机变量平均结果的极限定理

中心极限定理　：标准化

4.1　随机变量列的收敛性

4.1.1　收敛性

依概率收敛　对于 $\forall \varepsilon > 0$

$$\lim_{n\to\infty} P\{|\xi_n(\omega) - \xi(\omega)| \geq \varepsilon\} = 0$$

称 $\{\xi_n\}$ 依概率收敛到 ξ，记作 $\xi_n \xrightarrow{P} \xi$ 即当 n 充分大时，$\{\xi_n\}$ 与 ξ 两者有充分大差异的概率任意小。

几乎必然收敛　如果有

$$P\{\omega : \lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\} = 1$$

称随机变量数列 $\{\xi_n\}$ 几乎必然收敛到 ξ，记作 $\xi_n \xrightarrow{a.s.} \xi$

4.2　大数定律

4.3　中心极限定理

第五章　数理统计部分

5.1　总体与样本

总体与个体　研究对象的全体称为总体，总体中所包含的个体的个数称为总体的容量。总体中每个成员称为个体。

1. 由于每个个体的出现是随机的，所以相应的数量指标的出现也带有随机性。从而可以把这种数量指标看作一个随机变量 X，因此随机变量 X 的分布就是该数量指标在总体中的分布。
2. 总体就可以用一个随机变量及其分布来描述。因此在理论上可以把总体与概率分布等同起来。统计中，总体这个概念的要旨是：总体就是一个随机变量 (向量) 或一个概率分布。

样本　总体中抽出若干个体而成的集体，称为样本。样本中所含个体的个数，称为样本容量。

抽样　统计中，采用的抽样方法是随机抽样法，即子样中每个个体是从总体中随意地取出来的。

抽样的分类

1. 重复（返回）抽样：从总体中抽取个体检查后放回，总体成分不变（分布不变）。样本 $X_1, X_2, \cdots, X_n$ 相互独立，与总体有相同的分布。
2. 非重复(无返回)抽样：对有限总体取出样本后改变了总体的成分，所以 $X_1, X_2, \cdots, X_n$ 不相互独立；对无限总体而言做无返回抽取，并不改变总体的成分，$X_1, X_2, \cdots, X_n$ 相互独立，与总体有相同的分布。

常用方法　简单随机抽样。

1. 代表性（随机性）:。从总体中抽取样本的每一个分量 X_k 是随机的，每一个个体被抽到的可能性相同。
2. 独立同分布性：$X_1, X_2, \cdots, X_n$ 相互独立，其中每一个分量 X_k 与所考察的总体有相同的分布。

样本联合分布 若总体的分布函数为 $F(x)$ 、概率密度为 $f(x)$ ，则其简单随机样本的联合分布函数为

$$F_{X_1,X_2,\cdots,X_n}(x_1,x_2,\cdots,x_n)=F(1)F(2)\cdots F(n) \tag{5.1}$$

其简单随机样本的联合概率密度函数为

$$f_{X_1,X_2,\cdots,X_n}(x_1,x_2,\cdots,x_n)=f(1)f(2)\cdots f(n) \tag{5.2}$$

样本经验分布函数 在 n 次独立重复实验中，事件 $\{X\leqslant x\}$ 发生的频率

$$\hat{F}_n(x;X_1,X_2,\cdots,X_n)=\frac{1}{n}\sum_{i=1}^{n}I(X_i\leqslant x) \tag{5.3}$$

具有分布函数的一切性质。是在每个数据点 X_i 上权重相等的均匀分布的分布函数。

性质

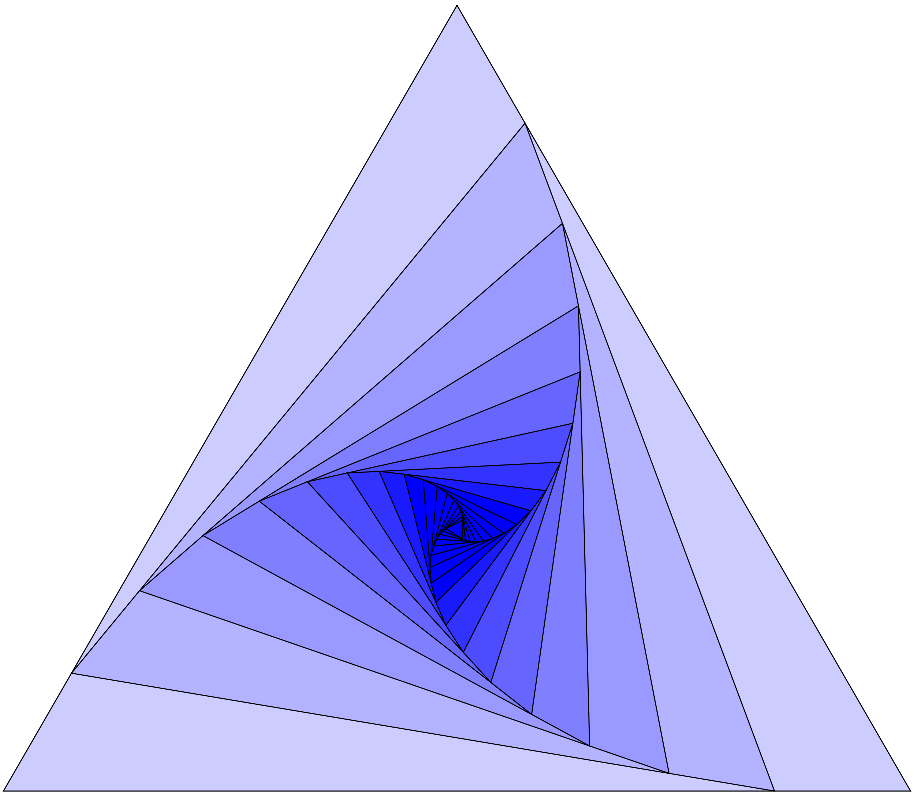
1. 给定 x，$\hat{F}_n(x)$ 是一个随机变量：$n\hat{F}_n(x)$ 服从二项分布 $b(n,F(x))$
2. $E(\hat{F}_n(x))=F(x)$
3. $D(\hat{F}_n(x))=\frac{F(x)(1-F(x))}{n}\to 0$
4. $\hat{F}_n(x)\xrightarrow{P}F(x)$
5. Dvoretzky-Kiefer-Wolfowitz (DKW) 不等式：如果 $X_1,X_2,\cdots,X_n\sim F$ ，则对任意 $\epsilon>0$

$$P\left\{\sup_x\left|\hat{F}_n(x))-F(x)\right|>\epsilon\right\}\leqslant 2e^{-2n\epsilon^2} \tag{5.4}$$

格列汶科定理 当 $n\to\infty$ 时，$\hat{F}_n(x)$ 以概率 1 关于 x 一致收敛于 $F(x)$ ，即

$$P\left\{\lim_{n\to\infty}\sup_{-\infty<x<\infty}\left|\hat{F}_n(x)-F(x)\right|=0\right\}=1 \tag{5.5}$$

当样本容量 n 足够大时，对所有的 x，$\hat{F}_n(x)$ 与 $F(x)$ 之差的绝对值都很小，这件事发生的概率为 1 。



5.2 回归分析与方差分析

5.3 统计决策与贝叶斯统计