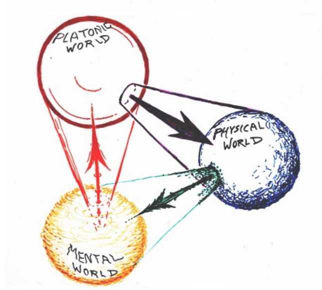
$$life = \int_{birth}^{death} study\, \mathrm{d}t$$

作者：王子毅

时间：June 23, 2023

版本：0.1



目录

1 随机事件与概率 **1**

1.1 随机事件及其运算 . . . 1
1.2 概率的定义及其确定方法 . . . 2
1.3 概率的公理化定义 . . . 4
1.4 条件概率与事件独立性 . . . 7
1.5 全概率公式与贝叶斯公式 . . . 9

2 随机变量及其分布 **11**

2.1 离散随机变量 . . . 11
2.2 随机变量的分布函数 . . . 13
2.3 连续型随机变量 . . . 13
2.4 随机变量函数的分布 . . . 17

3 多维随机变量及其分布 **19**

3.1 二维随机变量 . . . 19
3.2 边缘分布 . . . 20
3.3 二维随机变量的条件分布 . . . 21
3.4 二维随机变量的独立性 . . . 22
3.5 二维随机变量函数的分布 . . . 22

4 随机变量的特征数 **24**

4.1 数学期望 . . . 24
4.2 方差 . . . 25
4.3 协方差 . . . 26
4.4 相关系数 . . . 27

第 1 章 随机事件与概率

1.1 随机事件及其运算

定义 1.1

1. 随机现象：个体上表现为不确定性，而大量观察中呈现出统计规律的现象
2. 随机试验：对随机现象进行的观察或试验，通常用 T 来表示

♣

注 满足:

1. 重复性
2. 明确性 (所有结果)
3. 随机性 (不可预言)

定义 1.2 (样本点与样本空间)

1. 随机试验的每一个可能的结果称为这个试验的样本点记作 ω
2. 全体样本点的集合称为样本空间，记作 $\Omega = \{\omega\}$

♣

例题 1.1 抛一枚硬币的样本空间为 $\Omega_1 = \{\omega_1, \omega_2\}$，其中 ω_1 表示正面朝上，ω_2 表示反面朝上

定义 1.3 (随机事件)

1. 随机试验的某些样本点构成的集合称为随机事件 注 随机事件是样本空间 Ω 的子集
2. 基本事件是由一个样本点组成的集合
3. 复合事件是由多个样本点组成的集合
4. 必然事件：由全体样本点组成的集合 (Ω 的最大子集)，记为 Ω
5. 不可能事件：不包含任何样本点的集合 (Ω 的最小子集)。记为空集 $\varnothing$

♣

定义 1.4 (随机变量)

用来表示随机现象的结果的变量称为随机变量

♣

注 在同一个随机现象中不同的设置可获得不同的随机变量，如何设置可按需进行

定义 1.5 (事件之间的关系和运算)

1. 包含关系：事件 A 发生必然导致事件 B 发生，记作 $A \subset B$
2. 相等关系：A = B ⇔ A⊂B 且 B⊂A
3. 事件的和 (并)：事件 A 或 B 至少有一个发生，记作 $A \cup B(A + B)$
4. 事件的积 (交)：事件 A、B 同时发生，记作 $A \cap B(AB)$
5. 互斥关系: 事件 A、B 不可能同时发生，即 AB=∅

   注 基本事件是互不相容的
6. 互逆 (对立) 关系: 事件 A、B 只有一个发生且必有一个发生，即 $A \cap B(AB) = \varnothing$ 且 $A \cup B = \Omega$, 记作 $A = \overline{B}, B = \overline{A}$

   注 对立事件一定是互不相容的事件，但互不相容的事件不一定是对立事件
7. 事件的差：事件 A 发生而 B 不发生，这样的事件称为事件 A 与 B 的差，记作 A-B
8. 完备事件组：$A_i A_j = \varnothing (i \neq j, i, j = 1, 2, ...n)$ 或者 $\bigcup_{i=1}^{n} A_i = \Omega$

   注

   (a). 一事件的基本事件构成完备事件组

(b). A 与 $\overline{A}$ 构成完备事件组

(c). 概念推广：可列个事件 $A_k(k=1,2...n)$ 构成完备事件组

命题 1.1 (事件运算法则)

1. 分配律：
   (a). $(A\cup B)\cap C = AC\cup BC$
   (b). $(A\cap B)\cup C = (A\cup C)\cap(B\cup C)$
2. De Morgan 公式：
   (a). $\overline{A\cup B} = \overline{A}\cap\overline{B}$
   (b). $\overline{A\cap B} = \overline{A}\cup\overline{B}$
   注 该公式可以推广到多个事件以及可列个事件的场合
3. $A-B = A\overline{B} = A-AB$
   注
   (a). $A-B=\varnothing \Leftrightarrow A\subset B$
   (b). $A-B=A \Leftrightarrow A\cap B=\varnothing$
4. $\overline{A} = \Omega - A$

1.2 概率的定义及其确定方法

定义 1.6 (频率的定义)

频率定义：事件 A 在 n 重复试验出现了 n_A 次，则比值 n_A/n 称为事件 A 在 n 次重复试验中出现的频率，记作 $f_n(A)=n_A/n$

命题 1.2 (频率的性质)

可加性：若 $AB=\varnothing$, 则 $f_n(A\cup B)=f_n(A)+f_n(B)$

可列可加性公理：若 $A_1,A_2...A_n$ 互不相容，则

$$P\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i)$$

注 概率可以通过频率来“测量”，频率是概率的近似

定义 1.7 (频率的统计定义)

事件 A 的 n 次重复试验中出现了 n_A 次，频率 $f_n(A)=n_A/n$ 随着 n 的增大总在某一个固定的数值 p 附近摆动，称 p 为事件 A 发生的概率，记作 P(A)=p

注 $P(A)\approx f_n(A)$, 正如木棒长度客观存在，但无论测量仪器多么精确，将测量值当作真实值总是一种近似

命题 1.3 (排列组合公式)

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

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

注

$$P_n^k = C_n^k \cdot k!$$

定义 1.8 (古典概型)

若实验 T 满足

1、有限性：样本空间 Ω={$\omega_1, \omega_2, ...\omega_n$};

2、等可能性：$P(\omega_1) = P(\omega_2) = ... = P(\omega_n)$

则称 T 为古典概型也叫等可能概型

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

m 为事件 A 包含的样本点数，n 为样本点总数

例题 1.2 当班级有 n 个人 ($n \leq 365$), 问至少两个人生日在同一天概率是多少

解

$$P(A) = 1 - P(\overline{A}) = 1 - \frac{p_{365}^n}{365^n} = 1 - \frac{365!}{365^n \cdot (365-n)!}$$

例题 1.3 30 名学生中有 3 名运动员，将 30 名学生平均分三组，求

(1) 每组有一名运动员的概率 (2)3 名运动员在一组的概率

解

$$(1)P(A) = \frac{3! \cdot C_{27}^9 \cdot C_{18}^9 \cdot C_9^9}{C_{30}^{10} \cdot C_{20}^{10} \cdot C_{10}^{10}} = \frac{50}{203}$$

$$(2)P(B) = \frac{C_3^1 \cdot C_{27}^7 \cdot C_{20}^{10} \cdot C_{10}^{10}}{C_{30}^{10} \cdot C_{20}^{10} \cdot C_{10}^{10}} = \frac{18}{203}$$

定义 1.9 (几何概型)

(1) 设样本空间 Ω 是空间平面上某区域，其面积 S_Ω

(2) 向区域 Ω 上随机投掷一点，该点落入 Ω 内任何区域内的可能性只与该区域的面积成正比，而与这部分区域的位置和形状无关

(3) 设事件 A 是 Ω 的某个区域，其面积为 S_A, 则向区域 Ω 随机投掷一点该点落在 A 内的概率

$$P(A) = \frac{S_A}{S_\Omega}$$

(4) 若样本空间用一线段，或空间某个区域表示，仍可用长度或体积来类似定义

例题 1.4 甲、乙相约在 0 到 T 这段时间内在预定地点会面, 先到的人等候另一人 t 时间离去. 设每人在 0 到 T 这段时间内各时刻到达该地是等可能的, 且两人到达的时刻互不相连. 求甲、乙两人能会面的概率.

解 以 (x,y) 表示两个人到达的时刻，则 $0 \leq x \leq T$，$0 \leq y \leq T$

以 x、y 表示平面上点的坐标，所有可能到达时刻组成的点可用，两人可能会面的充要条件是

$$|x-y| \leq t$$

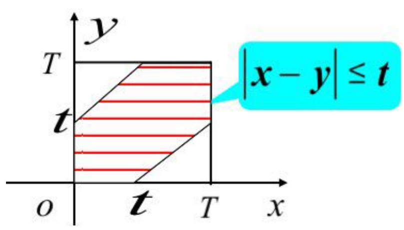


图 1.1: 例题 1.4

所求概率

$$P(A)=\frac{S_{black}}{S_{square}}=\frac{T^2-(T-t)^2}{T^2}=1-(1-\frac{t}{T})^2$$

例题 1.5 从五双不同尺码的鞋子里任取四只，求至少有两只配成一双的概率

解

1. 配成 2 双：C_5^2
2. 配成 1 双：$C_5^1(C_4^2\cdot C_2^1\cdot C_2^1)$ 或者 $C_5^1(C_8^2-C_4^1)$
3. 均不成双：$\sum_{i=0}^4 C_5^i C_{5-i}^{4-i}$(分左右) 或者 $C_5^4\cdot C_2^1\cdot C_2^1\cdot C_2^1\cdot C_2^1$(五选四每双任选一只)

$$\begin{aligned}P(A)&=\frac{C_5^1(C_4^2\cdot C_2^1\cdot C_2^1)+C_5^2}{C_{10}^4}\\&=\frac{C_5^1(C_8^2-C_4^1)+C_5^2}{C_{10}^4}\\&=1-\frac{\sum_{i=0}^4 C_5^i C_{5-i}^{4-i}}{C_{10}^4}\\&=1-\frac{C_5^4\cdot C_2^1\cdot C_2^1\cdot C_2^1\cdot C_2^1}{C_{10}^4}\\&=\frac{130}{210}\end{aligned}$$

注

1. 先保证 2 只成双：C_5^1，其余 2 只从余下的八只中任意选 C_8^2, 错误！有重复！
   改正：$C_5^1\cdot C_8^2-C_5^2$
2. 配成两双 C_5^2，配成一双:$C_5^1\cdot C_8^1\cdot C_6^1$
   共有 $C_5^1\cdot C_8^1\cdot C_6^1+C_5^2$，错误！有重复！
   改正:$C_5^1\frac{C_8^1\cdot C_6^1}{2}+C_5^2$

1.3 概率的公理化定义

定义 1.10

设随机试验 T 的样本空间为 Ω, 对于试验 T 的每一个事件 A, 总有唯一确定的实数 P(A) 与之对应, 若 P(A) 满足三条性质:

(1)$0\le P(A)\le 1$;

(2)$P(\Omega)=1$;

(3) 设 $A_1,A_2...$ 为两两互不相容事件，即 $A_iA_j=\varnothing(i\neq j)$

则

$$P\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i)$$

则称 P(A) 为事件 A 的概率

♣

命题 1.4 (概率的性质)

1. $P(\varnothing)=0$
2. 事件的有限可加性：$P\left(\bigcup_{i=1}^n A_i\right)=\sum_{i=1}^n P(A_i)$
3. 可减性:A、B 是两个事件，则 $P(A-B)=P(A)-P(AB)$
   注 若事件 $B\subset A$ 则 $P(A-B)=P(A)-P(B)$

4. 加法公式：对任意两事件 A、B 有 $P(A \cup B) = P(A) + P(B) - P(AB)$

   注 对任意事件 A、B、C, 有 $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$

5. 概率的连续性：若 $P(\cdot)$ 是事件域 $\mathscr{F}$ 上的一个概率函数，则 $P(\cdot)$ 既是下连续的，又是上连续的

定义 1.11 (事件序列的极限)

若事件序列 $\{F_n\}$，满足：$F_1 \subset F_2 \subset \dots F_n \subset \dots$ 则称 F_n 为单调不减事件序列，其极限事件为

$$\lim_{n\to\infty} F_n = \bigcup_{n=1}^{\infty} F_n$$

若事件序列 $\{E_n\}$，满足：$E_1 \supset E_2 \supset \dots E_n \supset \dots$ 则称 E_n 为单调不增事件序列，其极限事件为

$$\lim_{n\to\infty} E_n = \bigcap_{n=1}^{\infty} E_n$$

定义 1.12 (集合函数的连续性)

设 $P(\cdot)$ 是一个集合函数

1. 若任对单调不减集合序列 F_n，有

$$P(\lim_{n\to\infty} F_n) = \lim_{n\to\infty} P(F_n)$$

   则称 $P(\cdot)$ 是下连续的

2. 若任对单调不增集合序列 E_n，有

$$P(\lim_{n\to\infty} E_n) = \lim_{n\to\infty} P(E_n)$$

   则称 $P(\cdot)$ 是上连续的

证明 [命题 1.4.5]

(1) 先证明 $P(\cdot)$ 的下连续性

设 $\{F_n\}$ 是单调不减事件序列，根据下连续的定义，即证：

$$\begin{aligned} P(\lim_{n\to\infty} F_n) &= \lim_{n\to\infty} P(F_n) \\ &= P\left(\bigcup_{n=1}^{\infty} F_n\right) \end{aligned}$$

由于 $\{F_n\}$ 是单调不减，我们有

$$\bigcup_{n=1}^{\infty} F_n = \bigcup_{i=1}^{\infty} (F_i - F_{i-1})$$

其中 $F_0 = \varnothing$

$$
\begin{aligned}
\lim_{n\to\infty} P(F_n) &= \lim_{n\to\infty} P\left(\bigcup_{i=1}^{n}(F_i - F_{i-1})\right) \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} P(F_i - F_{i-1}) \\
&= \sum_{i=1}^{\infty} P(F_i - F_{i-1}) \\
&= P\left(\bigcup_{i=1}^{\infty}(F_i - F_{i-1})\right) \\
&= P\left(\bigcup_{n=1}^{\infty}(F_n)\right) \\
&= P(\lim_{n\to\infty} F_n)
\end{aligned}
$$

(2) 再证明 $P(\cdot)$ 的上连续性

由于 $\{E_n\}$ 是单调不增的事件序列，故 $\{\overline{E}_n\}$ 是单调不减的事件序列

$$
\begin{aligned}
\lim_{n\to\infty} P(\overline{E}_n) &= P(\lim_{n\to\infty} \overline{E}_n) \\
1 - \lim_{n\to\infty} P(E_n) &= P\left(\bigcup_{n=1}^{\infty} \overline{E}_n\right) \\
&= P\left(\overline{\bigcap_{n=1}^{\infty} E_n}\right) \\
&= 1 - P\left(\bigcap_{n=1}^{\infty} E_n\right) \\
&= 1 - P(\lim_{n\to\infty} E_n) \\
\lim_{n\to\infty} P(E_n) &= P(\lim_{n\to\infty} E_n)
\end{aligned}
$$

命题 1.5

若 $P(\cdot)$ 是事件域 $\mathscr{F}$ 上满足：非负、正则的集合函数，则 $P(\cdot)$ 有可列可加性的充要条件是它有有限可加性和下连续性

证明 “$\Rightarrow$” 显然

“$\Leftarrow$”：设 $A_i (i = 1, 2, \ldots)$ 为两两互不相容的事件序列，则由有限可加性可知

$$
\sum_{i=1}^{n} P(A_i) = P\left(\bigcup_{i=1}^{n} A_i\right) \leq 1
$$

即上式左端正项级数收敛，故

$$
\begin{aligned}
\sum_{i=1}^{\infty} P(A_i) &= \lim_{n\to\infty} \sum_{i=1}^{n} P(A_i) \\
&= \lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} A_i\right)
\end{aligned}
$$

令 $\bigcup_{i=1}^{n} A_i = F_n$, 则

$$\begin{aligned}
\sum_{i=1}^{\infty} P(A_i) &= \lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} A_i\right) \\
&= \lim_{n\to\infty]} P(F_n) \\
&= P\left(\lim_{n\to\infty} F_n\right) \\
&= P\left(\bigcup_{n=1}^{\infty} F_n\right) \\
&= P\left(A_1 \bigcup (A_1 \cup A_2) \bigcup (A_1 \cup A_2 \cup A_3) \bigcup \ldots\right) \\
&= P\left(\bigcup_{i=1}^{\infty} A_i\right)
\end{aligned}$$

1.4 条件概率与事件独立性

定义 1.13 (条件概率的定义)

已知事件 A 发生的条件下，事件 B 发生的概率称为事件 A 发生条件下 B 的条件概率，记作 $P(B|A)$

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

♣

命题 1.6 (乘法公式)

$$P(AB) = P(B)P(A|B) = P(A)P(B|A)$$

注

$$P(ABC) = P(A)P(B|A)P(C|AB)$$

♠

问题 1.1 盒中有 3 个红球、2 个白球, 每次从袋中任取一只, 观察其颜色后放回, 并再放入一只与所取球颜色相同的球, 若从盒中连续取球 4 次, 试求第 1、2 次取得白球、第 3、4 次取得红球的概率.

解 设 A_i= “第 i 次取球时取到白球” (i =1, 2, 3, 4), 则

$$\begin{aligned}
P(A_1) = P(A_1A_2\overline{A_3A_4}) &= P(A_1)P(A_2|A_1)P(\overline{A_3}|A_1A_2)P(\overline{A_4}|A_1A_2\overline{A_3}) \\
&= \frac{2}{5} \times \frac{3}{6} \times \frac{3}{7} \times \frac{4}{8} \\
&= \frac{3}{70}
\end{aligned}$$

定义 1.14 (独立事件的定义)

1. 直观定义：事件 A、B，若其中任一事件发生的概率不受另一事件发生的概率的影响，称事件 A、B 相互独立.

$$P(A|B) = P(A)$$

$$P(B|A) = P(B)$$

2. 数学定义：事件 A、B，若 $P(AB) = P(A)P(B)$，则称事件 A、B 相互独立

♣

定理 1.1

以下四个命题等价

1. 事件 A、B 相互独立
2. 事件 $\overline{A}$、B 相互独立
3. 事件 A、$\overline{B}$ 相互独立
4. 事件 $\overline{A}$、$\overline{B}$ 相互独立

定义 1.15 (多个事件独立的定义)

若三个事件 A、B、C 满足两两事件相互独立，在此基础上加上 $P(ABC) = P(A)P(B)P(C)$，则称 A、B、C 相互独立.

推论 1.1

一般的，设 $A_1, A_2 ... A_n$ 是 n 个事件，若对任意 $k(1 < k \leq n)$, 任意 $1 \leq i_1 < i_2 < ... < i_k \leq n$ 具有等式：

$$P\left(\prod_{i=1}^{n} A_{i_k}\right) = \prod_{i=1}^{n} P(A_{i_k})$$

则称这 n 个事件相互独立

问题 1.2 设事件 A、B、C、D 相互独立，则 $A \cup B$ 与 CD 独立吗?

解

$$\begin{aligned}
P[(A \cup B)CD] &= P(ACD \cup BCD) \\
&= P(ACD) + P(BCD) - P(ABCD) \\
&= P(A)P(C)P(D) + P(B)P(C)P(D) - P(A)P(B)P(C)P(D) \\
&= [P(A) + P(B) - P(AB)]P(C)P(D) \\
&= P(A \cup B)P(CD)
\end{aligned}$$

故 $A \cup B$ 与 CD 独立

命题 1.7 (独立事件的加法、乘法公式)

1. 若 A、B 相互独立，则

$$P(AB) = P(A)P(B)$$

$$P(A \cup B) = 1 - P(\overline{A})P(\overline{B})$$

2. 若事件 $A_1, A_2, ..., A_n$ 相互独立，则

$$P(A_1 A_2 ... A_n) = P(A_1)P(A_2)...P(A_n)$$

$$P(A_1 \cup A_2 \cup ... \cup A_n) = 1 - P(\overline{A_1})P(\overline{A_2})...P(\overline{A_n})$$

定义 1.16 (独立试验概型)

在概率论中, 有些试验在同样条件下可以重复进行, 且任何一次试验发生的结果都不受其它各次试验结果的影响. 这种概率模型称做独立试验概型

在 n 次独立试验概型中, 如果对于每一次试验只有两个可能的结果发生, 即 A 发生或 $\overline{A}$ 发生, P(A)>0, 称这样的独立试验概型为 Bernoulli 概型.

定理 1.2

在 Bernoulli 概型中, P(A)=p (0<p<1), 则事件 A 在 n 次试验中恰好发生 k 次的概率为:

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

注 该公式与 $[p+(1-p)]^n$ 二项展开式中第 k+1 项相同, 故称之为参数为 n 和 p 的二项概率公式.

1.5 全概率公式与贝叶斯公式

定理 1.3 (全概率公式)

设事件 $A_1,A_2...A_n$ 构成完备事件组, 且 $P(A_i)>0$, 则对任一事件 B, 有

$$P(B)=\sum_{i=1}^{n}P(A_i)P(B|A_i)$$

问题 1.3 甲、乙、丙三人同时对飞机进行射击, 三人击中的概率分别为 0.4、0.5、0.7. 飞机被一人击中而击落的概率为 0.2, 被两人击中而击落的概率为 0.6, 若三人都击中, 飞机必定被击落, 求飞机被击落的概率。

解 设 A_i= "飞机被 i 人击中" $(i=0,1,2,3)$, B= "飞机被击落"

依题意,$P(B|A_0)=0, P(B|A_1)=0.2, P(B|A_2)=0.6, P(B|A_3)=1$

为求 $P(A_i)$, 设 H_i= 飞机被第 i 人击中 (i=1,2,3)

$$P(A_1)=P(H_1\overline{H_2H_3}+\overline{H_1}+H_2\overline{H_3}+\overline{H_1H_2}H_3)=0.40.50.3+0.60.50.3+0.60.50.7=0.36$$

$$P(A_2)=P(H_1H_2\overline{H_3}+H_1\overline{H_2}H_3+\overline{H_1}H_2H_3)=0.40.50.3+0.40.50.7+0.60.50.7=0.41$$

$$P(A_3)=P(H_1H_2H_3)=0.40.50.7=0.14$$

于是

$$P(B)=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)+P(A_3)P(B|A_3)=0.360.2+0.410.6+0.141=0.458$$

即飞机被击落的概率为 0.458

定理 1.4 (贝叶斯公式)

设事件 $A_1,A_2,...A_n$ 构成完备事件组, 且 $P(A_i)>0$, 对事件 B, 若 $P(B)>0$, 则有

$$P(A_k|B)=\frac{P(A_kB)}{P(B)}=\frac{P(A_k)P(B|A_k)}{\sum_{i=1}^{n}P(A_i)P(B|A_i)}$$

注 在贝叶斯公式中,$P(A_i)$ 和 $P(A_i|B)$ 分别称为原因的先验概率和后验概率.

$P(A_i)(i=1,2,...,n)$ 是在没有进一步信息 (不知道事件 B 是否发生) 的情况下, 人们对诸事件发生可能性大小的认识.

当有了新的信息 (知道 B 发生), 人们对诸事件发生可能性大小 $P(A_i|B)$ 有了新的估计. 贝叶斯公式从数量上刻划了这种变化.

问题 1.4 商店论箱出售玻璃杯, 每箱 20 只, 其中每箱含 0,1,2 只次品的概率分别为 0.8, 0.1,0.1, 某顾客选中一箱, 从中任选 4 只检查, 结果都是好的, 便买下了这一箱. 问这一箱不含次品的概率是多少?

解 设 A_i= "任取一箱, 含 i 只次品" (i=1,2,3) B= "从一箱中任取 4 只检查, 结果都是好的"

$$P(B|A_0)=1$$

$$P(B|A_1)=\frac{C_{19}^4}{C_{20}^4}=\frac{4}{5}$$

$$P(B|A_2)=\frac{C_{18}^4}{C_{20}^4}=\frac{12}{19}$$

由 Bayes 公式

$$P(A_0|B) = \frac{P(A_0)P(B|A_0)}{\sum_{i=0}^{2} P(A_i)P(B|A_i)} \approx 0.848$$

问题 1.5 设甲袋中有 2 个白球,1 个红球, 乙袋中有 2 个红球,1 个白球, 手感上均无区别. 今从甲袋任取一球放入乙袋, 搅匀后再从乙袋任取一球, 求此球是红球的概率; 若已知取到红球, 求从甲袋放入乙袋的是白球的概率.

解 设 A=“从甲袋放入乙袋的是白球”, B=“从乙袋中任取一球是红球”

$$P(B) = P(A)P(B|A) + P(\overline{A})P(B|\overline{A}) = \frac{7}{12}$$

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} = \frac{4}{7}$$

第 2 章 随机变量及其分布

定义 2.1 (随机变量)

设 Ω 是试验 T 的样本空间, 若 $\forall\omega\in\Omega$ 可以通过一定的映射映到实数 (区间), 则称实值函数 $X(\omega)$ 为 Ω 上的随机变量，简记 **r.v.X**

注 r.v. 一般用大写字母 X,Y,Z,... 或小写希腊字母 $\zeta,\eta,\ldots$ 表示

命题 2.1 (随机变量的特点)

1. 随机性：r.v.X 的取值不止一个，试验前只能预知它的可能取值，但不能预知具体取哪一个值
2. X 以一定的概率取某一个值
3. 引入 r.v. 后可以用 r.v. 的等式或者不等式来表达随机事件
4. r.v. 的函数一般也是 r.v.
5. 在同一个样本空间可以同时定义多个 r.v. 各 r.v. 之间可能有一定的关系, 也可能没有关系—相互独立

2.1 离散随机变量

2.1.1 离散随机变量的概念

定义 2.2 (离散随机变量)

若随机变量 X 的可能取值是有限个或者可列个，则称 X 为离散型随机变量，称 $P(X=x_k)=p_k$ 为 r.v.X 的概率分布或者分布律

命题 2.2 (分布律的特征性质)

1. 非负性：$p_k\geq 0$
2. 规范性：$\sum_{k=1}^{\infty}p_k=1$

2.1.2 常见离散随机变量的分布

定义 2.3 (退化分布)

若随机变量 X 取常数 C 的概率为 1，$P(X=C)=1$, 这时称 X 服从退化分布或者单点分布

定义 2.4 (Bernoulli 分布)

若 $P(X=x)=p^x(1-p)^{1-x}$,x=0,1, 则称 r.v.X 的分布服从参数为 p 的 **0-1** 分布或两点分布或 **Bernoulli** 分布，记作 **X~ B(1,p)**

定义 2.5 (二项分布)

n 重 Bernoulli 试验中，X 是事件 A 在 n 次试验中发生的次数，$P(A)=p$，若 $b(k;n,p)=P(X=k)=C_n^kp^k(1-p)^{n-k}$, 则称 X 服从参数为 n，p 的二项分布，记作 **X~B(n,p)**

命题 2.3

1. 0-1 分布是 n=1 的二项分布
2. 若 $X_i\sim B(1,p)$, 则 $X=\sum_{i=1}^{n}X_i\sim B(n,p)$

3. $X \sim B(n,p) \Leftrightarrow$ X 可表示为 n 个独立服从 0-1 分布的随机变量之和.

定义 2.6 (Poisson 分布)

若 $P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda}$, 其中 $\lambda \geq 0$，则称 X 服从参数为 λ 的 Poisson 分布，记作 $\mathbf{X \sim P}(\lambda)$

定理 2.1 (Possion 逼近定理)

设 $np \approx \lambda \geq 0$, 则对固定的 k 有

$$\lim_{n\to\infty} C_n^k p^k(1-p)^{n-k}=\frac{\lambda^k}{k!}e^{-\lambda}$$

注 若 $X \sim B(n,p)$, 则当 n 较大 (≥ 20),p 较小（≤ 0.05）,$np=\lambda$ 适中，则可用近似公式

$$C_n^k p^k(1-p)^{n-k} \approx \frac{\lambda^k}{k!}e^{-\lambda}$$

证明

$$\begin{aligned}
C_n^k p^k(1-p)^{n-k} &= \frac{n(n-1)\cdots(n-k+1)}{k!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} \\
&= \frac{\lambda^k}{k!}\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{-\frac{n}{\lambda}(-\lambda)\frac{n-k}{n}} \\
&\to \frac{\boldsymbol{\lambda^k}}{\boldsymbol{k!}}\boldsymbol{e^{-\lambda}}(\boldsymbol{k=0,1,2,\cdots})
\end{aligned}$$

注

$$ln(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+o\left(x^4\right)$$

$$\frac{1}{x}\ln(1+x)=1-\frac{x}{2}+\frac{x^2}{3}-\frac{x^3}{4}+o\left(x^3\right)$$

$$e^{\frac{1}{x}\ln(1+x)}=e^{1-\frac{x}{2}+\frac{x^2}{3}-\frac{x^3}{4}+o\left(x^3\right)} \approx e \quad (x\to 0)$$

$$e^{-\frac{n}{\lambda}\ln\left(1-\frac{\lambda}{n}\right)}=e^{1+\frac{\lambda}{2n}+\frac{\lambda^2}{3n^2}+o\left(\left(\frac{\lambda}{n}\right)^3\right)} \approx e \quad (n\to\infty)$$

$$\left(1-\frac{\lambda}{n}\right)^{-\frac{n}{\lambda}(-\lambda)\frac{n-k}{n}} \approx \left(1-\frac{\lambda}{n}\right)^{-\frac{n}{\lambda}(-\lambda)} \approx e^{-\lambda}$$

Poisson 定理表明，Poisson 分布是二项分布的极限分布，当 n 很大，p 很小时，二项分布就可以近似地看成是参数 $\lambda=np$ 的 Possion 分布

问题 2.1 某厂产品不合格率为 0.03, 现将产品装箱, 若要以不小于 **90**% 的概率保证每箱中至少有 100 个合格品, 则每箱至少应装多少个产品?

解 设每箱至少应装 100+n 个, 每箱的不合格品个数为 X, 则 $X \sim \mathrm{B}(100+n, 0.03)$ 由题意 $P\{X \leq n\}=\sum_{k=0}^{n} b(k;100+n,0.03) \geq 0.9$；$(100+n)0.03=3+0.03n \approx 3$　取 $\lambda=3$

应用 Poisson 定理

$$\sum_{k=0}^{n} b(k;100+n,0.03) \approx \sum_{k=0}^{n}\frac{3^k}{k!}e^{-3}=1-\sum_{k=n+1}^{\infty}\frac{3^k}{k!}e^{-3} \geq 0.9$$

$$\Longrightarrow \sum_{k=n+1}^{\infty}\frac{3^k}{k!}e^{-3} \leq 0.1$$

查 Poisson 分布表, $\lambda=3$,$n+1=6$,$n=5$

故每箱至少装 105 个产品才能符合要求

2.2 随机变量的分布函数

定义 2.7 (分布函数的定义)

设 X 是 r.v.,x 是任意实数，称函数 $F(x)=P\{X\leq x\},-\infty<x<+\infty$ 为 X 的分布函数，也记作 $F_X(x)$
用分布函数计算 X 落在 (a,b] 里的概率：

$$\begin{aligned}P\{a<X\leq b\}&=P\{X\leq b\}-P\{X\leq a\}\\&=F(b)-F(a)\end{aligned}$$

注 因此只要知道了 r.v.X 的分布函数，他的统计特性就可以得到全面的描述.

命题 2.4 (分布函数的特征性质)

1. 非降性：F(x) 单调不减，即 $\forall x_1\leq x_2,F(x_1)\leq F(x_2)$
2. 有界性：$0\leq F(x)\leq 1$

$$\lim_{x\to-\infty}F(x)=F(-\infty)=F(X<-\infty)=0$$

$$\lim_{x\to+\infty}F(x)=F(+\infty)=F(X<+\infty)=1$$

3. 右连续性：F(x) 右连续，即 $\lim_{x\to x_0^+}F(x)=F(x_0)$

注 任一函数 F(x) 为分布函数的充要条件是：F(x) 满足上述三条性质

定理 2.2

一般地，若离散型随机变量 X 的分布律为 $P\{X=x_k\}=p_k$, 则其分布函数为 $F(x)=P\{X\leq x\}=\sum_{x_i\leq x}p_i$

注 特别地，若随机变量以概率 1 取常数, 即

$$P\{X=c\}=1$$

则称这个分布为单点分布或退化分布，它的分布函数为

$$F(x)=\begin{cases}0 & x<c\\1 & x\geq c\end{cases}$$

2.3 连续型随机变量

定义 2.8 (连续型随机变量的密度函数)

设 X 是随机变量, F(x) 是它的分布函数. 若存在一个非负可积函数 f(x)($\infty<x<\infty$), 使得

$$F(x)=P\{X\leq x\}=\int_{-\infty}^{x}f(u)du$$

则称 X 是连续型 r.v，f(x) 是它的概率密度函数 (p.d.f)
由定义可知, 连续型随机变量的分布函数是连续函数, 是密度函数的变上限的定积分. 在 $f(x)$ 的连续点

$$\frac{dF(x)}{dx}=f(x)$$

命题 2.5 (密度函数的特征性质)

1. 非负性：$f(x)\geq 0$
2. 规范性：$\int_{-\infty}^{+\infty}f(x)dx=1$

注 改变概率密度函数 f(x) 在个别点的函数值不影响公式 2 规范性, 故对固定的分布函数, 概率密度函数不唯一

注 满足上述两条性质的函数必是某一随机变量的密度函数. 故常利用这两个性质检验一个函数能否作为连续性 r.v. 的 p.d. f. (求 f(x) 中未知参数!)

命题 2.6 (分布函数的特征性质)

设连续型 r.v. X 的分布函数 (c.d.f.) 为 $F(x)$, 概率分布密度函数为 $f(x)$, 则

1. $F(x)$ 为连续函数; (求 $F(x)$ 中未知参数!)
2. 若 x 是 $f(x)$ 的连续点, 则 $\frac{dF(x)}{dx} = f(x)$
3. $P\{a \le X \le b\} = P\{a < X \le b\} = P\{a \le X < b\} = P\{a < X < b\} = F(b) - F(a) = \int_a^b f(x)dx$
4. 对任意实数 c, 则 $P\{X = c\} = 0$. 因为

$$P\{X = c\} = \lim_{\Delta x \to 0} P\{c \le X < c + \Delta x\} = \lim_{\Delta x \to 0} \int_c^{c+\Delta x} f(x)dx = 0$$

可见, 密度函数全面描述了连续型随机变量的规律.

注 由 $P(A) = 0$ 不能推出 $A = \varnothing$; 由 $P(B) = 1$ 不能推出 $B = \Omega$.

注 当 Δx 很小时, $P\{x < X < x + \Delta x\} = F(x + \Delta x) - F(x) \approx f(x)\Delta x$

密度函数值 f(a) 并不反映 X 取 a 值的概率. 但这个值越大,X 取 a 附近值的概率就越大. 也可以说, 在某点密度曲线的高度, 反映了概率集中在该点附近的程度.

定义 2.9 (均匀分布)

若 r.v.X 的 p.d.f. 为

$$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & other \end{cases}$$

则称 X 在 (a,b) 内服从均匀分布, 记作 $X \sim U(a, b)$

注 r.v.X 落在 (a,b) 区间上任一点的可能性都相同

命题 2.7 (均匀分布的特性)

1. X 服从均匀分布 U(a, b) 的充分必要条件是
   (1)X 落在 (a, b) 概率为 1, 落在区间外的概率为 0;
   (2)X 落在 (a, b) 子区间上概率与子区间长度成正比.
2. 均匀分布的分布函数:

$$F(x) = \begin{cases} 0 & x \le a \\ \frac{x-a}{b-a} & a < x \le b \\ 1 & x > b \end{cases}$$

定义 2.10 (指数分布)

若 r.v.X 的 p.d.f. 为

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

则称 X 服从参数为 $\lambda > 0$ 的指数分布. 其分布函数:

$$F(x) = P\{X \le x\} = \begin{cases} 0 & x \le 0 \\ \int_0^x \lambda e^{-\lambda x}dx & x > 0 \end{cases}$$

$$\mathrm{F(x)} = \begin{cases} 0 & x \le 0 \\ 1 - e^{-\lambda x} & x > 0 \end{cases}$$

注 非负的连续型 r.v.X 服从指数分布的充分必要条件是: 无记忆性

练习 2.1 某公路桥每天第一辆汽车过桥时刻为 T，设 [0，t] 时段内过桥的汽车数 X_t 服从参数为 λt 的泊松分布，求 T 的概率密度

解

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

$$F(t) = P\{T \le t\}$$

当 $t \le 0$ 时, $F(t) = 0$

当 $t > 0$ 时, $F(t) = P\{T \le t\} = 1 - P\{T > t\} = 1 - \{$ 在 t 时刻之前无汽车过桥 $\}$

$$= 1 - P\{X_t = 0\} = 1 - e^{-\lambda \cdot t}$$

$$f(t) = F'(t) = \begin{cases} \lambda e^{-\lambda t} & t > 0 \\ 0 & t \le 0 \end{cases}$$

定义 2.11 (Gamma 分布)

若 r.v.X 的 p.d.f 为

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} (x > 0)$$

$$\text{则称} X \sim \Gamma(\lambda, \alpha) \quad \text{其中} \Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx$$

$$\text{说明:} \begin{cases} \Gamma(\alpha+1) = \Gamma(\alpha)\alpha \\ \Gamma(n) = (n-1)! \\ \Gamma(1/2) = \sqrt{\pi} \end{cases}$$

注 $\Gamma(\lambda, 1) = Exp(\lambda) \Longrightarrow \alpha = 1$ 的 Γ 分布即为参数为 λ 的指数分布 $E(\lambda)$

定义 2.12 (正态分布)

若 r.v.X 的 p.d.f 为:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

则称 X 服从参数为 (μ, σ^2) 的正态分布，记作 $X \sim N(\mu, \sigma^2)$, 其中, μ 是均值, σ 是标准差。

正态分布的图像如下所示:

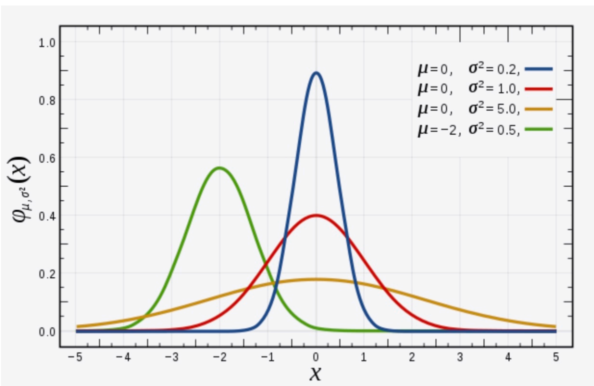


图 2.1: 正态分布

命题 2.8 (正态分布 p.d.f 的特点)

1. 关于直线 $x=\mu$ 对称
2. 在 $x=\mu$ 时, f (x) 取得最大值 $\frac{1}{\sqrt{2\pi}\sigma}$
3. f(x) 在 $x=\mu\pm\sigma$ 对应点处有拐点
4. f(x) 以 x 轴为渐近线

注

- μ 决定图形的中心位置，固定 σ, 图形形状不变, μ 改变, 图形平移.
- σ 决定随机变量取值的分散程度, 固定 μ, 图形由 σ 确定:
  (1) σ 越大, 图形越扁平,X 落在 μ 附近概率越小, 即取值越分散;
  (2) σ 越小, 图形越尖峭,X 落在 μ 附近概率越大, 即取值越集中.

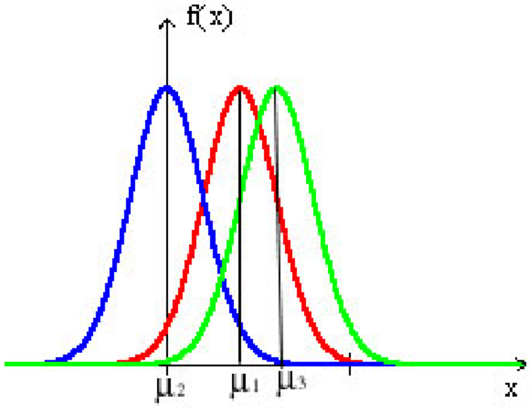


图 2.2: 平移

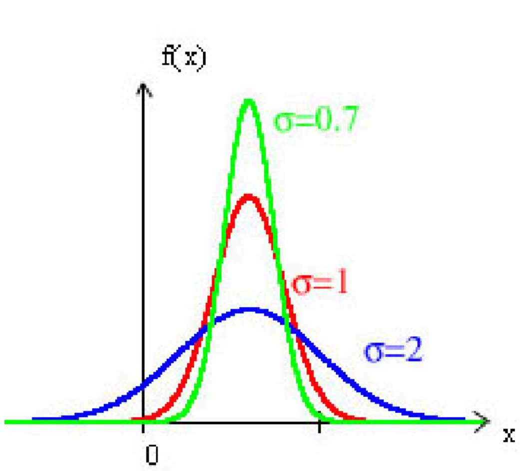


图 2.3: 陡峭程度

定义 2.13 (标准正态分布)

当 μ=0,σ=1 时，称 X 服从标准正态分布, 记为 $X\sim N(0,1)$, 此时，p.d.f 为偶函数:

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\quad -\infty<x<+\infty$$

分布函数:

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt\quad -\infty<x<+\infty$$

命题 2.9 (标准正态分布的性质)

若 $X\sim N(0,1)$，则

1. $\Phi(0)=0.5$
2. $\Phi(-x)=1-\Phi(x)$
3. $P(|X|<a)=2\Phi(a)-1$

命题 2.10 (正态分布标准化过程)

对一般的正态分布:$X\sim N(\mu,\sigma^2)$ 其分布函数

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

作变量代换 $s=\frac{t-\mu}{\sigma}$

$$F(x)=\Phi\left(\frac{x-\mu}{\sigma}\right)$$

即: $Y=\frac{X-\mu}{\sigma}\sim N(0,1)$

$$P(a<X<b)=F(b)-F(a)=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$P(X>a)=1-F(a)=1-\Phi\left(\frac{a-\mu}{\sigma}\right)$$

命题 2.11 (3σ 原则)

对于一般的正态分布 $X\sim N(\mu,\sigma^2)$

$$P(|X-\mu|\le\sigma)=0.6826$$

$$P(|X|-\mu\le 2\sigma)=0.9544$$

$$P(|X|-\mu\le 3\sigma)=0.9977$$

可以认为，Y 的取值几乎全部集中在 $[\mu-3\sigma,\mu+3\sigma]$ 区间内，这在统计学上称作 3σ（三倍标准差）原则

练习 2.2 一种电子元件的使用时间 X (小时) 服从正态分布 $(100,15^2)$, 某仪器上装有 3 个这种元件, 三个元件损坏与否是相互独立的. 求使用的最初 90 小时内无一元件损坏的概率.

解 设 Y 为使用的最初 90 小时内损坏的元件数, 则 $Y\sim B(3,p)$, 其中

$$p=P\{X<90\}=\Phi(\frac{90-100}{15})\approx\Phi(-0.67)=0.2514$$

故

$$P\{Y=0\}=(1-p)^3\approx 0.4159$$

2.4 随机变量函数的分布

命题 2.12 (离散型随机变量函数的分布律)

一般，若 X 是离散型 r.v.，X 的概率分布为

$$X\sim\begin{Bmatrix} x_1 & x_2\cdots & x_n \\ p_1 & p_2\cdots & p_n\end{Bmatrix}$$

$$\text{则 } Y=g(X)\sim\begin{Bmatrix} g(x_1) & g(x_2) & \cdots & g(x_n) \\ p_1 & p_2 & \cdots & p_n\end{Bmatrix}$$

注 若 $g(x_k)$ 中有一些是相同的, 则将它们作并项.

命题 2.13 (连续型随机变量函数的分布)

已知 X 的 p.d.f. p(x) 或分布函数, 求 Y = g(X) 的 p.d.f.

- 分布函数法

$$F_Y(y)=P\{Y\le y\}=P\{g(X)\le y\}=\int_{g(x)\le y}f(x)dx$$

$$f_Y(y)=\frac{dF_Y(y)}{dy}$$

- 公式法

  若 $X \sim f_X(x)$,$y = g(x)$ 是单调可导函数，记 $x = h(y)y = g(x)$ 的反函数，则

$$f_Y(y) = f_X[h(y)]|h'(y)|$$

注 注意定义域的选择

本质：$f_Y(y) = f_X(x)|\frac{dx}{dy}|$ i.e.$|f_Y(y)dy| = |f_X(x)dx|$

第 3 章 多维随机变量及其分布

3.1 二维随机变量

定义 3.1 (二维随机变量)

设 Ω 为随机试验的样本空间，若 $\forall\omega\in\Omega,\exists$ 一个法则使之映射到 $(X(\omega),Y(\omega))\in\mathbb{R}$, 则称 (X,Y) 为二维随机变量

定义 3.2 (联合分布函数)

$$F(x,y)=P< X<x,Y<y \qquad -\infty<x,y<+\infty$$

注 X 和 Y 的联合分布函数实际上是 $(X\le x)\cap(Y\le y)$ 的概率

命题 3.1 (联合分布函数的性质)

- 非降性：F(x,y) 对每个变量单调不减
- 有界性：$\forall x,y\in\mathbb{R}$ 有 $0\le F(x,y)\le 1$
  $F(-\infty,y)=F(x,-\infty)=0 \quad F(-\infty,-\infty)=0 \quad F(+\infty,+\infty)=1$
- 右连续性：$F(x_0,y_0)=F(x_0+0,y_0) \quad F(x_0,y_0)=F(x_0,y_0+0)$
- 对于任意 $a<b,c<d$,$F(b,d)-F(b,c)-F(a,d)+F(a,c)=P\{a<X\le b,c<Y\le d\}\ge 0$

注 对于二维随机变量 $P\{X>a,Y>c\}\neq 1-F(a,c)$

$$P\{X>a,Y>c\}=P\{a<X<+\infty,c<Y<+\infty\}=1-F(+\infty,c)-F(a,+\infty)+F(a,c)$$

定义 3.3 (二维离散型 r.v.)

若二维 r.v.(X ,Y) 所有可能的取值为有限多个或无穷可列多个, 则称 (X ,Y) 为二维离散型 r.v.

定义 3.4 (二维离散 r.v. 的联合分布函数)

设二维离散型随机向量 (X,Y) 的分布律为 p_{ij},i=1,2, ... ; j=1,2, ... 于是 (X,Y) 的分布函数

$$\begin{aligned}F(x,y)=P\{X\le x,Y\le y\}&=P\left(\bigcup_{x_i\le x,y_j\le y}\{X=x_i,Y=y_j\}\right)\\&=\sum_{x_i\le x}\sum_{y_j\le y}P\{X=x_i,Y=y_j\}\\&=\sum_{x_i\le x}\sum_{y_j\le y}p_{ij}\end{aligned}$$

定义 3.5 (二维连续 r.v. 及其概率特性)

设二维 r.v.(X ,Y) 的分布函数为 F(x ,y), 若存在非负可积函数 f (x,y) , 使得对于任意实数 x , y 有

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)dvdu$$

则称 (X ,Y) 为二维连续型 r.v. f (x,y) 为 (X ,Y) 的联合概率密度函数简称概率密度函数简记 p.d.f.

命题 3.2 (联合密度与联合分布函数的性质)

1. $f(x,y) \geq 0$
2. $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dydx = 1$
3. 在 f(x,y) 的连续点处 $\frac{\partial^2 F}{\partial x \partial y} = f(x,y)$
4. $P(X=a, Y=b) = 0$
   $P(X=a, -\infty < Y < +\infty) = 0$
   $P(-\infty < X < +\infty, Y=a) = 0$
5. 若 G 是平面上的区域，则 $P((X,Y) \in G) = \iint_G f(x,y)dxdy$

命题 3.3 (均匀分布)

设 G 是平面上的有界区域, 面积为 S, 若 r.v.(X ,Y) 的联合 p.d.f. 为

$$f(x,y) = \begin{cases} \frac{1}{S}, & (x,y) \in G \\ 0, & others \end{cases}$$

则称 (X,Y) 服从区域 G 上的均匀分布, 记 U(G). 若 (X,Y) 服从 G 上的均匀分布，其特征性质为

1. (X,Y) 落在 G 中某一区域 A 内的概率 $P\{(X,Y) \in A\}$ 与 A 的面积成正比而与 A 的位置和形状无关.
2. $P\{(X,Y) \in A\} = S_A/S$

命题 3.4 (二维正态分布)

若 r.v.(X ,Y) 的联合概率密度函数为:

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_x)^2}{\sigma_x^2} - 2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2}\right]\right)$$

其中 X 和 Y 是二维正态分布的随机变量，μ_x 和 μ_y 是 X 和 Y 的均值，σ_x 和 σ_y 是 X 和 Y 的标准差，ρ 是 X 和 Y 的相关系数。$\sigma_x, \sigma_y > 0, -1 < \rho < 1$

则称 (X ,Y) 服从参数为 $\mu 1, \sigma_1^2, \mu_2, \sigma_2^2, \rho$ 的正态分布, 记作 $(X,Y) \sim N(\mu 1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$

注 正态分布的边缘分布仍为正态分布

$$F_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left(-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right)$$

$$F_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(y-\mu_y)^2}{2\sigma_y^2}\right)$$

3.2 边缘分布

定义 3.6 (边缘分布函数)

设 $F(x,y)$ 为随机变量 (X,Y) 的分布函数, 则 $F(x,y) = P\{X \leq x, Y \leq y\}$.

令 $y \to \infty$, 称

$$P\{X \leq x\} = P\{X \leq x, Y < \infty\} = F(x,\infty)$$

为随机变量 (X,Y) 关于 X 的边缘分布函数. 记为 $F_X(x) = F(x,\infty)$.

同理令 $\boldsymbol{x} \to \infty$,

$$F_Y(y) = F(\infty,y) = P\{X < \infty, Y \leq y\} = P\{Y \leq y\}$$

定义 3.7 (二维离散型 r.v. 的边缘分布)

若随机变量 X 与 Y 的联合分布律为

$$P\left(X=x_i, Y=y_j\right)=p_{ij}, i=1,2,\cdots, j=1,2,\cdots$$

则 (X, Y) 关于 X、Y 的边缘分布律分别为:

$$P\left\{X=x_i\right\}=\sum_{j=1}^{+\infty} p_{ij}=p_{i\cdot}, \boldsymbol{i}=1,2,\cdots$$

$$P\left\{Y=y_j\right\}=\sum_{i=1}^{+\infty} p_{ij}=p_{\cdot j}, \boldsymbol{j}=\mathbf{1},2,\cdots$$

由边缘分布律易求边缘分布函数

$$F_X(\boldsymbol{x})=\boldsymbol{F}(x,+\infty)=\sum_{x_i \leq x} p_{i\cdot}=\sum_{x_i \leq x} \sum_{j=1}^{+\infty} p_{ij}$$

♣

定义 3.8 (二维连续型 r.v. 边缘分布)

设 $(X, Y) \sim f(x, y),(x, y) \in R^2, F(x, y)$ 为分布函数

则 (X, Y) 关于 X 的边缘密度函数为:

$$f_X(x)=\int_{-\infty}^{+\infty} f(x, y) d y$$

(X, Y) 关于 Y 的边缘密度函数为:

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x, y) d x$$

由边缘密度函数易求边缘分布函数

$$F_X(x)=F(x,+\infty)=\int_{-\infty}^x f_X(u) d u=\int_{-\infty}^x d u \int_{-\infty}^{+\infty} f(u, v) d v$$

♣

3.3 二维随机变量的条件分布

命题 3.5 (离散型随机变量的条件分布律)

设 $(X, Y) \sim P\left\{X=x_i, Y=y_j\right\}=p_{ij},(i, j=1,2, \cdots), X$ 和 Y 的边缘分布律分别为

$$\begin{aligned} P\left\{X=x_i\right\}=p_{i\cdot}=\textstyle\sum_{j \geq 1} p_{ij} \quad & i=1,2, \ldots \\ P\left\{Y=y_j\right\}=p_{\cdot j}=\textstyle\sum_{i \geq 1} p_{ij} \quad & j=1,2, \ldots \end{aligned}$$

对固定的 j, $p_{\cdot j}>0$, 在 $Y=y_j$ 的条件下,X 的条件分布律:

$$P\left\{X=x_i \mid Y=y_j\right\}=\frac{p_{ij}}{p_{\cdot j}}$$

对固定的 i, $p_{i\cdot}>0$, 在 $X=x_i$ 的条件下,Y 的条件分布律:

$$P\left\{Y=y_j \mid X=x_i\right\}=\frac{p_{ij}}{p_{i\cdot}}$$

♠

命题 3.6 (连续型随机变量的条件分布)

$Y=y$ 条件下关于 X 的条件密度函数

$$f_{X \mid Y}(x \mid y)=\frac{f(x, y)}{f_Y(y)}$$

$X = x$ 条件下关于 Y 的条件密度函数

$$f_{Y|X}(y \mid x) = \frac{f(x, y)}{f_X(x)}$$

从而

$$F_{X|Y}(x \mid y) = \int_{-\infty}^{x} \frac{f(u, y)}{f_Y(y)} du$$

$$F_{Y|X}(y \mid x) = \int_{-\infty}^{y} \frac{f(x, v)}{f_X(x)} dv$$

3.4 二维随机变量的独立性

定义 3.9 (二维随机变量独立性的定义)

如果对任意实数 $x, y, F(x, y) = F_X(x)F_Y(y)$, 即

$$P(X \le x, Y \le y) = P(X \le x) \cdot P(Y \le y)$$

称随机变量 X 与 Y 独立.

命题 3.7 (离散型 r.v. 的独立性)

离散型 r.v.(X,Y) 的概率分布为:

$$P\left(X = x_i, Y = y_j\right) = p_{ij}, i, j = 1, 2, \cdots$$

则 X, Y 独立的充分必要条件为

$$p_{ij} = p_{i\cdot} p_{\cdot j}, i, j = 1, 2, \cdots$$

即

$$P\left(X = x_i, Y = y_j\right) = P\left(X = x_i\right) \cdot P\left(Y = y_j\right)$$

命题 3.8 (连续型 r.v. 的独立性)

设 (X,Y) 是二维连续型随机变量，X 与 Y 独立的充分必要条件是

$$f(x, y) = f_X(x) f_Y(y)$$

推论 3.1

设 (X,Y) 是连续型随机变量，f (x, y) 为 (X, Y) 的概率密度函数, 则随机变量 X, Y 独立的充分必要条件为

$$f(x, y) = h(x)g(y)$$

3.5 二维随机变量函数的分布

3.5.1 二维离散型随机变量函数的分布

命题 3.9

设二维离散型随机变量 (X, Y) 的分布为 $P(X = x_i, Y = y_j) = p_{ij}, i, j = 1, 2, Z = g(X, Y)$,
则 $P\{Z = z_k\} = p_k = \sum_{g(x_i, y_i) = z_k} p_{ij}$

3.5.2 二维连续型随机变量函数的分布

命题 3.10

利用分布函数的定义将 Z 的分布函数转化为 (X,Y) 的事件

$$F_Z(z) = P\{Z \le z\} = P\{g(X,Y) \le z\} = \iint_{g(x,y\le z)} f(x,y)dxdy$$

然后对分布函数求导

定理 3.1 (和的分布：Z = X + Y)

若 X ,Y 相互独立, 则

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)dx = f_X(z)f_Y(z)$$

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y)f_Y(y)dy = f_X(z)f_Y(z)$$

称之为函数 $\mathrm{f}_X(z)f_Y(z)$

命题 3.11 (正态随机变量的结论)

若 X ,Y 相互独立， $X \sim N(\mu_1,\sigma_1^2), Y \sim N(\mu_2,\sigma_2^2)$, 则 $X+Y \sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$

推论 3.2

$X_1, X_2, \cdots, X_n$ 相互独立 $X_i \sim N\left(\mu_i,\sigma_i^2\right), i = 1,2,\cdots,n, c_i$ 是常数，则 $\sum_{i=1}^n c_iX_i \sim N\left(\sum_{i=1}^n c_i\mu_i, \sum_{i=1}^n c_i^2\sigma_i^2\right)$

注 正态随机变量的线性组合仍服从正态分布

第 4 章 随机变量的特征数

4.1 数学期望

定义 4.1 (数学期望的定义)

设 X 为离散 r.v. 其分布为 $P(X=x_k)=p_k, k=1,2...$, 若无穷级数 $\sum_{k=1}^{\infty} x_k p_k$ 绝对收敛，则称其和为 X 的数学期望，记作 $E(X)$, 即 $E(X)=\sum_{k=1}^{\infty} x_k p_k$

命题 4.1

- 若 $X \sim B(n,p)$, 则 $E(X)=np$
- 若 $X \sim P(\lambda)$, 则 $E(X)=\lambda$

定义 4.2 (连续型 r.v. 数学期望)

设连续 r.v. X 的 p.d.f. 为 f(x) 若广义积分 $\int_{-\infty}^{+\infty} x f(x) dx$ 绝对收敛，则称此积分为 X 的数学期望，记作 E(X)，即

$$E(X)=\int_{-\infty}^{+\infty} x f(x) dx$$

命题 4.2

- 若 $X \sim U(a,b)$, 则 $E(X)=\frac{a+b}{2}$.
- 若 $X \sim Exp(\lambda)$, 则 $E(X)=\frac{1}{\lambda}$.
- 若 $X \sim N(\mu\sigma^2)$, 则 $E(X)=\mu$

注 不是所有的 r.v. 都有数学期望

例如：柯西 (Cauchy) 分布的密度函数为

$$f(x)=\frac{1}{\pi(1+x^2)} \quad ,-\infty<x<+\infty$$

但是 $\int_{-\infty}^{+\infty} |x| f(x) dx=\int_{-\infty}^{+\infty} \frac{|x|}{\pi(1+x^2)} dx$ 发散，故数学期望不存在

命题 4.3 (离散 r.v. 函数的数学期望)

设离散 r.v. X 的概率分布为 $(X)=\sum_{k=1}^{+\infty} x_k p_k$, 若无穷级数 $\sum_{i=1}^{+\infty} g(x_i) p_i$ 绝对收敛，则 $E(Y)=E[g(x)]=\sum_{i=1}^{+\infty} g(x_i) p_i$

命题 4.4 (连续 r.v. 函数的数学期望)

设连续 r.v. 的 p.d.f. 为 f(x), 若广义积分 $\int_{-\infty}^{+\infty} g(x) f(x) dx$ 绝对收敛，则 $E(Y)=\int_{-\infty}^{+\infty} g(x) f(x) dx$

命题 4.5 (二元离散 r.v. 函数的数学期望)

设离散 r.v.(X,Y) 的概率分布为 $P(X=x_i, Y=y_j)=p_{ij}$, 若级数 $\sum_{i,j=1}^{+\infty} g(x_i,y_j) p_{ij}$ 绝对收敛，则 $E[g(X,Y)]=\sum_{i,j=1}^{+\infty} g(x_i,y_j) p_{ij}$

命题 4.6 (二元连续 r.v. 函数的数学期望)

设连续 r.v.(X,Y) 的联合 p.d.f. 为 f(x，y), 若广义积分 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y) f(x,y) dxdy$ 绝对收敛，则 $E[g(X,Y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y) f(x,y) dxdy$

定理 4.1 (数学期望的性质)

1. $E(C) = C$
2. $E(aX) = aE(X)$
3. $E(X+Y) = E(X) + E(Y)$
4. $E\left(\sum_{i=1}^{n} c_i X_i + C\right) = \sum_{i=1}^{n} c_i E(X_i) + C$
5. 当 X,Y 独立时，$E(XY) = E(X)E(Y)$. 逆命题不一定成立

定理 4.2 (概率积分)

$$\begin{aligned}\left(\int_{-\infty}^{+\infty} e^{-x^2} dx\right)^2 &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-(x^2+y^2)} dxdy \\ &= 4\int_{0}^{+\infty}\int_{0}^{+\infty} e^{-(x^2+y^2)} dxdy \\ &= 4\int_{0}^{\frac{\pi}{2}} d\theta \int_{0}^{+\infty} e^{-r^2} rdr \\ &= 4\times\frac{\pi}{2}\times\frac{1}{2} \\ &= \pi\end{aligned}$$

4.2 方差

定义 4.3 (方差的定义)

若 $E[X-E(X)]^2$ 存在, 则称其为随机变量 X 的方差, 记为 $D(X)$ 或 $Var(X)$, 即

$$D(X) = E[X-E(X)]^2$$

称 $\sqrt{D(X)}$ 为 X 的均方差或标准差

注 D(X) ——描述 r.v. X 的取值偏离平均值的平均偏离程度

命题 4.7 (离散型随机变量的方差)

若 X 为离散型 r.v.，分布律为 $P(X = x_k) = p_k$, 则

$$D(X) = \sum_{k=1}^{+\infty}(x_k - E(X))^2 p_k$$

命题 4.8 (连续型随机变量的方差)

若 X 为连续型 r.v. , 概率密度为 f (x)

$$D(X) = \int_{-\infty}^{+\infty}(x - E(x))^2 f(x)dx$$

定理 4.3 (重要公式)

$$D(X) = E(X^2) - [E(X)]^2$$

证明

$$
\begin{aligned}
D(X) &= E(X-E(X))^2 = E\{X^2 - 2XE(X) + [E(X)]^2\} \\
&= E(X^2) - 2E(X)E(X) + [E(X)]^2 \\
&= E(X^2) - [E(X)]^2
\end{aligned}
$$

命题 4.9

- 方差非负，即 $D(X) \geq 0$;
- E(X) 的取值相当于物理学上作一条直线，使所有的点均匀分布在直线的两边;
- D(X) 的取值相当于平均误差;
- D(X)=0 的充分必要条件为 r.v.X 的取值为常数.

命题 4.10

- $D(c) = 0$
- $D(cX) = c^2 D(X)$
- $D(c_1X \pm c_2Y) = c_1^2 D(X) + c_2^2 D(Y) \pm 2c_1c_2[E(XY) - E(X)\cdot E(Y)]$

  推论 4.1

  若 X,Y 相互独立，则 $D(c_1X \pm c_2Y) = c_1^2 D(X) + c_2^2 D(Y)$

- 对任意常数 C,$D(X) \leq E(X-C)^2$, 当且仅当 $C = E(X)$ 时等号成立
- $D(X) = 0 \Leftrightarrow P(X = E(X)) = 1$　称为 X 依概率 1 等于常数 E(X)
- 对任意常数 C,$D(X) \leq E(XC)^2$, 当且仅当 $C = E(X)$ 时等号成立

命题 4.11 (重要的方差)

- $X \sim B(n,p)$,　$D(X) = np(1-p)$
- $X \sim P(\lambda)$,　$D(X) = \lambda$
- $X \sim U(a,b)$　$D(X) = \frac{(b-a)^2}{12}$
- $X \sim EXP(\lambda)$,　$D(X) = \lambda^{-2}$

定义 4.4 (标准化随机变量)

设随机变量 X 的期望 E(X)、方差 D(X) 都存在, 且 $D(X) \neq 0$, 则称

$$
X^* = \frac{X - E(X)}{\sqrt{D(X)}}
$$

为 X 的标准化随机变量，显然 $E(X^*) = 0$,　$D(X^*) = 1$

注 仅知 r.v. 的期望与方差并不能确定其分布

4.3 协方差

定义 4.5

称 $E([X-E(X)][Y-E(Y)])$ 为 X,Y 的协方差. 记

$$
Cov(X,Y) = E([X-E(X)][Y-E(Y)])
$$

若 (X,Y) 为离散型

$$Cov(X,Y)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}[x_i-E(X)][y_j-E(Y)]p_{ij}$$

若 (X,Y) 为连续型

$$Cov(X,Y)=\int_{-\infty}^{+\infty]}\int_{-\infty}^{+\infty}[x-E(X)][y-E(Y)]f(x,y)dxdy$$

注

$$Cov(X,Y)=E(XY)-E(X)E(Y)$$

命题 4.12 (协方差的性质)

- $Cov(X,Y)=Cov(Y,X)$
- $Cov(X,X)=D(X);Cov(X,c)=0$
- $Cov(aX,bY)=abCov(X,Y)$, 其中 a, b 为常数
- $Cov(X+YZ)=Cov(X,Z)+Cov(Y,Z)$
- $D(X\pm Y)=DX+DY\pm 2Cov(X,Y)$

定义 4.6 (不相关)

当 $Cov(X,Y)=0$ 时，称 X 与 Y 不相关

注 “X 与 Y 独立” $\Longrightarrow$ “X 与 Y 不相关”，反之未必成立

命题 4.13

X 与 Y 为随机变量, 则下列结果等价

- X,Y 不相关
- $Cov(X,Y)=0$
- $E(XY)=E(X)E(Y)$
- $D(X\pm Y)=D(X)+D(Y)$

4.4 相关系数

定义 4.7 (相关系数)

若随机变量 X，Y 的方差和协方差均存在, 且 $DX>0,DY>0$, 则

$$\rho=\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

称为 X 与 Y 的相关系数

注 $E(X)^*=0,D(X)^*=1$

$\rho_{XY}=Cov(X^*,Y^*)=E(X^*Y^*)$

命题 4.14

$$\begin{aligned}\rho_{XY}=0 &\Longleftrightarrow X,Y\text{ 不相关}\\ &\Longleftrightarrow Cov(X,Y)=0\\ &\Longleftrightarrow E(XY)=E(X)E(Y)\\ &\Longleftrightarrow D(X\pm Y)=D(X)+D(Y)\end{aligned}$$

命题 4.15 (相关系数的性质)

- $|\rho_{XY}| \leq 1$
- $|\rho_{XY} = 1 \Leftrightarrow$ 存在常数 a, b 使 $P\{Y = aX + b\} = 1$
- X 与 Y 不相关 $\Leftrightarrow \rho_{XY} = 0$