

概率论与数理统计

喵喵喵

2020 年 9 月 17 日

目录

1 概率论的基本概念	2
1.1 随机实验	2
1.2 样本空间、随机事件	2
1.3 频率与概率	2
1.4 等可能概型（古典概型）	5
1.5 条件概率	6
1.6 独立性	8
2 随机变量及其分布	9
2.1 随机变量	9
2.2 离散型随机变量及其分布律	9
2.2.1 (0-1) 分布	10
2.2.2 伯努利试验、二项分布	10
2.2.3 泊松分布	10
2.3 随机变量的分布函数	11
2.4 连续型随机变量及其概率密度	12
2.4.1 均匀分布	12
2.4.2 指数分布	12
2.4.3 正态分布	13
2.5 随机变量的函数的分布	14

3 多维随机变量及其分布	15
3.1 二维随机变量	15
3.2 边缘分布	16
3.3 条件分布	18
3.4 相互独立的随机变量	18
3.5 两个随机变量的函数的分布	19

1 概率论的基本概念

1.1 随机实验

1.2 样本空间、随机事件

将随机试验 E 的所有可能结果组成的集合称为 E 的 **样本空间**，记为 S ，样本空间的元素，即 E 的每个结果，称为 **样本点**

称试验 E 的样本空间 S 的子集为 E **随机事件**，简称 **事件**。在每次试验中，当且仅当这一子集中的一个样本点出现时，称这一 **事件发生**

由一个样本点组成的单点集称为 **基本事件**

样本空间 S 包含所有的样本点，他是 S 自身的子集，在每次试验中它总是出现， S 称为 **必然事件**，空集 \emptyset 称为 **不可能事件**

设试验 E 的样本空间为 S ，而 $A, B, A_k \subseteq S$

1. 若 $A \subset B$ ，则称事件 B 包含事件 A ，事件 A 的发生必导致事件 B 发生
若 $A \subset B, B \subset A$ 即 $A = B$ ，则称事件 A 与事件 B **相等**
2. 事件 $A \cup B = \{x \mid x \in A \text{ or } x \in B\}$ 称为事件 A 与事件 B 的 **和事件**，当且仅当 A, B 中至少由一个发生时，事件 $A \cup B$ 发生
3. 事件 $A \cap B = \{x \mid x \in A \text{ and } x \in B\}$ 称为事件 A 与事件 B 的 **积事件**，也记作 AB

称 $\bigcap_{k=1}^{\infty} A_k$ 为可列个事件 A_1, A_2, \dots 的积事件

4. 事件 $A - B = \{x \mid x \in A \text{ and } x \notin B\}$ 称为事件 A 与事件 B 的 **差事件**
5. 若 $A \cap B = \emptyset$ ，则称事件 A 与事件 B **互不相容**或 **互斥**的
6. 若 $A \cup B = S, A \cap B = \emptyset$ ，则称事件 A 与事件 B 互为 **逆事件**，又称事件 A 与事件 B 互为 **对立事件**

1.3 频率与概率

Definition 1.1. 在相同的条件下，进行了 n 次试验，在这 n 次试验中，事件 A 发生次数 n_A 称为事件 A 发生的 **频数**，比值 n_A/n 称为事件 A 发生的 **频率**，并记成 $f_n(A)$

基本性质

1. $0 \leq f_n(A) \leq 1$
2. $f_n(S) = 1$
3. 若 A_1, \dots, A_k 是两两互不相容的事件, 则

$$f_n(A_1 \cup \dots \cup A_k) = f_n(A_1) + \dots + f_n(A_k)$$

Definition 1.2. 设 E 是随机试验, S 是它的样本空间, 对于 E 的每一事件 A 赋予一个实数, 记为 $P(A)$, 称为事件 A 的 **概率**, 如果集合函数 $P(\cdot)$ 满足下列条件

1. **非负性**: 对于每一个事件 A , 有 $P(A) \geq 0$
2. **规范性**: 对于必然事件, 有 $P(A) = 1$
3. **可列可加性**: 设 A_1, A_2, \dots 是两两互不相容事件, 有

$$P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$$

Proposition 1.3. $P(\emptyset) = 0$

证明. 令 $A_n = \emptyset$, 则 $\bigcup_{n=1}^{\infty} A_n = \emptyset$, 且 $A_i A_j = \emptyset$, 由可列可加性

$$P(\emptyset) = P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} P(\emptyset)$$

由概率的非负性, $P(\emptyset) = 0$ □

Proposition 1.4 (有限可加性). 若 A_1, \dots, A_n 是两两互不相容的事件, 则有

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + \dots + P(A_n)$$

证明. 令 $A_{n+1} = A_{n+2} = \dots = \emptyset$, 即有 $A_i A_j = \emptyset$

$$\begin{aligned} P(A_1 \cup A_2 \cup \dots \cup A_n) \\ &= P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k) \\ &= \sum_{k=1}^{\infty} P(A_k) + 0 = P(A_1) + \dots + P(A_n) \end{aligned}$$

□

Proposition 1.5. 设 A, B 是两个事件, 若 $A \subset B$, 则有

$$P(B - A) = P(B) - P(A)$$

$$P(B) \geq P(A)$$

证明. $B = A \cup (B - A)$

□

Proposition 1.6. 对于任一事件 A

$$P(A) \leq 1$$

证明. $P(A) \leq P(S) = 1$

□

Proposition 1.7 (逆事件的概率). 对于任一事件 A , 有

$$P(\bar{A}) = 1 - P(A)$$

证明. $P(S) = P(A \cup \bar{A})$

□

Proposition 1.8 (加法公式). 对于任一两事件 A, B 有

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

证明. $A \cup B = A \cup (B - AB)$

$$P(A \cup B) = P(A) + P(B - AB) = P(A) + P(B) - P(AB)$$

□

可推广到

$$\begin{aligned} P(A_1 \cup A_2 \cup \cdots \cup A_n) &= \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i A_j) \\ &+ \sum_{1 \leq i < j < k \leq n} P(A_i A_j A_k) + \cdots + (-1)^{n-1} P(A_1 \cdots A_n) \end{aligned}$$

1.4 等可能概型 (古典概型)

等可能概型 (古典概型)

1. 试验的样本空间只包含有限个元素
2. 试验中每个基本事件发生的可能性相同

设试验的样本空间为 $S = \{e_1, \dots, e_n\}$, 由于在试验中每个基本事件发生的可能性相同, 即

$$P(\{e_1\}) = \dots = P(\{e_n\})$$

又由于基本事件两两互不相容, 于是

$$\begin{aligned} 1 &= P(S) = P(\{e_1\} \cup \dots \cup \{e_n\}) \\ &= P(\{e_1\}) + \dots + P(\{e_n\}) \\ &= nP(\{e_1\}) \end{aligned}$$

于是

$$P(\{e_i\}) = \frac{1}{n}$$

若事件 A 包含 k 个基本事件, 即 $A = \{e_{i_1}\} \cup \dots \cup \{e_{i_k}\}$, 则有

$$P(A) = \sum_{j=1}^k P(\{e_{i_j}\}) = \frac{k}{n}$$

Example 1.1. 设有 N 件产品, 其中有 D 件次品, 今从中任取 n 件, 文其中恰有 $k(k \leq D)$ 件次品的概率

$$p = \frac{\binom{D}{k} \binom{N-D}{n-k}}{\binom{N}{n}}$$

Example 1.2. 袋中有 a 只白球, b 只红球, k 个人依次在袋中取一只球, 求第 i 人取到白球 (记为事件 B) 的概率 ($k \leq a+b$)

共有 A_{a+b}^k 个基本事件, 事件 B 发生时, 第 i 人取的应是白球, 有 a 中取法, 剩余 $k-1$ 只球有 A_{a+b-1}^{k-1} 种取法, 则

$$P(B) = \frac{a \cdot A_{a+b-1}^{k-1}}{A_{a+b}^k} = \frac{a}{a+b}$$

1.5 条件概率

Definition 1.9. 设 A, B 是两个事件, 且 $P(A) > 0$, 称

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

为在事件 A 发生的条件下事件 B 发生的 **条件概率**

条件概率 $P(\cdot|A)$ 符合

1. **非负性**: 对于每一事件 B , 有 $P(B|A) \geq 0$
2. **规范性**: 对于必然事件 S , 有 $P(S|A) = 1$
3. **可列可加性**: 设 B_1, B_2, \dots 是两两互不相容的事件, 则有

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

Theorem 1.10 (乘法定理). 设 $P(A) > 0$, 则有

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

一般地, 设 A_1, \dots, A_n 为 n 个事件, $n \geq 2$, 且 $P(A_1 \dots A_{n-1}) > 0$, 则有

$$P(A_1 \dots A_n) = P(A_n | A_1 \dots A_{n-1})P(A_{n-1} | A_1 \dots A_{n-2}) \dots P(A_2 | A_1)P(A_1)$$

Example 1.3. 设袋中装有 r 只红球, t 只白球, 每次自袋中任取一只球, 观察其颜色再放回, 并再放入 a 只与所取出的那只球同色的球, 若在袋中连续取球四次, 试求第一、二次取到红球且第三、四次取到白球的概率

以 A_i 表示事件 “第 i 次取到红球”, 则

$$\begin{aligned} P(A_1 A_2 \overline{A_3} \overline{A_4}) &= P(\overline{A_4} | A_1 A_2 \overline{A_3}) P(\overline{A_3} | A_1 A_2) P(A_2 | A_1) P(A_1) \\ &= \frac{t+a}{r+t+3a} \cdot \frac{t}{r+t+2a} \cdot \frac{r+a}{r+t+a} \cdot \frac{r}{r+t} \end{aligned}$$

Definition 1.11. 设 S 为试验 E 的样本空间, B_1, \dots, B_n 为 E 的一组事件, 若

1. $B_i B_j = \emptyset, i \neq j, i, j = 1, 2, \dots, n$
2. $B_1 \cup B_2 \cup \dots \cup B_n = S$

则称 B_1, \dots, B_n 为样本空间 S 的一个 **划分**

Theorem 1.12. 设试验 E 的样本空间为 S , A 为 E 的事件, B_1, \dots, B_n 为 S 的一个划分, 且 $P(B_i) > 0$, 则

$$P(A) = P(A|B_1)P(B_1) + \dots + P(A|B_n)P(B_n)$$

称为 **全概率公式**

证明.

$$A = AS = A(B_1 \cup \dots \cup B_n) = AB_1 \cup \dots \cup AB_n$$

□

Theorem 1.13. 设试验 E 的样本空间 S , A 为 E 的事件, B_1, \dots, B_n 为 S 的一个划分, 且 $P(A) > 0, P(B_i) > 0$, 则

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^n P(A|B_j)P(B_j)}$$

称为 **贝叶斯公式**

特别的, 取 $n = 2$, 则

$$P(A) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B})$$

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

Example 1.4. 患肺癌的概率约为 0.1%, 在人群中 20% 是吸烟者, 他们患肺癌的概率约为 0.4%, 求不吸者患肺癌的概率

以 C 记事件 “患肺癌”, 以 A 记事件 “吸烟”, 则 $P(C) = 0.001, P(A) = 0.2, P(C|A) = 0.004$, 由全概率公式

$$P(C) = P(C|A)P(A) + P(C|\bar{A})P(\bar{A})$$

因此

$$P(C|\bar{A}) = 0.00025$$

1.6 独立性

Definition 1.14. 设 A, B 是两事件, 如果满足

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

则称事件 A, B **相互独立**, 简称 A, B **独立**

Theorem 1.15. 设 A, B 是两事件, 且 $P(A) > 0$, 若 A, B 相互独立, 则 $P(B|A) = P(B)$

Theorem 1.16. 若事件 A, B 相互独立, 则 A 与 \bar{B} , \bar{A} 与 B , \bar{A} 与 \bar{B} 也相互独立

Definition 1.17. 设 A, B, C 是三个事件, 满足

$$\begin{cases} P(AB) = P(A)P(B) \\ P(BC) = P(B)P(C) \\ P(AC) = P(A)P(C) \\ P(ABC) = P(A)P(B)P(C) \end{cases}$$

则称事件 A, B, C **相互独立**

一般地, 设 A_1, \dots, A_n , 如果对于 n 中任意 $2, 3, \dots, n$ 个事件的积事件的概率, 都等于各事件概率之积, 则称事件 A_1, \dots, A_n **相互独立**

Example 1.5. 要验收一批 (100) 件乐器, 验收方案如下: 自该批乐器中随机地取 3 件测试 (相互独立), 如果 3 件中至少有一件在测试中被认为音色不纯, 则这批乐器被拒绝接收。设一件音色不纯的乐器经测试查出其音色不纯的概率为 0.95, 而一件音色纯的乐器被误认为不纯的概率为 0.01, 已知 100 中有 4 件音色不纯, 试问这批乐器被接收的概率是多少

设以 H_i 表示 3 件中恰有 i 件不纯, A 表示这批批乐器被接收, 则

$$P(A|H_0) = 0.99^3, P(A|H_1) = 0.99^2 \times 0.01$$

$$P(A|H_2) = 0.99 \times 0.01^2, P(A|H_3) = 0.01^3$$

而

$$P(H_0) = \frac{\binom{96}{3}}{\binom{100}{3}}, P(H_1) = \frac{\binom{4}{1}\binom{96}{2}}{\binom{100}{3}}$$

$$P(H_2) = \frac{\binom{4}{2}\binom{96}{1}}{\binom{100}{3}}, P(H_3) = \frac{\binom{4}{3}}{\binom{100}{3}}$$

故

$$P(A) = \sum P(A|H_i)P(H_i)$$

2 随机变量及其分布

2.1 随机变量

Definition 2.1. 设随机试验的样本空间为 $S = \{e\}$, $X = X(e)$ 是定义在样本空间 S 上的实值单值函数, 称 $X = X(e)$ 为随机变量

2.2 离散型随机变量及其分布律

设离散型随机变量 X 所有可能取的值为 $x_k (k = 1, 2, \dots)$, X 取各个可能值的概率, 即事件 $\{X = x_k\}$ 的概率, 为

$$P\{X = x_k\} = p_k, k = 1, 2, \dots \quad (2.2.1)$$

由概率的定义, p_k 满足如下两个条件

1. $p_k \geq 0, k = 1, 2, \dots$
2. $\sum_{k=1}^{\infty} p_k = 1$

我们称 (2.2.1) 为离散型随机变量 X 的 **分布律**, 分布律也可以用表格表示

X	x_1	x_2	\dots	x_n	\dots
p_k	p_1	p_2	\dots	p_n	\dots

2.2.1 (0-1) 分布

设随机变量 X 只可能取 0 与 1 两个值, 它的分布律是

$$P\{X = k\} = p^k(1-p)^{1-k}, k = 0, 1 \quad (0 < p < 1)$$

则称 X 服从以 p 为参数的 (0-1) 分布或两点分布

2.2.2 伯努利试验、二项分布

设试验 E 只有两个可能结果: A 及 \bar{A} , 则称 E 为 **伯努利试验**。设 $P(A) = p$ ($0 < p < 1$), 此时 $P(\bar{A}) = 1 - p$ 。将 E 独立重复地进行 n 次, 则称这一串重复的独立试验为 **n 重伯努利试验**

以 X 表示 n 重伯努利事件 A 发生的次数, X 是一个随机变量。记 $q = 1 - p$, 即有

$$P\{X = k\} = \binom{n}{k} p^k q^{1-k}$$

注意到 $\binom{n}{k} p^k q^{1-k}$ 刚好是 $(p + q)^n$ 的展开式中出现 p^k 的那一项, 我们称随机变量 X 服从参数 n, p 的 **二项分布**, 并记为 $X \sim b(n, p)$

2.2.3 泊松分布

设随机变量 X 所有可能的值为 $0, 1, 2, \dots$, 而各个值的概率为

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

其中 $\lambda > 0$ 是常数, 则称 X 服从参数 λ 的 **泊松分布**, 记为 $X \sim \pi(\lambda)$

易知 $P\{X = k\} \geq 0$, $k = 0, 1, 2, \dots$, 且有

$$\sum_{k=0}^{\infty} P\{X = k\} = \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1$$

Theorem 2.2 (泊松定理). 设 $\lambda > 0$ 是一个常数, n 是任意正整数, 设 $np_n = \lambda$, 则对于任意固定的非负整数 k , 有

$$\lim_{n \rightarrow \infty} \binom{n}{k} p_n^k (1 - p_n)^{n-k} = \frac{\lambda^k e^{-\lambda}}{k!}$$

证明. 设 $p_n = \frac{\lambda}{n}$, 有

$$\begin{aligned} \binom{n}{k} p_n^k (1-p_n)^{n-k} &= \frac{n(n-1)\dots(n-k+1)}{k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k}{k!} \left[\left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{k-1}{n}\right)\right] \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} \end{aligned}$$

当 $n \rightarrow \infty$ 时 $\left(1 - \frac{\lambda}{n}\right)^n \rightarrow e^{-\lambda}$, 故有

$$\lim_{n \rightarrow \infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{\lambda^k e^{-\lambda}}{k!}$$

□

也就是说以 n, p 为参数的二项分布的概率值可以由参数为 $\lambda = np$ 的泊松分布的概率值近似

2.3 随机变量的分布函数

Definition 2.3. 设 X 是一个随机变量, x 是任意实数, 函数

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

称为 X 的 **分布函数**

对于任意实数 $x_1, x_2 (x_1 < x_2)$, 有

$$\begin{aligned} P\{x_1 < X \leq x_2\} &= P\{X \leq x_2\} - P\{X \leq x_1\} \\ &= F(x_2) - F(x_1) \end{aligned}$$

分布函数 $F(x)$ 具有以下的基本性质

1. $F(x)$ 是一个不减函数
2. $0 \leq F(x) \leq 1$, 且

$$F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0$$

$$F(\infty) = \lim_{x \rightarrow \infty} F(x) = 1$$

3. $F(x+0) = F(x)$

2.4 连续型随机变量及其概率密度

如果对于随机变量 X 的分布函数 $F(x)$, 存在非负函数 $f(x)$ 使对于任意实数 x 有

$$F(x) = \int_{-\infty}^x f(t)dt$$

则称 X 为 **连续型随机变量**, 其中函数 $f(x)$ 称为 X 的 **概率密度函数**, 简称 **概率密度**

概率密度 $f(x)$ 具有以下性质

1. $f(x) \geq 0$
2. $\int_{-\infty}^{\infty} f(x)dx = 1$
3. 对于任意实数 $x_1, x_2 (x_1 \leq x_2)$,

$$P\{x_1 < X \leq x_2\} = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x)dx$$

4. 若 $f(x)$ 在点 x 处连续, 则有 $F'(x) = f(x)$

2.4.1 均匀分布

若连续型随机变量 X 有概率密度

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

则称 X 在区间 (a, b) 上服从 **均匀分布**, 记为 $X \sim U(a, b)$ 。分布函数

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

2.4.2 指数分布

若连续型随机变量 X 的概率密度为

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta} & x > 0 \\ 0 & \text{其他} \end{cases}$$

其中 $\theta > 0$ 为常数, 则称 X 服从参数 θ 的 **指数分布**

$$F(x) = \begin{cases} 1 - e^{-x/\theta} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

对于任意 $s, t > 0$, 有

$$P\{X > s + t | X > s\} = P\{X > t\}$$

事实上

$$\begin{aligned} P\{X > s + t | X > s\} &= \frac{P\{(X > s + t) \cap (X > s)\}}{P\{X > s\}} \\ &= \frac{P\{X > s + t\}}{P\{X > s\}} = \frac{1 - F(s + t)}{1 - F(s)} \\ &= e^{-t/\theta} = P\{X > t\} \end{aligned}$$

这个性质称为无记忆性

2.4.3 正态分布

若连续型随机变量 X 的概率密度为

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

其中 $\mu, \sigma (\sigma > 0)$ 为常数, 则称 X 服从参数为 μ, σ 的 **正态分布** 或 **高斯分布**, 记为 $X \sim N(\mu, \sigma^2)$

令 $(x - \mu)/\sigma = t$, 记 $I = \int_{-\infty}^{\infty} e^{-t^2/2} dt$, 则有 $I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(t^2+u^2)/2} dt du$, 利用极坐标得

$$I^2 = \int_0^{2\pi} \int_0^{\infty} r e^{-r^2/2} dr d\theta = 2\pi$$

$f(x)$ 有以下性质

1. 曲线关于 $x = \mu$ 对称
2. 当 $x = \mu$ 时取得最大值

$$f(\mu) = \frac{1}{\sqrt{2\pi}\sigma}$$

特别的, 当 $\mu = 0, \sigma = 1$ 时称随机变量 X 服从 **标准正态分布**, 其概率密度和分布函数分别用 $\varphi(x), \Phi(x)$ 表示, 即

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

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

易知

$$\Phi(-x) = 1 - \Phi(x)$$

Lemma 2.4. 若 $X \sim N(\mu, \sigma^2)$, 则 $Z = \frac{X-\mu}{\sigma} \sim N(0, 1)$

于是, 若 $X \sim N(\mu, \sigma^2)$, 则

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

2.5 随机变量的函数的分布

Theorem 2.5. 设随机变量 X 具有概率密度 $f_X(x), -\infty < x < \infty$, 又设函数 $g(x)$ 处处可导且恒有 $g'(x) > 0$ (或恒有 $g'(x) < 0$), 则 $Y = g(x)$ 是连续型随机变量, 其概率密度为

$$f_Y(y) = \begin{cases} f_X[h(y)]|h'(y)| & \alpha < y < \beta \\ 0 & \text{其他} \end{cases}$$

其中 $\alpha = \min\{g(-\infty), g(\infty)\}, \beta = \max\{g(-\infty), g(\infty)\}$, $h(y)$ 是 $g(x)$ 的反函数

Proposition 2.6. 设随机变量 $X \sim N(\mu, \sigma^2)$, 试证明 X 的线性函数 $Y = aX + b (a \neq 0)$ 也服从正态分布

证明. X 的概率密度为

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

由 $Y = aX + b$

$$x = \frac{y-b}{a}, h'(y) = \frac{1}{a}$$

因此

$$\begin{aligned} f_Y(y) &= \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right), -\infty < y < \infty \\ &= \frac{1}{|a|\sigma\sqrt{2\pi}} e^{-\frac{(\frac{y-b}{a}-\mu)^2}{2(a\sigma)^2}} \end{aligned}$$

即有 $Y = aX + B \sim N(a\mu + b, (a\sigma)^2)$

□

3 多维随机变量及其分布

3.1 二维随机变量

设 E 是一个随机试验，它的样本空间是 $S = \{e\}$ ，设 $X = X(e), Y = Y(e)$ 是定义在 S 上的随机变量，它们构成的一个向量 (X, Y) 叫做 **二维随机向量** 或 **二维随机变量**

Definition 3.1. 设 (X, Y) 是二维随机变量，对于任意实数 x, y ，二元函数

$$F(x, y) = P\{(X \leq x) \cap (Y \leq y)\} \text{ (written as } P\{X \leq x, Y \leq y\})$$

称为二维随机变量的 **分布函数**，或称为随机变量 X, Y 的 **联合分布函数**

分布函数 $F(x, y)$ 具有以下性质

1. $F(x, y)$ 是变量 x, y 的不减函数
2. $0 \leq F(x, y) \leq 1$ ，且

$$\forall y, F(-\infty, y) = 0$$

$$\forall x, F(x, -\infty) = 0$$

$$F(-\infty, -\infty) = 0, F(\infty, \infty) = 1$$

3. $F(x+0, y) = F(x, y), F(x, y+0) = F(x, y)$ ，即 $F(x, y)$ 关于 x 又连续，关于 y 也右连续
4. 对于任意 $(x_1, y_1), (x_2, y_2), x_1 < x_2, y_1 < y_2$ ，下列不等式成立

$$F(x_2, y_2) - F(x_2, y_1) + F(x_1, y_1) - F(x_1, y_2) \geq 0$$

如果二维随机变量 (X, Y) 全部可能取到的值是有限对或可列无限多对, 则称 (X, Y) 是 **离散型的随机变量**

设二维离散型随机变量 (X, Y) 所有可能取的值为 $(x_i, y_j), i, j = 1, 2, \dots$, 记 $P\{X = x_i, Y = y_j\} = p_{ij}$, 称为二维离散型随机变量 (X, Y) 的 **分布律**, 或随机变量 X, Y 的 **联合分布律**

Y, X	x_1	\dots	x_i	\dots
y_1	p_{11}	\dots	p_{i1}	\dots
\vdots				
y_j	p_{1j}		p_{ij}	

对于二维随机变量 (X, Y) 的分布函数 $F(x, y)$, 如果存在非负的函数 $f(x, y)$ 使对于任意 x, y 有

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

则称 (X, Y) 是 **连续型的二维随机变量**, 函数 $f(x, y)$ 称为二维随机变量的 **概率密度**, 或称为随机变量 X, Y 的 **联合概率密度**

概率密度 $f(x, y)$ 具有以下性质

1. $f(x, y) \geq 0$
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$
3. 设 G 是 xOy 平面的区域, 点 (X, Y) 落在 G 内的概率为

$$P\{(X, Y) \in G\} = \iint_G f(x, y) dx dy$$

4. 若 $f(x, y)$ 在点 (x, y) 连续, 则有

$$\frac{\partial^2 F(x, y)}{\partial x \partial y} = f(x, y)$$

3.2 边缘分布

$F_X(x), F_Y(y)$ 称为二维随机变量 (X, Y) 关于 X, Y 的 **边缘分布函数**

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

记

$$p_{i\cdot} \sum_{j=1}^{\infty} p_{ij} = P\{X = x_i\}$$

$$p_{\cdot j} \sum_{i=1}^{\infty} p_{ij} = P\{X = y_j\}$$

分别称为 $p_{i\cdot}$ 和 $p_{\cdot j}$ 为 (X, Y) 关于 X, Y 的 **边缘分布律**

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

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

分别为 X, Y 的 **边缘概率密度**

Example 3.1. 设二维随机变量 (X, Y) 的概率密度

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ \frac{-1}{2(1-\rho^2)} \left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2} \right] \right\}$$

其中 $\mu_1, \mu_2, \sigma_1, \sigma_2, \rho$ 都是常数, 且 $\sigma_1, \sigma_2 > 0, -1 < \rho < 1$, 我们称 (X, Y) 服从参数为 $\mu_1, \mu_2, \sigma_1, \sigma_2, \rho$ 的 **二维正态分布**, 记为 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, 试求二维正态分布随机变量的边缘概率密度

令 $t = \frac{1}{\sqrt{1-\rho^2}} \left(\frac{y-\mu_2}{\sigma_2} - \rho \frac{x-\mu_1}{\sigma_1} \right)$, 则有

$$f_X(x) = \frac{1}{2\pi\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt$$

即

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$$

3.3 条件分布

Definition 3.2. 设 (X, Y) 是二维离散型随机变量, 对于固定的 j , 若 $P\{Y = y_j\} > 0$, 则称

$$P\{X = x_i | Y = y_j\} = \frac{P\{X = x_i, Y = y_j\}}{P\{Y = y_j\}} = \frac{p_{ij}}{p_j}$$

为在 $Y = y_j$ 条件下随机变量 X 的 **条件分布律**

Definition 3.3. 设二维随机变量 (X, Y) 的概率密度为 $f(x, y)$, (X, Y) 关于 Y 的边缘概率密度为 $f_Y(y)$, 若对于固定的 y , $f_Y(y) > 0$, 则称 $\frac{f(x, y)}{f_Y(y)}$ 为在 $Y = y$ 的条件下 X 的 **条件概率密度**, 记为

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

称 $F_{X|Y}(x|y) = \int_{-\infty}^x f_{X|Y}(x|y) dx = \int_{-\infty}^x \frac{f(x, y)}{f_Y(y)} dx$ 为在 $Y = y$ 下 X 的条件分布函数

3.4 相互独立的随机变量

Definition 3.4. 设 $F(x, y)$ 及 $F_X(x), F_Y(y)$ 分别是二维随机变量 (X, Y) 的分布函数及边缘分布函数, 若对于所有 x, y 有

$$P\{X \leq x, Y \leq y\} = P\{X \leq x\}P\{Y \leq y\}$$

$$F(x, y) = F_X(x)F_Y(y)$$

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

则称 X, Y 是 **相互独立的**

下面考查二维正态随机变量 (X, Y) , 它的概率密度为

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ \frac{-1}{2(1-\rho^2)} \left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2} \right] \right\}$$

3 多维随机变量及其分布

如果 $\rho = 0$ 则对于所有 $x, y, f(x, y) = f_X(x)f_Y(y)$ 。如果 X, Y 相互独立, 令 $x = \mu_1, y = \mu_2$, 则 $\rho = 0$

对于二维正态随机变量 (X, Y) , X, Y 相互独立的充要条件是参数 $\rho = 0$

Theorem 3.5. 设 (X_1, \dots, X_m) 和 (Y_1, \dots, Y_n) 相互独立, 则 X_i 和 Y_j 相互独立。又若 h, g 是连续函数, 则 $h(X_1, \dots, X_m)$ 和 $g(Y_1, \dots, Y_n)$ 相互独立

3.5 两个随机变量的函数的分布