

1. 随机事件及其概率	4
1.1. 随机试验和随机事件	4
1.1.1. 随机试验	4
1.1.2. 样本空间	4
1.1.3. 随机事件	4
1.2. 事件的关系、运算及其运算律	4
1.2.1. 事件 A 与 B 之和 (并)	4
1.2.2. 事件 A 与 B 的积	4
1.2.3. 事件 A 与 B 的差	4
1.2.4. 事件的包含	4
1.2.5. 事件相等	5
1.2.6. 互斥事件	5
1.2.7. 对立事件	5
1.2.8. 事件的运算律	5
1.3. 概率及其性质	5
1.3.1. 概率的定义	5
1.3.2. 概率的性质	5
1.4. 三大概型	6
1.4.1. 古典概型	6
1.4.2. 几何概型	6
1.4.3. 伯努利(Bernoulli)模型	6
1.5. 三大概率公式	6
1.5.1. 条件概率与乘法公式	6
1.5.2. 全概率公式	7
1.5.3. 贝叶斯(Bayes)公式	7
1.6. 事件的独立性	7
1.6.1. 两个事件的独立性	7
1.6.2. 两事件相互独立的充要条件	8
1.6.3. n 个事件的两两独立于相互独立 ($n \geq 3$)	8
2. 随机变量及其分布	8
2.1. 分布函数	8
2.2. 离散型随机变量	8
2.2.1. 分布律	9
2.2.2. 分布律的性质	9
2.2.3. 分布函数 $F(x)$	9
2.3. 连续型随机变量	9
2.3.1. 定义	9
2.3.2. 概率密度函数的性质	9
2.3.3. 连续型随机变量分布函数的性质	10
2.4. 离散型分布 (五大分布)	10

2.4.1. 0-1分布	10
2.4.2. 二项分布	10
2.4.3. 几何分布	10
2.4.4. 超几何分布	10
2.4.5. 泊松分布	11
2.5. 连续型分布 (三大分布)	11
2.5.1. 均匀分布	11
2.5.2. 指数分布	11
2.5.3. 正态分布	12
2.6. 随机变量函数的分布	13
2.6.1. X 是离散型随机变量	13
2.6.2. X 是连续型随机变量	13
3. 多维随机变量及其分布	13
3.1. 二维随机变量联合分布函数及其性质	13
3.1.1. 二维RV联合分布函数定义	13
3.1.2. 二维RV联合分布函数性质	14
3.1.3. 联合分布函数的几何意义	14
3.1.4. 边缘分布函数	14
3.2. 二维离散型随机变量的联合概率分布、边缘概率分布及条件概率分布	14
3.2.1. 联合分布律	14
3.2.2. 边缘分布律	15
3.2.3. 条件分布律	15
3.3. 二维连续型随机变量的联合概率密度、边缘概率密度及条件概率密度	16
3.3.1. 连续型随机变量	16
3.3.2. 联合概率密度 $f(x, y)$ 的性质:	16
3.3.3. 边缘概率密度	16
3.3.4. 条件概率密度	16
3.4. 常见的二维分布	17
3.4.1. 二维均匀分布	17
3.4.2. 二维正态分布	17
3.5. 随机变量的独立性	17
3.5.1. 一般型随机变量	17
3.5.2. 离散型随机变量	17
3.5.3. 连续型随机变量	18
3.5.4. 随机变量函数的独立性	18
3.5.5. 二维正态分布关于独立的相关结论	18
3.6. 两个随机变量函数的分布	18
3.6.1. 两个随机变量函数的定义	18
3.6.2. 二维离散型随机变量函数的分布律	18
3.6.3. 二维连续型随机变量函数的分布	19
3.6.4. 混合型	19

3.7. 极值分布	19
4. 随机变量的数字特征	19
4.1. 随机变量的数学期望	19
4.1.1. 一维随机变量的数学期望	19
4.1.2. 随机变量函数的数字期望	20
4.1.3. 数学期望的性质	21
4.2. 随机变量的方差	21
4.2.1. 方差及标准差的概念	21
4.2.2. 方差的计算公式	22
4.2.3. 方差的性质	22
4.3. 协方差	22
4.3.1. 协方差的概念	23
4.3.2. 协方差的计算公式	23
4.3.3. 协方差的性质	23
4.4. 相关系数	23
4.4.1. 相关系数的定义	23
4.4.2. 相关系数的性质	23
4.4.3. 随机变量 X, Y 等价的五个结论	23
5. 大数定律与中心极限定理	24
5.1. 切比雪夫不等式	24
5.2. 大数定律	24
5.2.1. 依概率收敛	24
5.2.2. 伯努利大数定律 (即频率依概率收敛于概率)	24
5.2.3. 辛钦大数定律	24
5.2.4. 切比雪夫大数定律	25
5.3. 中心极限定理	25
5.3.1. 独立同分布中心极限定理	25
5.3.2. 拉普拉斯中心极限定理	25
6. 数理统计的基本概念	26
7. 参数估计与假设检验	26

1. 随机事件及其概率

1.1. 随机试验和随机事件

1.1.1. 随机试验

概率论中将满足下面三个条件的试验称为随机试验，简称试验：

- (1) 可在相同的条件下重复进行；
- (2) 所有的结果是明确可知的；
- (3) 试验之前不能确定哪一个结果会发生。

1.1.2. 样本空间

随机试验的所有可能结果所组成的集合称为样本空间，常记为 Ω ， Ω 中的元素称为样本点。

1.1.3. 随机事件

样本空间的子集，即试验的结果称为随机事件，简称事件，由一个样本点组成的单点集，称为基本事件。

另外，两个特殊的事件为：

必然事件 Ω ——每次试验中一定发生的事件；

不可能事件 \emptyset ——每次试验中一定不发生的事件。

1.2. 事件的关系、运算及其运算律

1.2.1. 事件 A 与 B 之和（并）

$A \cup B$ （或 $A + B$ ）表示事件 A 与 B 至少有一个发生

1.2.2. 事件 A 与 B 的积

$A \cap B$ （或 AB ）表示事件 A 与 B 同时发生

1.2.3. 事件 A 与 B 的差

$A - B$ 表示事件 A 发生而 B 不发生

性质： $A - B = A\bar{B}$ 或 $A \cap \bar{B}$ 。

1.2.4. 事件的包含

若事件 A 发生必然导致事件 B 发生，则称事件 B 包含 A （或 A 包含于 B ），记为 $B \supset A$ 。

1.2.5. 事件相等

若 $A \supset B$ 且 $B \subset A$, 则称事件 A 与 B 相等, 记为 $A = B$.

1.2.6. 互斥事件

在试验中, 若事件 A 与 B 不能同时发生, 即 $A \cap B = \emptyset$, 则称 A 、 B 为互斥事件.

1.2.7. 对立事件

每次试验中, “事件 A 不发生”的事件称为 A 的对立事件或逆事件, A 的对立事件记为 \bar{A} .
性质:

- (1) $A + \bar{A} = \Omega$ (必然事件)
- (2) $A\bar{A} = \emptyset$ (不可能事件)

1.2.8. 事件的运算律

- (1) 吸收律 若 $A \subset B$, 则 $A \cup B = B$, $AB = A$, $A\bar{B} = \emptyset$.
- (2) 交换律 $A \cup B = B \cup A$, $AB = BA$.
- (3) 结合律 $(A \cup B) \cup C = A \cup (B \cup C)$, $(A \cap B) \cap C = A \cap (B \cap C)$.
- (4) 分配律 $(A \cup B)C = (AC) \cup (BC)$, $A \cup (BC) = (A \cup B)(A \cup C)$.
- (5) 德摩根律¹ $\overline{A_1 \cup A_2} = \bar{A}_1 \cap \bar{A}_2$, $\overline{A_1 \cap A_2} = \bar{A}_1 \cup \bar{A}_2$.

1.3. 概率及其性质

1.3.1. 概率的定义

设随机试验 E 的样本空间为 Ω , 则称满足下列条件的事件集上的函数 $P(\cdot)$ 为概率:

- (1) 对于任意事件 A , $P(A) \geq 0$;
- (2) 对于必然事件 Ω , $P(\Omega) = 1$;
- (3) 设 $A_1, A_2, \dots, A_n, \dots$ 为两两互不相容的事件, 即 $A_i A_j = \emptyset (i \neq j, i, j = 1, 2, \dots)$, 则

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k).$$

1.3.2. 概率的性质

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

¹ 长杆变短杆, 开口换方向

- (2) $P(A - B) = P(A) - P(AB)$, 特别, 当 $B \subset A$ 时, $P(A - B) = P(A) - P(B)$, 且 $P(B) \leq P(A)$;
- (3) $P(A \cup B) = P(A) + P(B) - P(AB)$,
 $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(BC) - P(AC) + P(ABC)$;
 特别的, 若 A, B 互斥, 则 $P(A \cup B) = P(A) + P(B)$, $P(A - B) = P(A)$

1.4. 三大概型

1.4.1. 古典概型

如果随机试验 E 满足下列条件:

- (1) 试验的样本空间 Ω 的元素只有有限个,
- (2) 样本空间中每个元素, 即基本事件发生的可能性相同, 则称此试验为古典概型, 对于古典概型, 事件 A 的概率有下列计算公式:

$$P(A) = \frac{A \text{ 中基本事件数}}{\Omega \text{ 中基本事件总数}}$$

1.4.2. 几何概型

如果随机试验 E 的样本空间 Ω 为欧式空间中的一个区域, 且每个样本点的出现具有等可能性, 则称此试验为几何概型, 对于几何概型, 事件 A 的概率有下列计算公式:

$$P(A) = \frac{A \text{ 的度量(长度、面积、体积)}}{\Omega \text{ 的度量(长度、面积、体积)}}$$

1.4.3. 伯努利(Bernoulli)模型

如果试验 E 的结果只有两个, A 与 \bar{A} , 则称此试验为伯努利概型 (试验), 若将伯努利试验独立重复 n 次, 则称为 n 重伯努利概型, 简称伯努利概型, 在伯努利概型中, 若 $P(A) = p$, 则 n 次试验中事件 A 发生 k 次的概率为

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

1.5. 三大概率公式

1.5.1. 条件概率与乘法公式

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 发生条件下事件 B 发生的条件概率.

乘法公式: 设 $P(A_1) > 0, P(A_2) > 0$, 则

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

一般地, 设 $P(A_1 A_2 \cdots A_{n-1}) > 0$, 则

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

1.5.2. 全概率公式

设 B_1, B_2, \cdots, B_n 为一完备事件组, 即 $B_i B_j = \emptyset, i \neq j$,

$\bigcup_{i=1}^n B_i = \Omega, P(B_i) > 0, i = 1, 2, \cdots, n$, 则对事件 A 有

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

1.5.3. 贝叶斯(Bayes)公式

设 B_1, B_2, \cdots, B_n 为一完备事件组, $P(B_i) > 0, i = 1, 2, \cdots, n, P(A) > 0$, 则有

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^n P(A|B_i)P(B_i)}, j = 1, 2, \cdots, n.$$

1.6. 事件的独立性

1.6.1. 两个事件的独立性

设 A, B 是两个事件, 若有等式

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

则称 A 与 B 相互独立.

1.6.2. 两事件相互独立的充要条件

- ① $P(AB) = P(A)P(B)$
- ② $P(B|A) = P(B)$ ($P(A) > 0$) ; 或 $P(A|B) = P(A)$ ($P(B) > 0$)
- ③ 四对事件 A 与 B , A 与 \bar{B} , \bar{A} 与 B , \bar{A} 与 \bar{B} 中有一对相互独立, 则另外三对也相互独立.

1.6.3. n 个事件的两两独立于相互独立 ($n \geq 3$)

- (1) n 个事件 A_1, A_2, \dots, A_n 两两独立: 其中任何两个事件相互独立, 即有下列

$$C_n^2 = \frac{n(n-1)}{2} \text{ 个等式成立: } P(A_i A_j) = P(A_i)P(A_j) \quad (1 \leq i < j \leq n)$$

- (2) n 个事件 A_1, A_2, \dots, A_n 相互独立, 其中任意 k 个不同事件都相互独立, 即有下列

$$2^n - n - 1 \text{ 个等式成立; } P(A_{i_1} A_{i_2} \cdots A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k}), \text{ 其中}$$

$$i \leq i_1 < i_2 < \cdots < i_k \leq n, k = 2, 3, \dots, n,$$

- (3) n 个事件两两独立于相互独立的关系:

$$A_1, A_2, \dots, A_n \text{ 相互独立} \Rightarrow A_1, A_2, \dots, A_n \text{ 两两独立}$$

$$A_1, A_2, \dots, A_n \text{ 两两独立} \Rightarrow A_1, A_2, \dots, A_n \text{ 相互独立}$$

- (4) n 个事件相互独立的结论: 若 A_1, A_2, \dots, A_n 相互独立, 则由其中任意部分事件所产生的事件与另一部分事件所产生的事件相互独立.

2. 随机变量及其分布

2.1. 分布函数

- (1) 定义 设 X 为随机变量, x 为任意实数, 则函数 $F(x) = P(X \leq x)$ 称为随机变量 X 的分布函数

- (2) 性质:

$$① 0 \leq F(x) \leq 1, -\infty < x < +\infty;$$

$$② F(x) \text{ 是单调不减的函数, 即 } x_1 < x_2 \text{ 时, 有 } F(x_1) \leq F(x_2);$$

$$③ \lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow +\infty} F(x) = 1;$$

$$④ \lim_{x \rightarrow x_0^+} F(x) = F(x_0), \text{ 即 } F(x) \text{ 是右连续的;}$$

2.2. 离散型随机变量

2.2.1. 分布律

设X的所有取值为 $x_k (k = 1, 2, \dots)$ ，则称事件 $\{X = x_k\}$ 的概率即 $P\{X = x_k\} = p_k$ ，或下列表格。

X	x_1	x_2	, ... ,	x_n	...
P	p_1	p_2	, ... ,	p_n	...

为X的分布律

2.2.2. 分布律的性质

$$(1) p_i \geq 0, i = 1, 2, \dots, n, \dots$$

$$(2) \sum_{k=1}^{\infty} p_k = 1$$

2.2.3. 分布函数 $F(x)$

设X的分布律为 $P\{X = x_k\} = p_k, k = 1, 2, \dots$ ，则X的分布函数为

$$F(x) = P(X \leq x) = \sum_{x_k \leq x} P\{X = x_k\}.$$

若已知X的分布函数 $F(x)$ ，则可求得X的分布律：

$$P\{X = x_k\} = F(x_k) - \lim_{x \rightarrow x_k^-} F(x), k = 1, 2, \dots$$

2.3. 连续型随机变量

2.3.1. 定义

若随机变量X的分布函数 $F(x)$ ，可以表示成非可积负函数 $f(x)$ 的积分形式：

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

则称X为连续型随机变量， $f(x)$ 称为X的概率密度函数，简称概率密度。

2.3.2. 概率密度函数的性质

$$(1) f(x) \geq 0$$

$$(2) \int_{-\infty}^{+\infty} f(t)dt = 1$$

2.3.3. 连续型随机变量分布函数的性质

(1) $F(x)$ 是关于 x 的连续函数, 对于任何实数 c , 有 $P\{X = c\} = 0$

(2) 若 $f(x)$ 在 x 处连续, 则有 $F'(x) = f(x)$

$$\begin{aligned} P(x_1 \leq X \leq x_2) &= P(x_1 < X \leq x_2) = P(x_1 \leq X < x_2) = P(x_1 < X < x_2) \\ &= F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(t)dt \end{aligned}$$

2.4. 离散型分布（五大分布）

2.4.1. 0-1分布

设一次伯努利试验中事件 A 发生 X 次, 则 X 服从0-1分布, 其分布律为

X	1	0
P	p	$1-p$

其中 p 为事件 A 出现的概率, $0 < p < 1$.

2.4.2. 二项分布

设 n 重伯努利试验中事件 A 发生 X 次, 则 X 服从二项分布, 记作 $X \sim B(n, p)$, 其分布律为

$$P(X = k) = C_n^k p^k q^{n-k}, \quad k = 0, 1, 2, \dots, n,$$

2.4.3. 几何分布

进行独立重复试验, 设每次试验成功的概率为 p , 失败的概率为 $q = 1 - p$ ($0 < p < 1$), 将试验进行到出现一次成功为止, 以 X 表示所需的试验次数, 则 X 服从参数为 p 的几何分布, 记为 $X \sim G(p)$, 其分布律为

$$P(X = k) = q^{k-1}p, \quad k = 1, 2, 3, \dots, \text{ 其中 } p \geq 0, q = 1 - p$$

2.4.4. 超几何分布

N 件产品, 其中 M 件次品, 从中任取 n 件, 有 X 个次品, 则随机变量 X 服从参数为 n, N, M 的超几何分布, 通常记为 $X \sim H(n, N, M)$, 其分布律为

$$P(X = i) = \frac{C_M^i C_{N-M}^{n-i}}{C_N^n}, 0 \leq i \leq n \leq N, i \leq M$$

2.4.5. 泊松分布

设随机变量X的分布律为

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \lambda > 0, k = 0, 1, 2, \dots,$$

则称随机变量X服从参数为 λ 的泊松分布，记为 $X \sim P(\lambda)$ 。

2.5. 连续型分布（三大分布）

2.5.1. 均匀分布

设随机变量X的值等可能地落在 $[a, b]$ 内，则称X在 $[a, b]$ 上服从均匀分布，记为 $X \sim U(x, y)$ ，其密度分布为

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b, \\ 0, & \text{others.} \end{cases}$$

分布函数为

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

2.5.2. 指数分布

若连续型随机变量X的概率密度为

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

其中 $\lambda > 0$ ，则称X服从参数为 λ 的指数分布，记为 $X \sim E(\lambda)$ ，其分布函数

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

2.5.3. 正态分布

(1) 一般正态分布

(a) 定义 若随机变量 X 的密度函数为

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

其中 $\sigma > 0$, μ 与 σ 均为常数, 则称 X 服从参数为 μ 和 σ 的正态分布, 记为 $X \sim N(\mu, \sigma^2)$. 其分布函数

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx, \quad (-\infty < x < +\infty)$$

(b) $f(x)$ 和 $F(x)$ 的性质如下:

(i) $f(x)$ 的图形是关于 $x = \mu$ 对称, 即 $f(\mu + x) = f(\mu - x)$;

(ii) $F(\mu + x) + F(\mu - x) = 1$;

(iii) $F(\mu) = \frac{1}{2}$;

(iv) 当 $x = \mu$ 时, $f(\mu) = \frac{1}{\sqrt{2\pi}\sigma}$ 为最大值;

(v) $f(x)$ 以 x 轴为渐近线.

(2) 标准正态分布

(a) 定义 当 $\mu = 0$, $\sigma^2 = 1$, 即 X 的密度函数为

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

则称 X 服从标准正态分布, 记为 $X \sim N(0, 1)$, 其分布函数为

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

(b) $\varphi(x)$ 和 $\Phi(x)$ 的性质如下:

(i) $\varphi(-x) = \varphi(x)$

(ii) $\Phi(-a) + \Phi(a) = 1$

(iii) $\Phi(0) = \frac{1}{2}$

(3) 一般正态分布与标准正态分布之间的关系: 如果 $X \sim N(\mu, \sigma^2)$, 则

$$\frac{X - \mu}{\sigma} \sim N(0, 1), \quad P(x_1 < X \leq x_2) = \Phi\left(\frac{x_2 - \mu}{\sigma}\right) - \Phi\left(\frac{x_1 - \mu}{\sigma}\right).$$

2.6. 随机变量函数的分布

随机变量 Y 是随机变量 X 的函数 $Y = g(X)$ ，其中 $g(x)$ 为连续函数或分段函数，现要求 $Y = g(X)$ 的概率分布，分两种情形讨论：

2.6.1. X 是离散型随机变量

已知 X 的分布列为

X	x_1	x_2	, ... ,	x_n	...
P	p_1	p_2	, ... ,	p_n	...

显然， $Y = g(x)$ 的取值只可能是 $g(x_1), g(x_2), \dots, g(x_n), \dots$ ，若 $g(x_i)$ 互不相等，则 Y 的分布列如下：

Y	$g(x_1)$	$g(x_2)$, ... ,	$g(x_n)$...
P	p_1	p_2	, ... ,	p_n	...

若有某些 $g(x_i)$ 相等，则应将对应的 p_i 相加作为 $g(x_i)$ 的概率

2.6.2. X 是连续型随机变量

分布函数法：先按分布函数的定义求得 Y 的分布函数 $F_Y(y)$ ，再通过求导得到 Y 的概率密度 $f_Y(y)$ ，即

$$F_Y(y) = P(Y \leq y) = P(g(x) \leq y) = \int_{g(x) \leq y} f_X(x) dx, \quad f_Y(y) = \frac{dF_Y(y)}{dy};$$

3. 多维随机变量及其分布

3.1. 二维随机变量联合分布函数及其性质

3.1.1. 二维RV联合分布函数定义

设 (X, Y) 是二维随机变量，对任意实数 x 和 y ，称

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

为 (X, Y) 的分布函数，又称联合分布函数.

3.1.2. 二维RV联合分布函数性质

- (1) $0 \leq F(x, y) \leq 1$;
- (2) $F(x, y)$ 对 x 和 y 都是单调非减的，即对任意的 x_1, x_2 ，若 $x_1 > x_2$ ，则有

$$F(x_1, y) \geq F(x_2, y);$$
- (3) $\lim_{\substack{x \rightarrow +\infty \\ y \rightarrow +\infty}} F(x, y) = 1, \quad \lim_{\substack{x \rightarrow -\infty \\ y \rightarrow -\infty}} F(x, y) = \lim_{x \rightarrow -\infty} F(x, y) = \lim_{y \rightarrow -\infty} F(x, y) = 0$;
- (4) $F(x, y)$ 对 x 和 y 都是右连续.

3.1.3. 联合分布函数的几何意义

$F(x, y)$ 在 (x, y) 的函数值就是随机点 (X, Y) 在 $X = x$ 左侧和 $Y = y$ 下方的无穷矩形内的概率，对有限矩形域有：

$$P\{x_1 < X < x_2, y_1 < Y < y_2\} = F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1).$$

3.1.4. 边缘分布函数

在试验涉及二维随机变量 (X, Y) 的前提下，单个随机变型 X 与 Y 的分布函数 $F_X(x)$ 和 $F_Y(y)$ 称为二维随机变量 (X, Y) 的关于 X 和关于 Y 的边缘分布函数，即

$$F_X(x) = P\{X \leq x\} = \lim_{y \rightarrow +\infty} F(x, y)$$

$$F_Y(y) = P\{Y \leq y\} = \lim_{x \rightarrow +\infty} F(x, y)$$

3.2. 二维离散型随机变量的联合概率分布、边缘概率分布及条件概率分布

3.2.1. 联合分布律

(1) 定义

设离散型随机变量 (X, Y) 的所有可能取值为 $(x_i, y_j)(i, j = 1, 2, \dots)$ 且事件 $\{(X, Y) = (x_i, y_j)\}$ 的概率为 p_{ij} ，称

$$P(X = x_i, Y = y_j) = P_{ij}(i, j = 1, 2, \dots)$$

为 (X, Y) 的联合分布律，联合分布律有时也用下面的概率分布表来表示：

	Y						
X		y_1	y_2	\dots	y_j	\dots	$P\{X = x_i\}$

x_1	P_{11}	P_{12}	\cdots	P_{1j}	\cdots	$P\{X = x_1\}$
x_2	P_{21}	P_{22}	\cdots	P_{2j}	\cdots	$P\{X = x_2\}$
\vdots	\vdots	\vdots		\vdots	\vdots	\vdots
x_i	P_{i1}	P_{i2}	\cdots	P_{ij}	\cdots	$P\{X = x_i\}$
\vdots	\vdots	\vdots		\vdots	\vdots	\vdots
$P\{Y = y_j\}$	$P\{Y = y_1\}$	$P\{Y = y_2\}$	\cdots	$P\{Y = y_j\}$	\cdots	1

(2) 性质

(a) $P_{ij} \geq 0 (i, j = 1, 2, \cdots)$;

(b) $\sum_i \sum_j p_{ij} = 1$.

3.2.2. 边缘分布律

设随机变盘 (X, Y) 的联合分布律为

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

则 X 的边缘分布为

$$P(X = x_i) = \sum_j p_{ij} (i, j = 1, 2, \cdots);$$

Y 的边缘分布为

$$P(Y = y_j) = \sum_i p_{ij} (i, j = 1, 2, \cdots).$$

3.2.3. 条件分布律

二维离散型随机变型 (X, Y) 的联合分布律为

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

在已知 $X = x_i$ 的条件下, Y 取值的条件分布律为

$$P(Y = y_j | X = x_i) = \frac{P\{X = x_i, Y = y_j\}}{P(X = x_i)},$$

在已知 $Y = y_j$ 的条件下, X 取值的条件分布律为

$$P(Y = y_i | X = x_j) = \frac{P\{X = x_i, Y = y_j\}}{P(Y = y_j)},$$

其中 $P(X = x_i)$, $P(Y = y_j)$ 分别为 X , Y 的边缘分布律.

3.3. 二维连续型随机变量的联合概率密度、边缘概率密度及条件概率密度

3.3.1. 连续型随机变量

设 $F(x, y)$ 为二维随机变量 (X, Y) 的分布函数, 如果存在非负函数 $f(x, y)$ 使得对任意实数 x, y 有

$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv$, 则称 (X, Y) 为二维连续型随机变量, $f(x, y)$ 为 (X, Y) 的联合概率密度.

3.3.2. 联合概率密度 $f(x, y)$ 的性质:

- (1) $f(x, y) \geq 0$;
- (2) $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$.

3.3.3. 边缘概率密度

当 (X, Y) 为连续型随机变量, 并且其联合概率密度为 $f(x, y)$, 则 X 和 Y 的边缘概率密度为

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy, \quad f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx.$$

3.3.4. 条件概率密度

二维连续型随机变量 (X, Y) 的联合概率密度为 $f(x, y)$, 则在已知 $Y = y$ 的条件下, X 的条件概率密度为

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

在已知 $X = x$ 的条件下， Y 的条件概率密度为

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

其中 $f_X(x) > 0, f_Y(y) > 0$ 分别为 X, Y 的边缘概率密度.

3.4. 常见的二维分布

3.4.1. 二维均匀分布

设随机变量 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} 1/S_D, & (x, y) \in D \\ 0, & \text{others} \end{cases}$$

其中 S_D 为区域 D 的面积，则称 (X, Y) 服从 D 上的均匀分布，记为 $(X, Y) \sim U(D)$.

3.4.2. 二维正态分布

设随机变量 (X, Y) 的概率密度为

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

其中 $\mu_1, \mu_2, \sigma_1 > 0, \sigma_2 > 0, |\rho| \leq 1$ ，则称 (X, Y) 服从二维正态分布，记为 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

3.5. 随机变量的独立性

3.5.1. 一般型随机变量

相互独立：设 (X, Y) 的分布函数为 $F(x, y)$ ，边缘分布函数为 $F_X(x)$ 和 $F_Y(y)$ ，如果一切 x, y 有 $F(x, y) = F_X(x)F_Y(y)$ 则称 X 与 Y 相互独立.

3.5.2. 离散型随机变量

X 与 Y 相互独立的充要条件是： $P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$

3.5.3. 连续型随机变量

X 与 Y 是相互独立的充要条件是 $f(x, y) = f_X(x)f_Y(y)$

3.5.4. 随机变量函数的独立性

若 X 与 Y 独立, h, g 为 X 与 Y 连续函数, 则: $h(X)$ 和 $g(Y)$ 独立.

3.5.5. 二维正态分布关于独立的相关结论

1. 若 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, 则 $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$; 但是若 $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$, 则 (X, Y) 未必是二维正态分布.
2. 若 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, 则随机变量 X, Y 的非零线性组合 $Z = aX + bY \neq 0$ 服从一维正态分布.
3. 若 $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$, 且 (X, Y) 相互独立, 则 (X, Y) 服从二维正态分布.
4. 若 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, 则 X 和 Y 相互独立的充要条件是 $\rho = 0$
5. 若 X_1, X_2, \dots, X_n 相互独立且服从正态分布, 则 X_1, X_2, \dots, X_n 的线性组合服从正态分布

3.6. 两个随机变量函数的分布

3.6.1. 两个随机变量函数的定义

设 X, Y 为两个随机变量, $z = g(x, y)$ 为二元连续函数, 则称 $Z = g(X, Y)$ 为随机变量 X, Y 的函数, 显然两随机变量函数是一维随机变量.

3.6.2. 二维离散型随机变量函数的分布律

$Z = g(X, Y)$ 也是离散型随机变量, 其分布律为

$$P\{Z = g(x_i, y_j)\} = P\{X = x_i, Y = y_j\} = P_{ij},$$

如果有若干个 $g(x_i, y_j)$ 相同, 则合并诸项为一项, 并将相应的概率相加作为 Z 取值为 $g(x_i, y_j)$ 的概率.

Z 的分布函数为: $F_Z(z) = P\{Z \leq z\} = P\{g(X, Y) \leq z\} = \sum_{g(x_i, y_j) \leq z} P(X = x_i, Y = y_j).$

3.6.3. 二维连续型随机变量函数的分布

设 (X, Y) 的联合概率密度为 $f(x, y)$ ，则 $Z = g(X, Y)$ 的分布函数为

$$F_Z(z) = P\{Z \leq z\} = P\{g(X, Y) \leq z\} = \iint_{g(x,y) \leq z} f(x, y) dx dy$$

3.6.4. 混合型

(X 和 Y 中有一个是离散型随机变量，另一个为连续型随机变量) 利用分布函数法求之.

3.7. 极值分布

设 X 与 Y 是两个相互独立的随机变量，其分布函数分别为 $F_X(x)$ 和 $F_Y(y)$ ，则 $Z = \max(X, Y)$ 的分布函数为：

$$F_{\max}(z) = F_X(z)F_Y(z);$$

$Z = \min(X, Y)$ 的分布函数为：

$$F_{\min}(z) = 1 - [1 - F_X(z)][1 - F_Y(z)].$$

特别地当 X_1, X_2, \dots, X_n 相互独立时

$Z = \max(X_1, X_2, \dots, X_n)$ 的分布函数为：

$$F_{\max}(z) = F_{X_1}(z)F_{X_2}(z) \cdots F_{X_n}(z)$$

$Z = \min(X_1, X_2, \dots, X_n)$ 的分布函数分别为：

$$F_{\min}(z) = 1 - [1 - F_{X_1}(z)][1 - F_{X_2}(z)] \cdots [1 - F_{X_n}(z)]$$

4. 随机变量的数字特征

4.1. 随机变量的数学期望

4.1.1. 一维随机变量的数学期望

(1) 离散型随机变量的数学期望

设离散型随机变量 X 的分布律为 $P\{X = X_i\} = p_i (i = 1, 2, \dots)$. 若无穷级数 $\sum_{i=1}^{\infty} x_i p_i$ 绝对收敛,

则称无穷级数 $\sum_{i=1}^{\infty} x_i p_i$ 的和为随机变量 X 的数学期望, 记作 $E(X)$, 即

$$E(X) = \sum_{i=1}^{\infty} x_i p_i$$

(2) 连续型随机变量的数学期望

设 X 为连续型随机变量, 其概率密度为 $f(x)$, 若反常积分 $\int_{-\infty}^{+\infty} x f(x) dx$ 绝对收敛, 则

称反常积分 $\int_{-\infty}^{+\infty} x f(x) dx$ 的值为随机变量 X 的数学期望, 记为 $E(X)$, 即

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

4.1.2. 随机变量函数的数字期望

(1) 一维随机变量函数的数学期望

设 Y 是随机变量 X 的函数: $Y = g(X)$ (g 是连续函数或分段连续函数).

设离散型随机变量 X 具有分布律 $p\{X = x_i\} = p_i, i = 1, 2, \dots$, 且无穷级数

$\sum_{i=1}^{\infty} g(x_i) p_i$ 绝对收敛; 则有:

$$E(Y) = E[g(X)] = \sum_{i=1}^{\infty} g(x_i) p_i$$

设连续型随机变量 X 具有概率密度函数 $f(x)$, 且反常积分 $\int_{-\infty}^{+\infty} g(x) f(x) dx$ 绝对收敛,

则有:

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

(2) 二维随机变量函数的数学期望

设 Z 是随机变量 X, Y 的函数: $Z = g(X, Y)$ (g 是连续函数), 那么, Z 是一个一维随机变量.

设二维离散型随机变量 (X, Y) 具有分布律 $P\{X = x_i, Y = y_j\} = p_{ij} (i, j = 1, 2, \dots)$; 则有

$$E(Z) = E[g(X, Y)] = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} g(x, y) P_{ij}$$

或者, 设 Z 具有分布律 $P\{Z = z_i\} = p_i (i = 1, 2, \dots)$, 则有:

$$E(Z) = \sum_{i=1}^{\infty} z_i p_i$$

这里的无穷级数绝对收敛.

设二维连续型随机变量 (X, Y) 具有联合概率密度 $f(x, y)$, 则有

$$E(Z) = E[g(X, Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f(x, y) dx dy$$

设 Z 具有概率密度 $f_Z(z)$, 则有:

$$E(Z) = \int_{-\infty}^{+\infty} z f_Z(z) dz$$

这里的反常积分绝对收敛.

4.1.3. 数学期望的性质

- (1) $E(C) = C$, (C 为常数)
- (2) $E(CX) = CE(X)$;
- (3) $E(X + C) = E(X) + C$;
- (4) $E(X + Y) = E(X) + E(Y)$;

综合(1)(2)(4)有 $E(\sum_{i=1}^n C_i X_i) = \sum_{i=1}^n C_i E(X_i)$;

4.2. 随机变量的方差

4.2.1. 方差及标准差的概念

设 X 是一个随机变量, 若 $E\{[X - E(X)]^2\}$ 存在, 则称 $E\{[X - E(X)]^2\}$ 为 X 的方差, 记为 $D(X)$, 即 $D(X) = E\{[X - E(X)]^2\}$

方差的算术平方根 $\sqrt{D(X)}$ 称为 X 的均方差或标准差.

- (1) 离散型随机变量的方差

$$D(X) = \sum_k [x_k - E(X)]^2 p_k$$

- (2) 连续型随机变量的方差

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

4.2.2. 方差的计算公式

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

4.2.3. 方差的性质

- (1) $D(C) = 0$, (C 为常数) ;
- (2) $D(CX) = C^2 D(X)$;
- (3) $D(X + C) = D(X)$;
- (4) 若 X_1, X_2, \dots, X_m 两两独立或两两不相关, 则

$$D(C_1 X_1 + C_2 X_2 + \dots + C_m X_m) = C_1^2 D(X_1) + C_2^2 D(X_2) + \dots + C_m^2 D(X_m).$$

分布	记号	期望	方差
0-1分布	$B(1, p)$	p	$p(1-p)$
二项分布	$B(n, p)$	np	$np(1-p)$
泊松分布	$P(\lambda)$	λ	λ
几何分布	$G(p)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
超几何分布	$H(n, M, N)$	$\frac{nM}{N}$	$\frac{nM}{N} \left(1 - \frac{M}{N}\right) \left(\frac{N-n}{N-1}\right)$
均匀分布	$U(a, b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
指数分布	$E(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
正态分布	$N(\mu, \sigma^2)$	μ	σ^2

4.3. 协方差

4.3.1. 协方差的概念

设 (X, Y) 是二维随机变量, 如果 $E[(X - EX)(Y - EY)]$ 存在, 则称它为随机变量 X 与 Y 的协方差, 记作 $Cov(X, Y) = E[(X - EX)(Y - EY)]$.

4.3.2. 协方差的计算公式

对任意两个随机变量 X, Y , 有

- (1) $Cov(X, Y) = E(XY) - E(X)E(Y)$;
- (2) $D(X \pm Y) = D(X) + D(Y) \pm 2Cov(X, Y)$.

4.3.3. 协方差的性质

- (1) $Cov(X, Y) = Cov(Y, X)$;
- (2) $Cov(X, X) = D(X)$;
- (3) $Cov(X, c) = 0$ (c 为任意常数)
- (4) $Cov(aX, bY) = ab Cov(X, Y)$ (a, b 是常数) ;
- (5) $Cov(X_1 + X_2, Y) = Cov(X_1, Y) + Cov(X_2, Y)$;

4.4. 相关系数

4.4.1. 相关系数的定义

对于随机变量 X 与 Y , 如果 X, Y 的方差都不等于零, 则称 $\rho_{XY} = \frac{Cov(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$ 为随机变量 X 与 Y 的相关系数, 记作 ρ_{XY} (简记为 ρ)

4.4.2. 相关系数的性质

- (1) $|\rho_{XY}| \leq 1$;
- (2) 当 $\rho_{XY} = 0$ 时, 称 X 与 Y 不相关.
若 X, Y 相互独立, 则 $\rho_{XY} = 0$, 反之不成立;
- (3) $|\rho| = 1 \Leftrightarrow X$ 与 Y 以概率1线性相关.
若 \exists 常数 $a > 0, b$, 使 $P\{X = aY + b\} = 1$, 则 $\rho_{XY} = 1$, 此时称 X, Y 正相关;
若 \exists 常数 $a < 0, b$, 使 $P\{X = aY + b\} = 1$, 则 $\rho_{XY} = -1$, 此时称 X, Y 负相关;

4.4.3. 随机变量 X, Y 等价的五个结论

- (1) X 与 Y 不相关
- (2) $\rho_{XY} = 0$
- (3) $Cov(X, Y) = 0$

$$(4) E(XY) = E(X)E(Y)$$

$$(5) D(X \pm Y) = D(X) + D(Y)$$

5. 大数定律与中心极限定理

5.1. 切比雪夫不等式

设随机变量 X 具有 $E(X)$ 和 $D(X)$, 则任给 $\varepsilon > 0$, 有

$$P\{|X - E(X)| \geq \varepsilon\} \leq \frac{D(X)}{\varepsilon^2}, \text{ 或 } P\left\{|X - E(X)| < \varepsilon\right\} \geq 1 - \frac{D(X)}{\varepsilon^2}.$$

5.2. 大数定律

5.2.1. 依概率收敛

设 a 是一个常数, X_n 为一随机变量序列, $\forall \varepsilon > 0$, $\exists P\{|X_n - a| < \varepsilon\} = 1$ 或 $P\{|X_n - a| \geq \varepsilon\} = 0$, 则称 $\{X_n\}$ 依概率收敛于 a , 记为 $X_n \xrightarrow{P} a$.

5.2.2. 伯努利大数定律 (即频率依概率收敛于概率)

设 n_A 是 n 次独立重复试验中事件 A 发生的次数, $P(A) = p$, 则 $\forall \varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} = 1 \text{ 或 } \lim_{n \rightarrow \infty} P\left\{\left|\frac{n_A}{n} - p\right| \geq \varepsilon\right\} = 0$$

5.2.3. 辛钦大数定律

设随机变量 X_1, X_2, \dots, X_n 相互独立, 服从同一分布 (任意分布), 且具有相同的数学期望

$E(X_k) = \mu$, $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$, 则 $\forall \varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{1}{n} \sum_{k=1}^n x_k - \mu\right| < \varepsilon\right\} = 1$$

5.2.4. 切比雪夫大数定律

设随机变量 X_1, X_2, \dots, X_n 相互独立, 数学期望 $E(X_k)$, 方差 $D(X_k)$ 均存在, 且方差有公共的

上界即 $\forall k, D(X_k) \leq C, \bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$, 则 $\forall \varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{k=1}^n X_k - \mu \right| < \varepsilon \right\} = 1$$

其中 $\mu = \frac{1}{n} \sum_{k=1}^n E(X_k)$.

5.3. 中心极限定理

5.3.1. 独立同分布中心极限定理

设 $X_1, X_2, \dots, X_n, \dots$ 是独立同分布的随机变量序列, $EX_i = \mu, DX_i = \sigma^2, i = 1, 2, \dots$, 则

随机变量 $Y_n = \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma}$ 的分布函数 $F_n(x), \forall x \in R$, 有

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

即, 当 n 充分大时, $\sum_{i=1}^n X_i$ 近似地服从以它的均值为均值, 它的方差为方差的正态分布, 即正态分布为 $N(n\mu, n\sigma^2)$.

5.3.2. 拉普拉斯中心极限定理

设 n_A 表示 n 重伯努利试验中事件 A 出现次数, $P(A) = p$, 则随机变量 $Y_n = \frac{n_A - np}{\sqrt{np(1-p)}}$ 的分布函数 $F_n(x), \forall x \in R$, 有

$$\lim_{n \rightarrow +\infty} F_n(x) = \Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

即，当 n 充分大时， n_A 近似地服从它的均值为均值，它的方差为方差的正态分布，即正态分布 $N(np, np(1-p))$ 。

6. 数理统计的基本概念

7. 参数估计与假设检验