

# 概率论与数理统计

**Probability and Mathematical Statistics**

- ◆ 第一章 随机事件及其概率
- ◆ 第二章 随机变量及其分布
- ◆ 第三章 多维随机变量及其分布
- ◆ 第四章 随机变量的数字特征
- ◆ 第五章 大数定理与中心极限定理
  
- ◆ 第六章 数理统计基础
- ◆ 第七章 参数估计
- ◆ 第八章 假设检验

# 概率论的起源和发展

- 1、16—17世纪，欧洲许多国家盛行赌博之风。
- 2、17世纪中叶，在法国出现了分赌注问题的研究，正是对这个问题的研究，推动了数学的发展，使一门崭新的学科——概率论诞生了。

1654年，法国有个叫德·梅雷的赌徒向法国的天才数学家帕斯卡提出了如下分赌注的问题：

甲乙两个赌徒每人各拿出一定数量的金币作为赌注，按某种方式赌了起来。若规定：甲乙谁先胜三局谁就赢得所有的金币。但实际情况是：甲两胜一负而因故中止了。

那么如何分这些赌注呢？

**3、17世纪末期到18世纪，在概论论上作出较大贡献的是瑞士的数学家族——贝努里家族。**

**4、18世纪到19世纪，概率论走出赌局，在生物、物理等各个领域发挥作用。**

**5、20世纪至今，成为一门严谨的学科。**



# 第一章 随机事件及其概率

- 随机事件
- 随机事件的概率
- 古典概型与几何概型
- 乘法公式与全概率公式
- 事件的独立性

# 什么是概率论

## ◆确定性现象 Certainty phenomena

- 在标准大气压下，将纯净水加热到 $100^{\circ}\text{C}$ 时必然沸腾.
- 垂直上抛一重物，该重物会垂直下落.

## ◆随机现象 Random phenomena

- 掷一颗骰子，可能出现1, 2, 3, 4, 5, 6点.
- 抛掷一枚均匀的硬币，会出现正面向上、反面向上两种结果.

概率论就是研究随机现象统计规律性的数学学科.

# 随机试验 E (*Experiment* )

对随机现象进行的试验或观察称为随机试验,简称试验。

这里试验的含义十分广泛,它包括各种各样的科学实验,也包括对事物的某一特征的观察。其典型的例子有:

$E_1$ : 抛一枚硬币,观察正面、反面出现的情况。

$E_2$ : 将一枚硬币抛掷三次,观察出现正面的次数。

$E_3$ : 抛一颗骰子,观察出现的点数。

$E_4$ : 记录寻呼台一分钟内接到的呼唤次数。

$E_5$ : 在一批灯泡中任意抽取一只,测试它的寿命。

$E_6$ : 记录某地一昼夜的最高温度和最低温度。

## 试验的基本特点：

- (1) 试验可以在相同条件下重复进行；
- (2) 试验的所有可能结果是明确可知的，并且不止一个；
- (3) 每次试验总是恰好出现这些可能结果中的一个，但在一次试验之前不能肯定这次试验会出现哪一个结果。

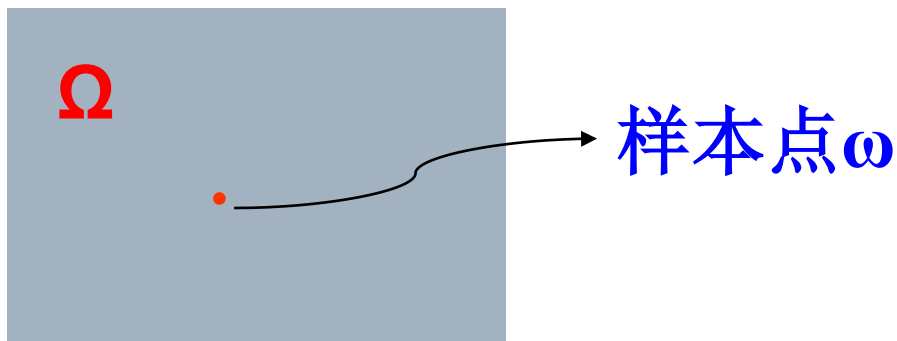
# 样本空间(Space)

## ■ 样本点 Sample Point

随机试验中的每一个可能出现的试验结果称为这个试验的一个样本点，记作 $\omega_i$ 。

## ■ 样本空间 Sample Space

全体样本点组成的集合称为这个试验的样本空间，记作 $\Omega$ 。即  $\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$



现代集合论为表述其关系提供了一个方便的工具。

## 写出下列试验的样本空间

**E1:** 掷一颗匀质骰子，观察骰子出现的点数

$\Omega = \{1, 2, 3, 4, 5, 6\}$       点数：一维离散型随机变量

**E2:** 射手向一目标射击，直到击中为止，观察射击次数

$\Omega = \{1, 2, 3, \dots\}$       射击次数：一维离散型随机变量

**E3:** 从四张扑克牌J,Q,K,A任意抽取两张。

$\Omega = \{(J,Q), \dots, (Q,A)\}$       二维离散型随机变量

**E4:** 在一批灯泡中任意抽取一只，测试它寿命

$\Omega = \{t \mid t \geq 0\}$       寿命：一维连续型随机变量

## 样本空间的说明:

1. 同一试验, 若试验目的不同, 则对应的样本空间也不同.

对于同一试验: “将一枚硬币抛掷三次”.

(1) 若观察正面  $H$ 、反面  $T$  出现的情况, 则样本空间为

$$\Omega = \{HHH, HHT, HTH, THH, HTT, TTH, THT, TTT\}.$$

(2) 若观察出现正面的次数, 则样本空间为

$$\Omega = \{0, 1, 2, 3\}.$$

## 说明:

2. 试验不同, 对应的样本空间可能不同.
3. 建立样本空间, 事实上就是建立随机现象的数学模型. 因此, 一个样本空间可以概括许多内容大不相同的实际问题。

例如: 只包含两个样本点的样本空间  $\Omega = \{H, T\}$  它既可以作为抛掷硬币出现**正面**或出现**反面**的模型, 也可以作为产品检验中**合格**与**不合格**的模型, 又能用于排队现象中**有人排队**与**无人排队**的模型等.



# 随机事件 random event

- 试验的结果称为事件。

- 事件的分类：

- 如果在每次试验的结果中,某事件可能发生，也可能不发生，则这一事件叫做随机事件(random event)，简称事件(event)。通常用大写英文字母  $A, B, C, \dots$  等表示。
- 在试验中一定会发生的事件叫做必然事件，用  $\Omega$  表示。
- 在试验中一定不会发生的事件叫不可能事件，用  $\Phi$  表示。

**例：** 抛掷一枚骰子, 观察出现的点数.



(1) 试验中, 骰子 “出现1点”, “出现2点”, “点数不大于4”, “点数为偶数” 等都为随机事件.

可以表示为  $A = \{\text{出现1点}\}$ ,

$B = \{\text{出现2点}\}$ ,  $C = \{\text{点数为偶数}\}$ ,  $D = \{\text{点数不大于4}\}$ .

(2) 骰子 “出现8点”, 这一事件是不可能事件.

可用  $\Phi = \{\text{出现8点}\}$  表示.

(3) 骰子 “点数小于10”, 这一事件是必然事件.

可用  $\Omega = \{\text{点数小于10}\}$  表示.

# 随机事件与样本空间的关系

**例：**抛掷一枚骰子，观察出现的点数.

样本空间为 $\Omega = \{1, 2, 3, 4, 5, 6\}$

不可能事件 $\Phi = \{\text{出现8点}\}$ 为空集，必然事件 $\Omega = \{\text{点数小于10}\}$ 即为样本空间， $A = \{1\}$ ， $B = \{2\}$ ， $C = \{\text{点数为偶数}\} = \{2, 4, 6\}$ ，为随机事件，且都是样本空间的一个子集.

**任一随机事件A是样本空间 $\Omega$ 的一个子集**

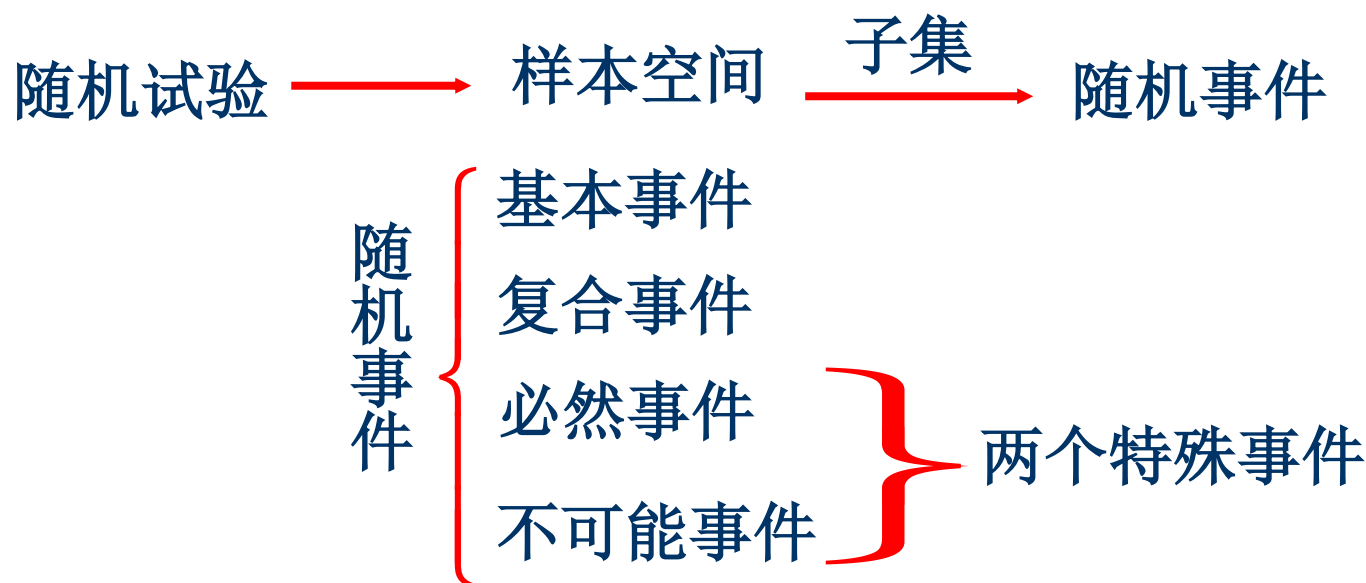
$$\Phi \subset A \subset \Omega$$

**基本事件：**由单个样本点构成的集合.

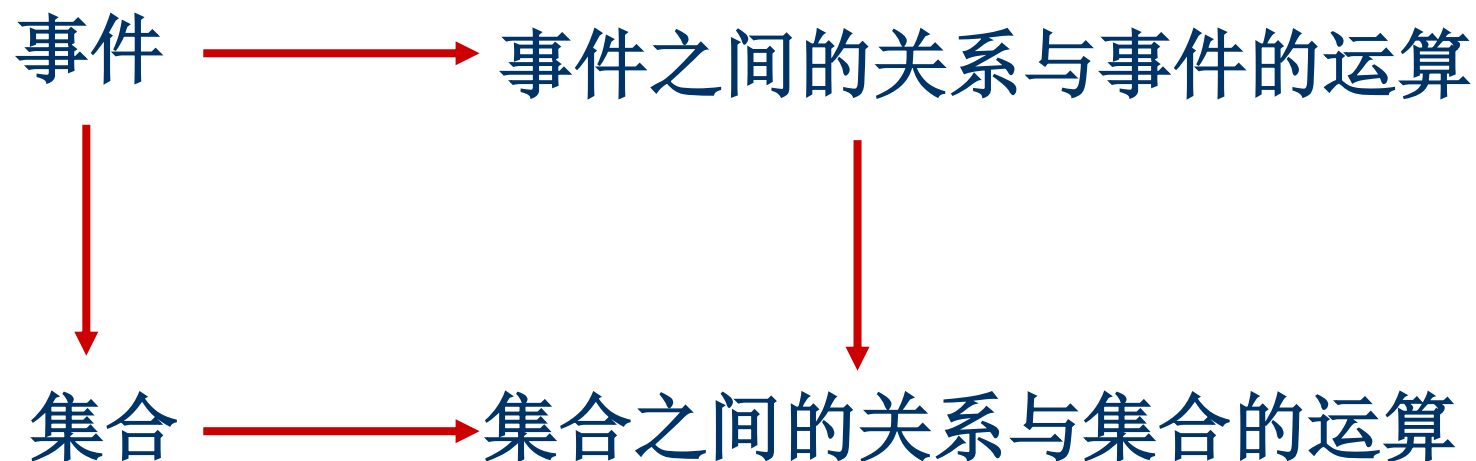
## 两点说明:

1. 当且仅当集合A中的一个样本点出现时, 称事件A发生.
2. 随机试验、样本空间与随机事件的关系

每个随机试验相应有一个样本空间, 样本空间的子集是随机事件.



# 随机事件的关系与运算

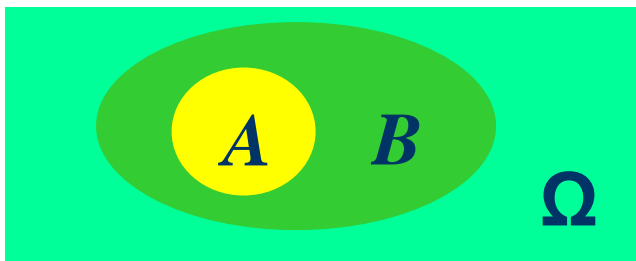


给定一个随机试验，设 $\Omega$ 为其样本空间，事件 $A$ ， $B$ ， $A_k$  ( $k=1,2,3,\dots$ ) 都是  $\Omega$  的子集.

# 1. 事件的包含

若事件  $A$  发生, 必然导致  $B$  发生, 即事件  $A$  的样本点都是事件  $B$  的样本点, 则称事件  $B$  包含事件  $A$ , 也称  $A$  是  $B$  的子集, 记作  $B \supset A$  或  $A \subset B$ .

图示  $B$  包含  $A$

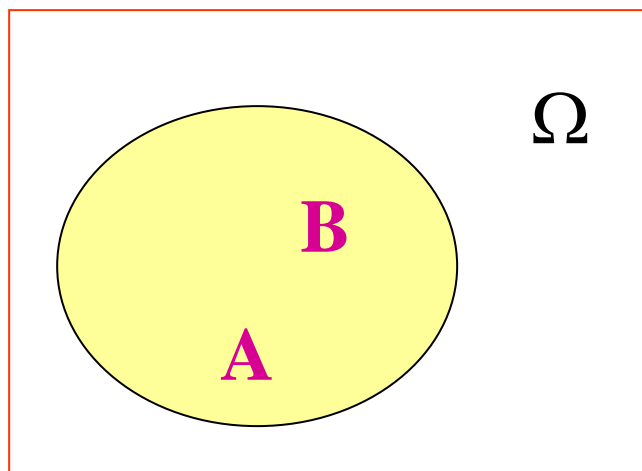


换一说法: 若事件  $B$  不发生, 必然导致 事件  $A$  也不发生。

如: 掷一颗骰子,  $A=\{2\}$ ,  $B=\{2, 4, 6\}$ , 显然  $A \subset B$

## 特别：相等事件

$$B \supset A \text{ 且 } A \supset B \iff A=B$$



事件**A**与事件**B**含有相同的样本点

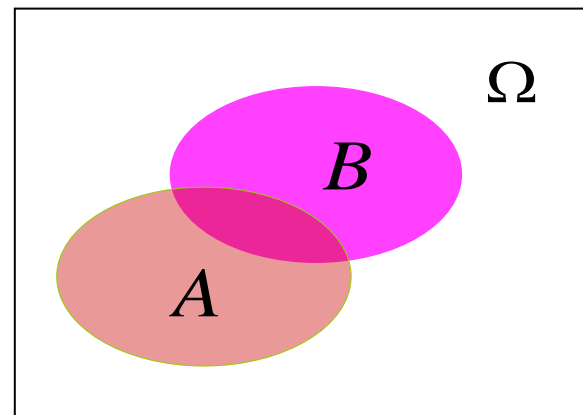
例如：在投掷一颗骰子的试验中，事件“出现偶数点”与事件“出现**2，4或6**点”是相等事件。

## 2. 事件的和（并）

和事件 $A \cup B$ 发生  $\longleftrightarrow$   $A$ 发生或 $B$ 发生

- ◆ 事件 $A$ 与事件 $B$ 至少有一个发生
- ◆ 由事件 $A$ 与事件 $B$ 所有样本点组成

关键词：至少，或



- ◆ 多个事件的和：表示 $n$ 个事件至少有一个发生

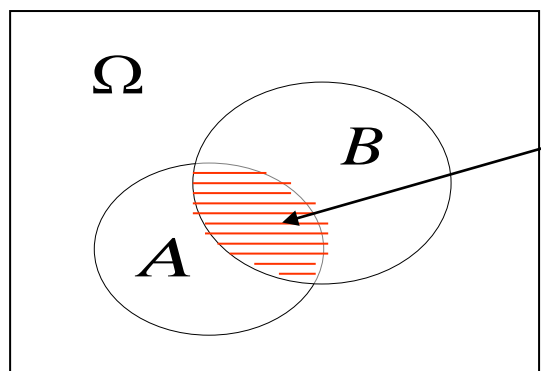
$$A_1 \cup A_2 \cup \cdots \cup A_n = \bigcup_{i=1}^n A_i$$



### 3.事件的积 (交)

积事件 $AB$ 发生  $\longleftrightarrow$  事件  $A$  和事件  $B$  同时发生

- ◆ 由事件  $A$  和事件  $B$  的公共样本点组成



$A B$  或者  $A \cap B$

关键词：同时，都，而且

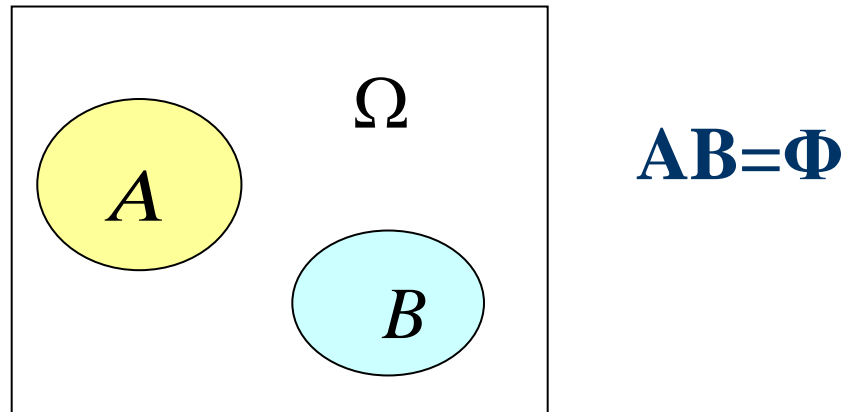
- ◆ 多个事件的积：表示  $n$  个事件同时发生

$$A_1 A_2 \cdots A_n = \bigcap_{i=1}^n A_i$$

## 4. 互不相容事件 (互斥事件)

事件A与事件B互斥  $\longleftrightarrow AB = \Phi$

- ◆ 事件A与事件B不能同时发生
- ◆ 事件A与事件B没有公共的样本点

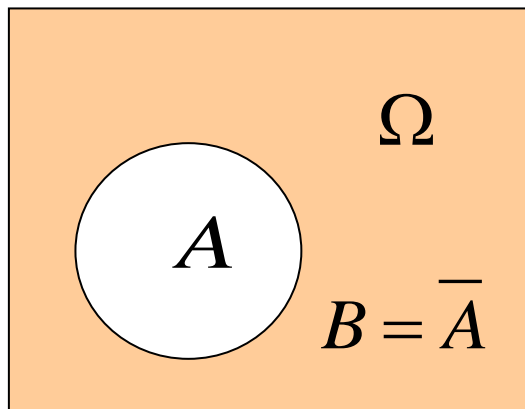


## 5. 对立事件 (互逆事件)

事件A与事件B对立  $\longleftrightarrow$  A与B中有且只有一事发生

即  $A \cup B = \Omega$  且  $AB = \Phi$  则称B是A的对立事件(逆事件), 记

作  $B = \bar{A}$



即  $\bar{A}$  是由所有不属于A的样本点组成

若A表示“事件A出现”,

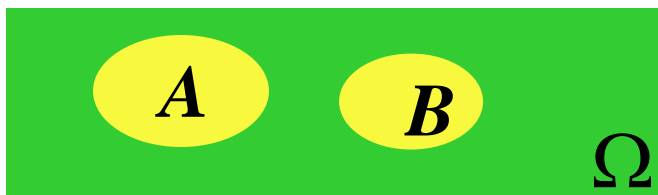
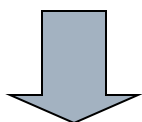
则  $\bar{A}$  表示“事件A不出现”.

◆ 性质

$$A\bar{A} = \Phi \quad A \cup \bar{A} = \Omega \quad \overline{(\bar{A})} = A$$

# 对立事件与互斥事件的区别

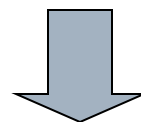
$A$ 、 $B$  互斥



$$AB = \Phi$$

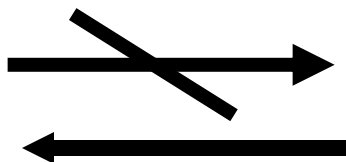
互斥

$A$ 、 $B$  对立



$$A \cup B = \Omega \text{ 且 } AB = \Phi$$

对立



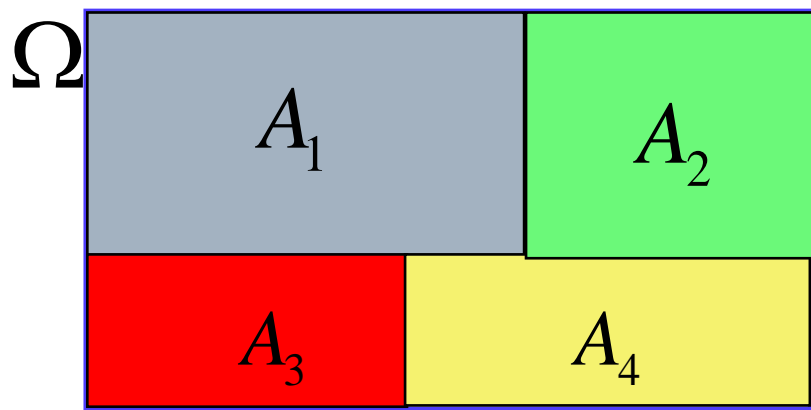
## 6. 完备事件组

完备事件组  $A_1, A_2, \dots, A_n$  满足

(1)  $A_1, A_2, \dots, A_n$  互不相容

(2)  $A_1 \cup A_2 \cup \dots \cup A_n = \Omega$

例如



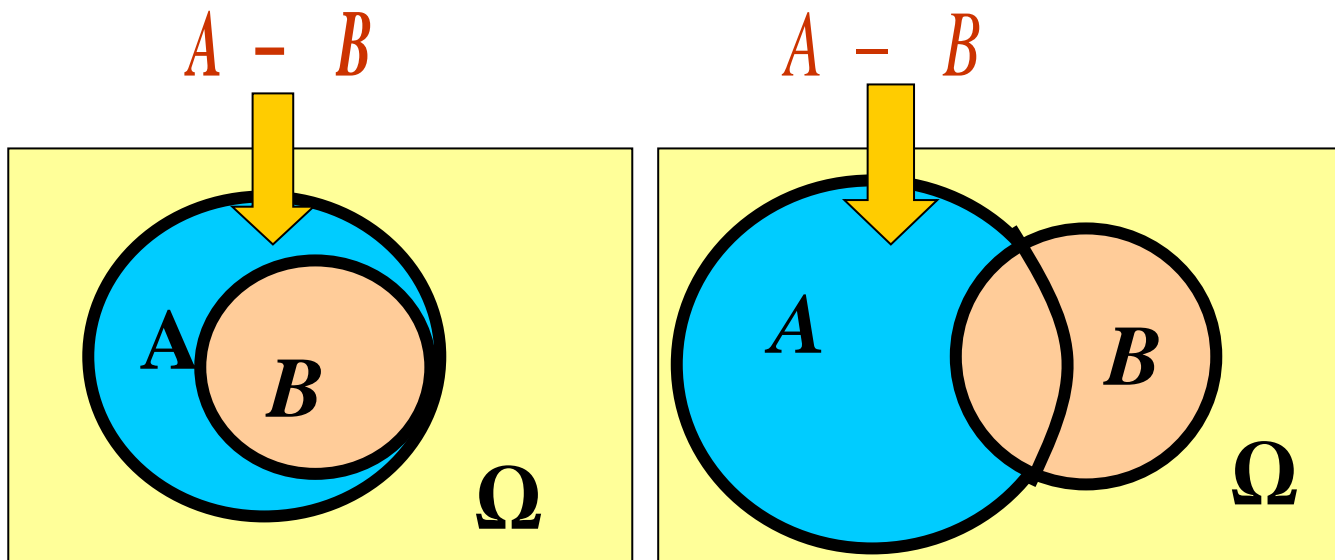
$A_1, A_2, A_3, A_4$

构成一完备事件组

## 7. 事件的差

事件A与事件B的差  $\longleftrightarrow A-B$

- ◆ 事件A发生，而事件B不发生
- ◆ 由属于事件A但不属于事件B的样本点构成



性质：

$$A-B = A - AB$$

小

结

## 概率论

样本空间(必然事件)  $\Omega$

不可能事件  $\Phi$

子事件  $A \subset B$

和事件  $A \cup B$

积事件  $A \cap B$

差事件  $A - B$

对立事件  $\bar{A}$

## 集合论

全集

空集  $\Phi$

子集  $A \subset B$

并集  $A \cup B$

交集  $A \cap B$

差集  $A - B$

补集  $\bar{A}$

## 事件之间的运算律

◆ 交换律  $A \cup B = B \cup A$        $AB = BA$

◆ 结合律  $(A \cup B) \cup C = A \cup (B \cup C)$

◆ 分配律  $A(B \cup C) = (AB) \cup (AC)$

$$A \cup (BC) = (A \cup B)(A \cup C)$$

◆ 摩根律  $\overline{AB} = \overline{A} \cup \overline{B}$        $\overline{A \cup B} = \overline{A} \overline{B}$

可推广  $\overline{\bigcup_i A_i} = \bigcap_i \overline{A_i}, \quad \overline{\bigcap_i A_i} = \bigcup_i \overline{A_i}$



## 例1：书本第七页习题四

设  $\Omega = \{1, 2, \dots, 10\}$ ,  $A = \{2, 3, 5\}$ ,  $B = \{3, 5, 7\}$ ,  $C = \{1, 3, 4, 7\}$ ,  
求下列事件：

$$(1) \overline{\overline{A} \overline{B}}$$

$$(2) \overline{A(\overline{BC})}$$

解：由摩根律

$$(1) \overline{\overline{A} \overline{B}} = \overline{\overline{A}} \cup \overline{\overline{B}} = A \cup B = \{2, 3, 5, 7\}$$

$$\begin{aligned} (2) \overline{A(\overline{BC})} &= \overline{A} \cup (BC) = \{1, 4, 6, 7, 8, 9, 10\} \cup \{3, 7\} \\ &= \{1, 3, 4, 6, 7, 8, 9, 10\} \end{aligned}$$

## 例2：复合事件的表示

某射手向目标射击三次，用 $A_i$ 表示第 $i$ 次击中目标 ( $i=1,2,3$ )，试用 $A_i$  及其运算符表示下列事件：

1. 三次都击中目标： $A_1 A_2 A_3$
2. 至少有一次击中目标： $A_1 \cup A_2 \cup A_3$
3. 恰好有两次击中目标： $\bar{A}_1 A_2 A_3 \cup A_1 \bar{A}_2 A_3 \cup A_1 A_2 \bar{A}_3$
4. 最多击中一次： $\bar{A}_1 \bar{A}_2 \cup \bar{A}_1 \bar{A}_3 \cup \bar{A}_2 \bar{A}_3$
5. 至少有一次没有击中目标： $\bar{A}_1 \cup \bar{A}_2 \cup \bar{A}_3 = \overline{A_1 A_2 A_3}$
6. 三次都没有击中目标： $\bar{A}_1 \bar{A}_2 \bar{A}_3 = \overline{A_1 \cup A_2 \cup A_3}$

## 课堂练习

1. 设事件 $A = \{\text{甲种产品畅销, 乙种产品滞销}\}$ , 则 $A$ 的对立事件为 ( ④ )

①甲种产品滞销, 乙种产品畅销;

②甲、乙两种产品均畅销;

③甲种产品滞销;

④甲种产品滞销或者乙种产品畅销。

2. 设 $x$ 表示一个沿数轴做随机运动的质点位置, 试说明下列各对事件间的关系

① $A = \{|x-a| < \sigma\}$ ,  $B = \{x-a < \sigma\}$  ( $\sigma > 0$ )  **$A \subset B$**

② $A = \{x > 20\}$ ,  $B = \{x \leq 20\}$   **$A$ 与 $B$ 对立**

③ $A = \{x > 22\}$ ,  $B = \{x < 19\}$   **$A$ 与 $B$ 互斥**

# 第二节

第二节

概 率

# 频 率

**定义** 设随机事件A在n次试验中发生了m次，则称比值 $m/n$ 为随机事件A发生的频率，记作 $W(A)$ ，用公式表示如下：

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

**性质**

$$(1) \quad 0 \leq W(A) \leq 1$$

$$(2) \quad W(\Omega) = 1, \quad W(\emptyset) = 0$$

**例** 将一枚硬币抛掷 5 次、50 次、500 次, 各做 7 遍, 观察正面出现的次数及频率.

试验 序号	$n = 5$		$n = 50$		$n = 500$	
	$n_H$	$f$	$n_H$	$f$	$n_H$	$f$
1 2 3 4 5 6 7	2	0.4	22	0.44	251	0.502
	6	0.6	24	0.48	240	0.498
	3	0.2	21	0.42	250	0.512
	5	1.0	25	0.50	247	0.494
	1	0.2	25	0.50	251	0.502
	2	0.4	18	0.36	251	0.502
	4	0.8	27	0.54	258	0.516

随 $n$ 的增大, 频率 $f$ 呈现出稳定性

在 $\frac{1}{2}$ 处波动较小

波动最小

# 抛掷硬币的试验

## 历史纪录

试 验 者	抛掷次数n	正面次数 m	频率m/n
德摩根	2048	1061	0.518
蒲 丰	4040	2048	0.5069
K.皮尔逊	12000	6019	0.5016
K.皮尔逊	24000	12012	0.5005
维 尼	30000	14994	0.4998

$W(\text{正面}) \xrightarrow{n \text{ 的增大}} \frac{1}{2}$



# 频率和概率

## ◆ 频率的稳定性

随机事件A在相同条件下重复多次时，事件A发生的频率在一个固定的数值 $p$ 附近摆动，随试验次数的增加更加明显.

## ◆ 事件的概率

事件A的频率稳定在数值 $p$ ，说明了数值 $p$ 可以用来刻画事件A发生可能性大小，可以规定为事件A的概率.



# 概率的统计定义

对任意事件  $A$ ，在相同的条件下重复进行  $n$  次试验，事件  $A$  发生的频率  $m/n$ ，随着试验次数  $n$  的增大而稳定地在某个常数  $p$  附近摆动那么称  $p$  为事件  $A$  的概率  $P(A) = p$

当试验次数足够大时，可以用事件  $A$  发生的频率近似的代替事件  $A$  的概率

事件发生的  
频繁程度

事件发生的可  
能性的大小

频 率  $\longrightarrow$  稳 定 值  $\longrightarrow$  概 率

# 概率的公理化定义

前面分别介绍了统计概率定义、古典概率及几何概率的定义，它们在解决各自相适应的实际问题中，都起着很重要的作用，但它们各自都有一定局限性。

为了克服这些局限性，1933年，前苏联数学家柯尔莫哥洛夫在综合前人成果的基础上，抓住概率共有特性，提出了概率的公理化定义，为现代概率论的发展奠定了理论基础。

## 概率的公理化的定义:

设  $\Omega$  是给定的实验  $E$  的样本空间, 对其中的任意一

事件  $A$ , 规定一个实数  $P(A)$ , 若  $P(A)$  满足:

(1) 非负性  $0 \leq P(A)$

(2) 规范性  $P(\Omega) = 1$

(3) 可列可加性: 设  $A_1, A_2, \dots, A_n, \dots$  两两互不相容.

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.

# 概率的性质

性质1 不可能事件概率零:  $P(\emptyset) = 0$ .

性质2 如果  $A_1, A_2, \dots, A_n$  两两互不相容, 则有

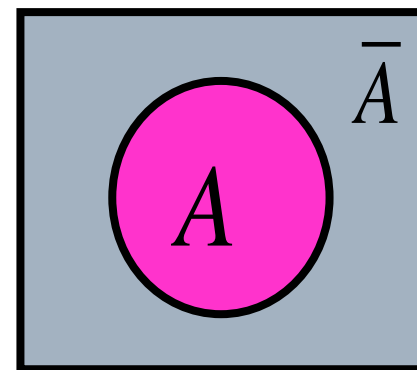
$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n).$$

↓  
概率的有限可加性

特别: (1) 如果  $A_1, A_2, \dots, A_n$  构成完备事件组, 则有

$$P(A_1) + P(A_2) + \dots + P(A_n) = 1$$

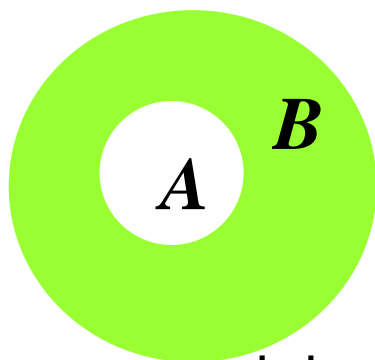
(2)  $P(A) + P(\bar{A}) = 1$ , 常用  $P(\bar{A}) = 1 - P(A)$



**性质3 单调不减性：**若事件  $B \supset A$  ， 则

$$P(B - A) = P(B) - P(A), \quad \text{且} \quad P(A) \leq P(B)$$

证明：



$$B = A \cup (B - A)$$

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

于是

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

又因  $P(B - A) \geq 0$ ，故  $P(A) \leq P(B)$ .

若A，B是任意两个随机事件，  $P(B - A) = P(B) - P(AB)$

性质4 (加法公式) 对于任意两个随机事件A,B有

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

证明: 由图可得



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

$$P(B - AB) = P(B) - P(AB)$$

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

## 推广 三个事件和的情况

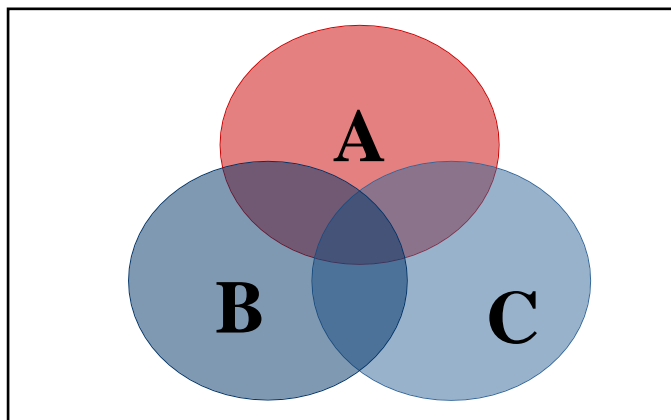
$$P(A \cup B \cup C) = P(A) + P(B) + P(C)$$

$$- P(AB) - P(BC) - P(AC) + P(ABC)$$

$$P(A \cup B \cup C \cup D) = P(A) + P(B) + P(C) + P(D)$$

$$- P(AB) - P(AC) - P(AD) - P(BC) - P(BD) - P(CD)$$

$$+ P(ABC) + P(ABD) + P(ACD) + P(BCD) - P(ABCD)$$



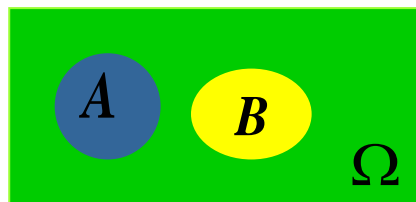
$n$ 个事件和的情况

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

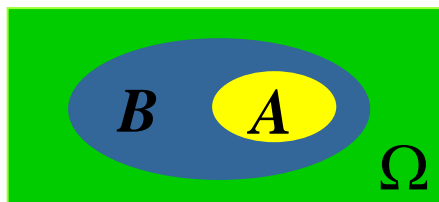


**例1:** 设事件A, B的概率分别为1/3和1/2, 求在下列三种情况下  $P(B\bar{A})$  的值

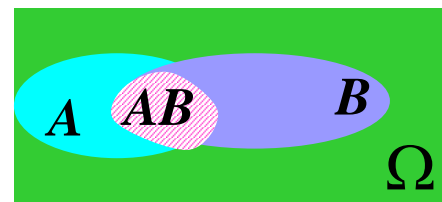
(1) A与B互斥;



(2)  $A \subset B$ ;



(3)  $P(AB)=1/8$



**解:**  $P(B\bar{A}) = P(B - AB) = P(B) - P(AB)$

(1) 当A与B互斥时  $P(B\bar{A}) = P(B) - P(AB) = P(B) = 1/2$

(1) 当A包含于B时  $P(B\bar{A}) = P(B) - P(AB) = P(B) - P(A) = 1/6$

(2) 当A与B相交时  $P(B\bar{A}) = P(B) - P(AB) = 3/8$

例2: 设  $A$ 、 $B$ 、 $C$  是三事件, 且  $P(A) = P(B) = P(C) = \frac{1}{4}$ ,  
 $P(AC) = \frac{1}{8}$ ,  $P(AB) = P(BC) = 0$ , 求  $A$ 、 $B$ 、 $C$  至少有一个发生的概率.

$$\begin{aligned} \text{解: } & P(A \cup B \cup C) \\ &= P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC) \\ &= 3 \cdot \frac{1}{4} - \frac{1}{8} + 0 = \frac{5}{8} \end{aligned}$$

**例3:** A、B都出现的概率与 A、B 都不出现的概率相等,  $P(A)=p$ , 求  $P(B)$ .

**解:**

$$\begin{aligned} P(B) &= P(AB) + P(\overline{A}B) \\ &= P(\overline{A}\overline{B}) + P(\overline{A}B) \\ &= P(\overline{A}) = 1 - P(A) \end{aligned}$$

所以,  $P(B) = 1 - P(A) = 1 - p$

# 第三节

第三节

基本概型

# 古典概率模型(等可能概型)

## ◆ 有限性

每次试验中, 所有可能发生的结果只有有限个,  
即样本空间 $\Omega$ 是个有限集

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$$

## ◆ 等可能性

每次试验中, 每一种可能结果的发生的可能性相同, 即

$$P(A_1) = P(A_2) = \dots = P(A_n) = \frac{1}{n}$$

$$\text{其中 } A_i = \{\omega_i\}, \quad i = 1, 2, \dots, n$$

# 古典概型的概率计算

## ◆ 确定试验的基本事件总数

设试验结果共有 $n$ 个基本事件 $\omega_1, \omega_2, \dots, \omega_n$ ，  
而且这些事件的发生具有相同的可能性

## ◆ 确定事件 $A$ 包含的基本事件数

事件 $A$ 由其中的 $m$ 个基本事件组成

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

**【注】** 求解古典概型问题的关键是弄清样本空间中的基本事件总数和事件A中包含的基本事件数. 在考虑事件数的时候, 必须分清研究的问题是组合问题还是排列问题, 掌握以下关于排列组合的知识是有用的:

**(1) 加法原理:** 设完成一件事有 $k$ 类方法, 每类又分别有 $m_1, m_2, \dots, m_k$ 种方法, 而完成这件事只需其中一种方法, 则完成这件事共有 $m_1 + m_2 + \dots + m_k$ 种方法.

**(2) 乘法原理:** 设完成一件事有 $n$ 个步骤. 第一步有 $m_1$ 种方法、第二步有 $m_2$ 种方法, ...第 $n$ 步有 $m_n$ 种方法, 则完成这件事共有 $m_1 \times m_2 \times \dots \times m_n$ 种方法.

### (3) 不同元素的组合

从含有 $n$ 个不同元素的集合中随机抽取 $k$ 个，共有

$$C_n^k \equiv \binom{n}{k} = \frac{A_n^k}{k!} = \frac{n!}{k!(n-k)!} \quad \text{种取法.}$$

### (4) 不同元素的全排列

无重复排列：从含有 $n$ 个不同元素的集合中随机抽取 $k$ 次，每次取一个，取后不放回，将所取元素排成一列，共有

$$A_n^k = n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!} \quad \text{种排列方式.}$$

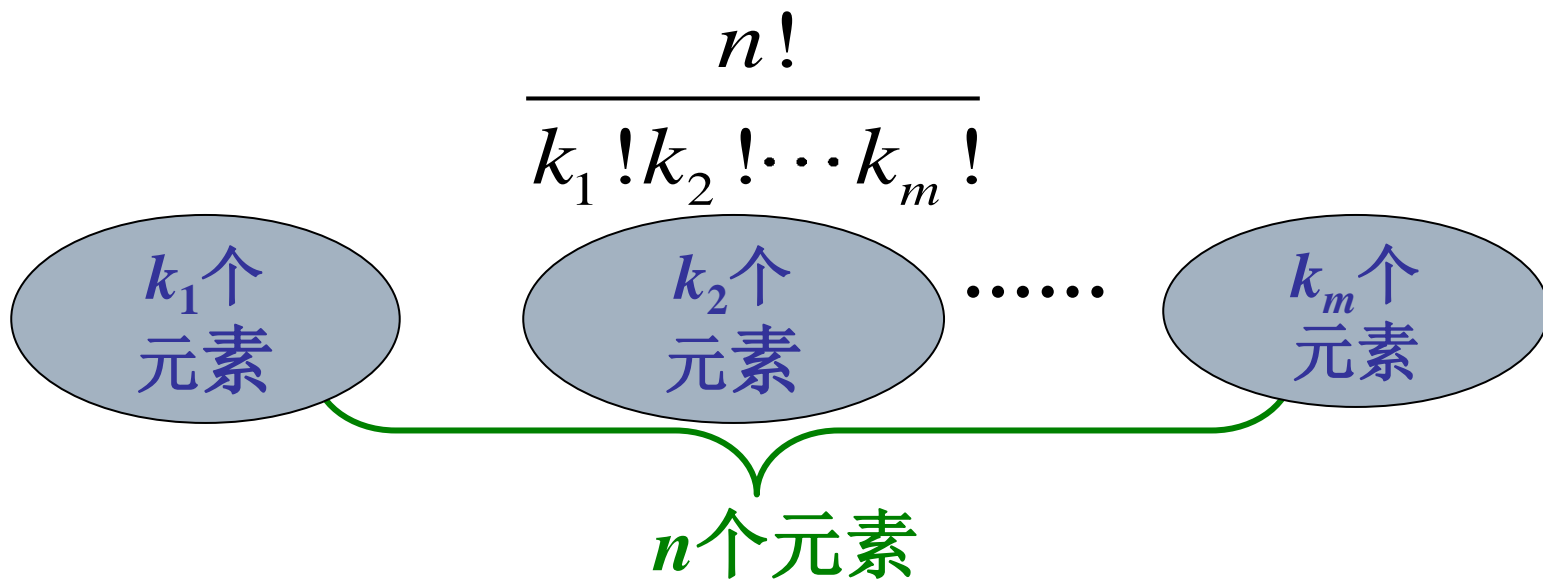


## (5) 不同元素的重复排列

从 $n$ 个不同的元素中，有放回地取 $k$ 个元素进行的排列，共有 $n^k$ 种(元素允许重复 $1 \leq k \leq n$ )。

## (6) 不全相异元素的排列

在 $n$ 个元素中，有 $m$ 类不同元素、每类各有 $k_1, k_2, \dots, k_m$ 个，将这 $n$ 个元素作全排列，共有如下种方式：



# 古典概率的计算

**例1** 将10本不同的书任意地放在书架上，求指定的3本书放在一起的概率。

**解：** 设 $A=\{\text{指定的3本书放在一起}\}$ ，则所求概率为

$$P(A) = \frac{m}{n} = \frac{8!3!}{10!} = \frac{1}{15}$$

**例2** 在1~2000的整数中随机地取一个数,问取到的整数

(1) 能被6或8整除的概率是多少?

(2) 既不能被6整除,又不能被8整除的概率是多少?

**解:** 设  $A$  为事件“取到的数能被6整除”,

$B$ 为事件“取到的数能被8整除”, 则

(1) 所求概率  $P(A \cup B) = P(A) + P(B) - P(AB)$

又因  $P(A) = \frac{333}{2000}, \quad P(B) = \frac{250}{2000}, \quad P(AB) = \frac{83}{2000}.$

为

于是所求概率为  $P(A \cup B) = \frac{333}{2000} + \frac{250}{2000} - \frac{83}{2000} = \frac{1}{4}.$

(2)所求概率为  $P(\overline{A \cup B})$ .

$$P(\overline{A \cup B}) = P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 - \frac{1}{4} = \frac{3}{4}$$

**例3** 在桥牌比赛中，把52张牌任意地发给东南西北四家，每家13张，求北家分到的13张牌中

- (1) 红心不多于2张的概率是多少？
- (2) 至少缺一种花色的概率是多少？
- (3) 四种花色都有的概率是多少？

**解：** (1) 设  $A=\{\text{红心不多于2张}\}$ ，则 所求概率为

$$P(A) = \frac{C_{39}^{13} + C_{13}^1 C_{39}^{12} + C_{13}^2 C_{39}^{11}}{C_{52}^{13}}$$

(2) 至少缺  $P(\text{四种花色都有}) = 1 - P(\text{至少缺一种花色})$

(3) 四种花  $= 1 - \frac{4C_{39}^{13} - 6C_{26}^{13} + 4}{C_{52}^{13}}$

解(2) 设

$A_1 = \{\text{缺黑桃}\}, A_2 = \{\text{缺红心}\}, A_3 = \{\text{缺草花}\}, A_4 = \{\text{缺方块}\}$

则  $P(\text{至少缺一种花色}) = P(A_1 \cup A_2 \cup A_3 \cup A_4)$

$$= \sum_{i=1}^4 P(A_i) - \sum_{i \neq j} P(A_i A_j) + \sum_{i \neq j \neq k} P(A_i A_j A_k) - P(A_1 A_2 A_3 A_4)$$

$$= C_4^1 \frac{C_{39}^{13}}{C_{52}^{13}} - C_4^2 \frac{C_{26}^{13}}{C_{52}^{13}} + C_4^3 \frac{C_{13}^{13}}{C_{52}^{13}} - 0 = \frac{4C_{39}^{13} - 6C_{26}^{13} + 4}{C_{52}^{13}}$$

# 几何概型 Geometric Probability

◆ 将古典概型中的有限性推广到无限性，而保留等可能性，就得到几何概型。

◆ 特点

- 有一个可度量的几何图形 $\Omega$
- 试验E看成在 $\Omega$ 中随机地投掷一点
- 事件A就是所投掷的点落在 $\Omega$ 中的可度量图形A中

$$P(A) = \frac{A \text{ 的几何度量}}{\Omega \text{ 的几何度量}} = \frac{L(A)}{L(\Omega)}$$

◆ 几何度量-----指长度、面积或体积

# 几何概型的计算

一个质地均匀的陀螺的圆周上均匀地刻有 $[0, 5)$ 上诸数字，在桌面上旋转它，求当它停下来时，圆周与桌面接触处的刻度位于区间 $[2, 3]$ 上的概率。

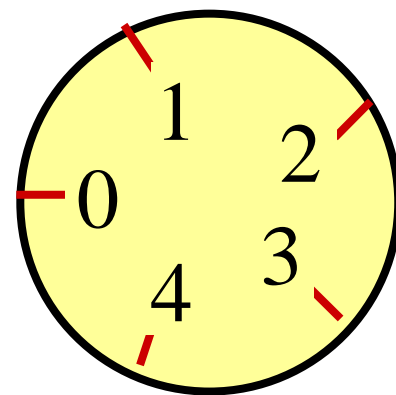
$$\Omega = [0, 5)$$

$$A = [2, 3]$$

$$L(\Omega) = 5 - 0 = 5$$

$$L(A) = 3 - 2 = 1$$

$$P(A) = \frac{L(A)}{L(\Omega)} = \frac{1}{5}$$



## 会面问题

**例1** 甲、乙两人相约在 0 到  $T$  这段时间内, 在预定地点会面. 先到的人等候另一个人, 经过时间  $t$  ( $t < T$ ) 后离去. 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牵连. 求甲、乙两人能会面的概率.

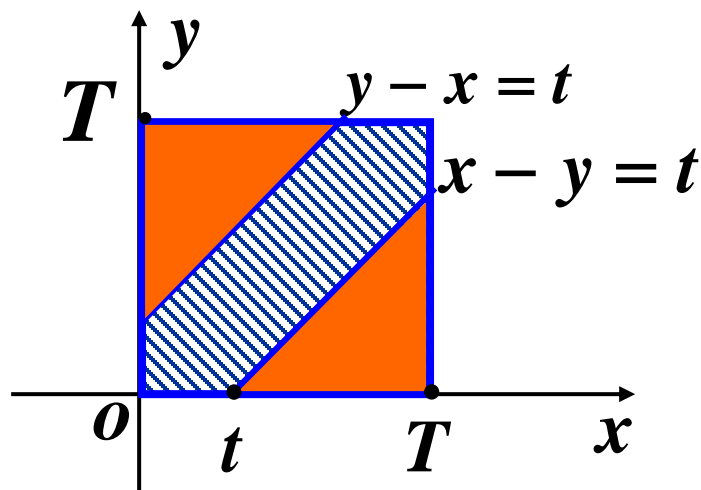
**解:** 设  $x, y$  分别为甲、乙两人到达的时刻, 那么  $0 \leq x \leq T, 0 \leq y \leq T$ .  
两人会面的充要条件为  $|x - y| \leq t$ ,



若以  $x, y$  表示平面上点的坐标, 则有  
故所求的概率为

$$p = \frac{\text{阴影部分面积}}{\text{正方形面积}}$$

$$= \frac{T^2 - (T - t)^2}{T^2} = 1 - \left(1 - \frac{t}{T}\right)^2.$$



**例2** 两船欲停靠同一个码头，设两船到达码头的时  
间各不相干，而且到达码头的时在一昼夜内是等  
可能的. 如果两船到达码头后需在码头停留的时间分  
别是2 小时与1小时，试求在一昼夜内，任一船到达  
时，需要等待空出码头的概率.

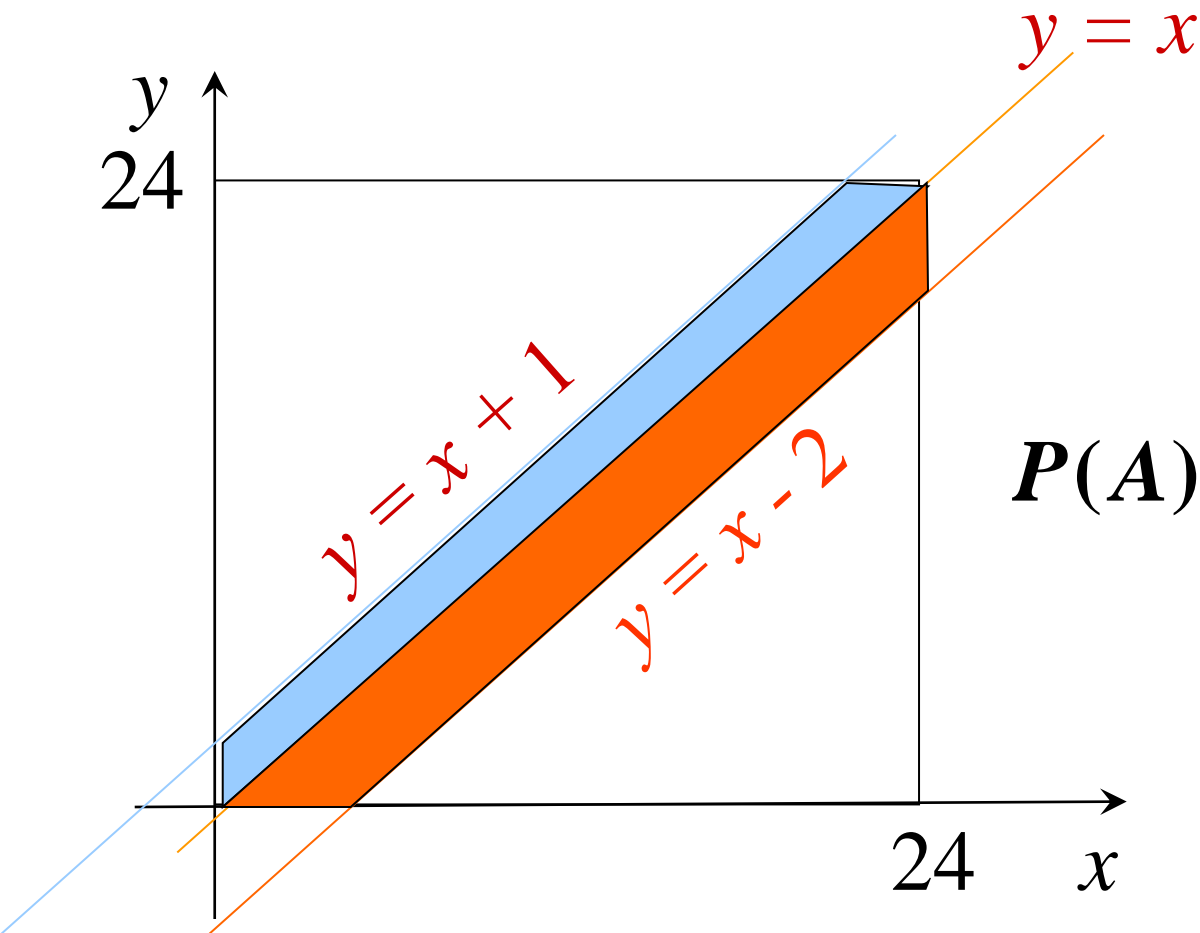
**解：** 设船1 到达码头的瞬时为  $x$  ,  $0 \leq x < 24$   
船2 到达码头的瞬时为  $y$  ,  $0 \leq y < 24$

设事件  $A$  表示任一船到达码头时需要等待  
空出码头.

$$\Omega = \{(x, y) \mid 0 \leq x < 24, 0 \leq y < 24\}$$

$$A = \{(x, y) \mid (x, y) \in \Omega,$$

$$0 \leq y - x \leq 1, 0 \leq x - y \leq 2\}$$



$$S_{\Omega} = 24^2$$

$$S_{\bar{A}} = \frac{1}{2}(23^2 + 22^2)$$

$$P(A) = 1 - \frac{S_{\bar{A}}}{S_{\Omega}} = 0.1207$$

# 小 结

最简单的随机现象 → 古典概型 <sup>试验结果  
连续无穷</sup> → 几何概型

↓  
古典概率  
↓

$$P(A) = \frac{m}{n} = \frac{A \text{ 所包含基本事件数}}{\text{基本事件总数}}$$

## 第四节

# 乘法公式与全概率公式

# 条件概率

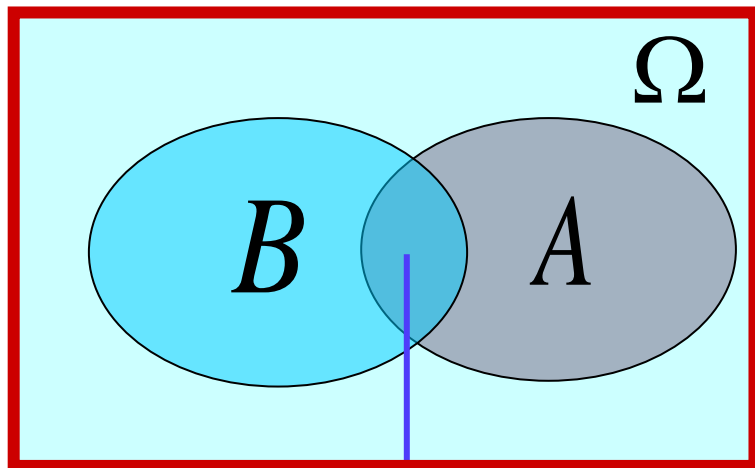
## ■ 定义

设A，B为同一个随机试验中的两个随机事件，  
且 $P(B) > 0$ ，则称

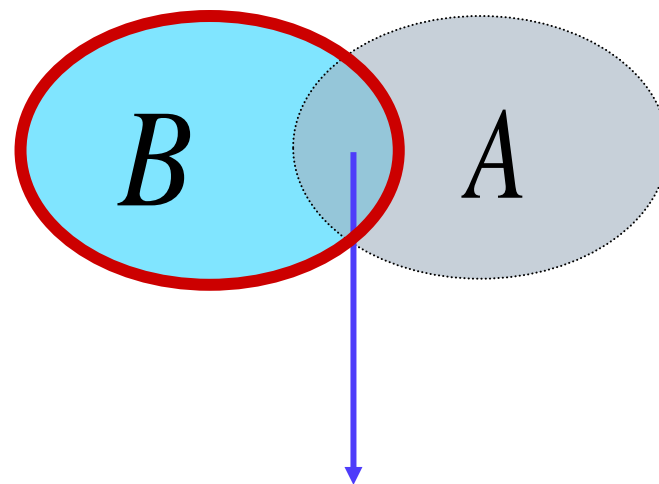
$$P(A|B) = \frac{P(AB)}{P(B)}$$

为在事件B发生的条件下，事件A发生的条件概率。

# 条件概率 $P(A|B)$ 的样本空间



$P(AB)$   
在 $\Omega$ 中去考虑 $P(AB)$



$P(A|B)$   
在 $B$ 中去考虑 $P(A|B)$

## 乘法定理

$$\begin{aligned} P(AB) &= P(A)P(B|A) &< \leftarrow P(B|A) = \frac{P(AB)}{P(A)} \\ &= P(B)P(A|B) &< \leftarrow P(A|B) = \frac{P(AB)}{P(B)} \end{aligned}$$

■ 推广

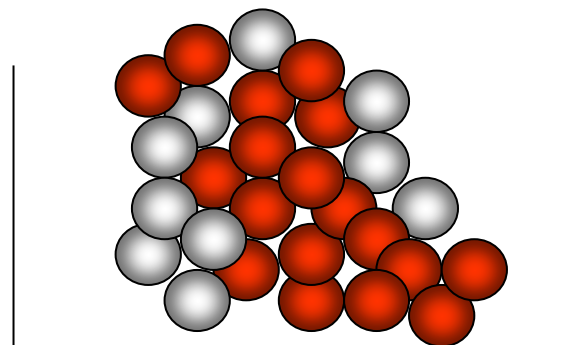
$$P(ABC) = P(A)P(B|A)P(C|AB)$$

$$\begin{aligned} P(A_1 A_2 \cdots A_n) &= P(A_1)P(A_2|A_1)P(A_3|A_1 A_2) \\ &\quad \cdots P(A_n|A_1 A_2 \cdots A_{n-1}) \end{aligned}$$



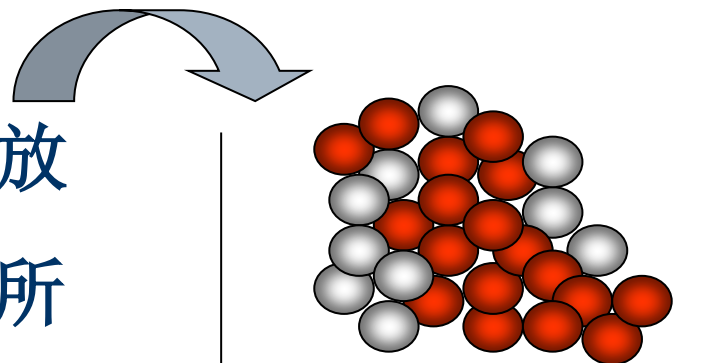
**例：**盒中有3个红球2个白球，每次从盒中任取一只，观察其颜色后放回，并再放入一只与所取之球颜色相同的球，若从盒中连续取球4次,试求第1、2次取得白球、第3、4次取得红球的概率。

波里亚罐子模型



$b$ 个白球,  $r$ 个红球

随机取一个球，观看颜色后放回罐中，并且再加进  $c$  个与所抽出的球具有相同颜色的球。



$b$  个白球,  $r$  个红球

当  $c > 0$  时，由于每次取出球后会增加下一次也取到同色球的概率。这是一个**传染病模型**。每次发现一个传染病患者，都会增加再传染的概率。

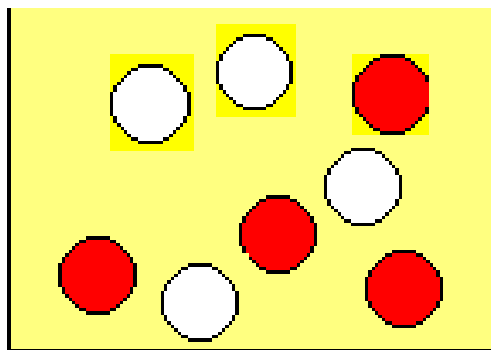
**例** 盒中有3个红球2个白球，每次从盒中任取一只，观察其颜色后放回，并再放入一只与所取之球颜色相同的球，若从盒中连续取球4次，试求第1、2次取得白球、第3、4次取得红球的概率。

**解：** 设 $A_i$ 为第 $i$ 次取球时取到白球，则

$$P(A_1 A_2 \bar{A}_3 \bar{A}_4) = P(A_1)P(A_2 | A_1)P(\bar{A}_3 | A_1 A_2)P(\bar{A}_4 | A_1 A_2 \bar{A}_3) = \frac{3}{70}$$

$$P(A_1) = \frac{2}{5}$$

$$P(A_2 | A_1) = \frac{3}{6}$$



$$P(\bar{A}_3 | A_1 A_2) = \frac{3}{7}$$

$$P(\bar{A}_4 | A_1 A_2 \bar{A}_3) = \frac{4}{8}$$

# 全概率公式与贝叶斯公式

全概率公式和贝叶斯公式主要用于计算比较复杂事件的概率, 它们实质上是加法公式和乘法公式的综合运用。

综合运用

加法公式

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

$A$ 、 $B$ 互斥

乘法公式

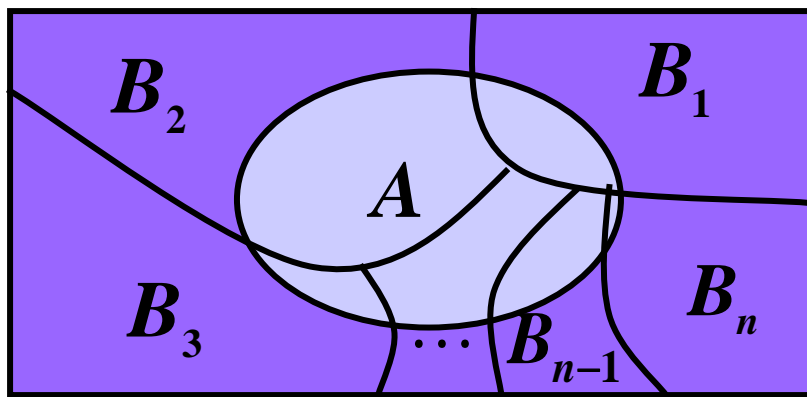
$$P(AB) = P(A)P(B|A)$$

$$P(A) > 0$$

证明

$$\begin{aligned} P(A) &= P(A\Omega) = P\{A \cap (B_1 \cup B_2 \cup \cdots \cup B_n)\} \\ &= P(AB_1 \cup AB_2 \cup \cdots \cup AB_n) \\ &= P(AB_1) + P(AB_2) + \cdots + P(AB_n) \\ &= P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \cdots + P(B_n)P(A|B_n). \end{aligned}$$

$\Omega$

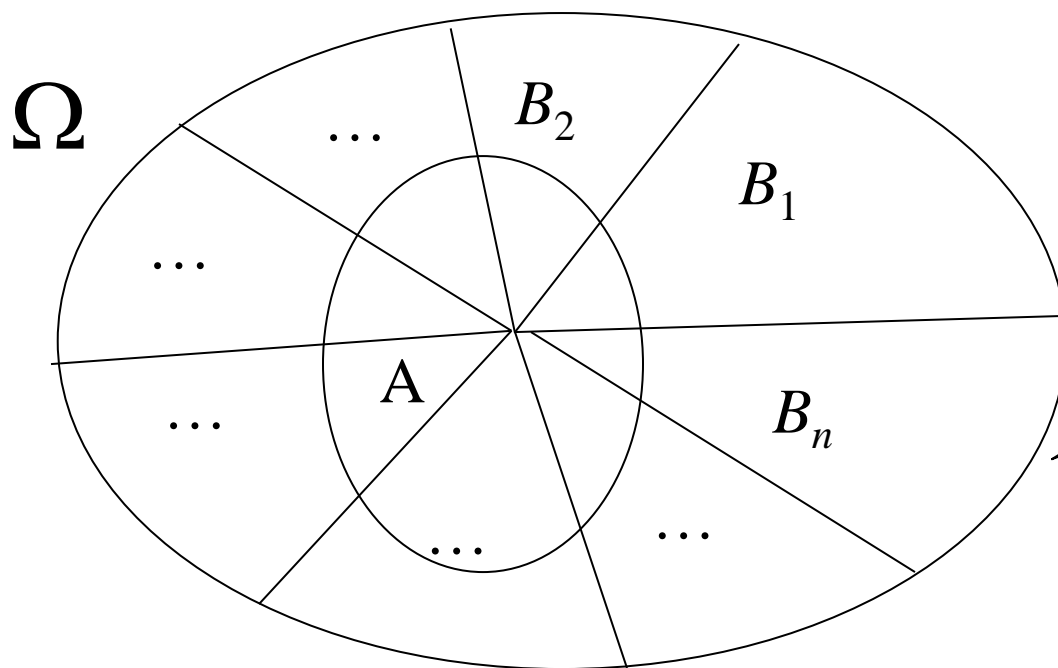


即  $A = AB_1 \cup AB_2 \cup \dots \cup AB_n$ ,  
且  $AB_1, AB_2, \dots, AB_n$  两两互斥

# 全概率公式

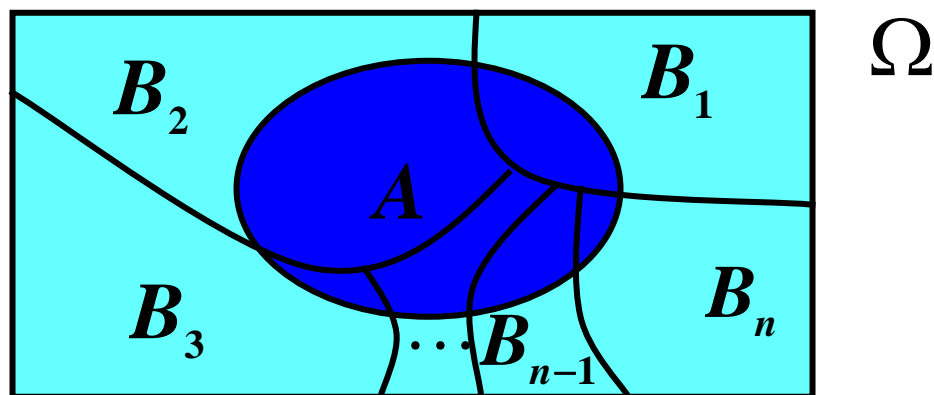
设 $B_1, B_2, \dots, B_n$  构成一个完备事件组, 且  $P(B_i) > 0$ ,  $i=1, 2, \dots, n$ , 则对任一随机事件 $A$ , 有

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



$A$ 发生总是伴随着 $B_1, B_2, \dots, B_n$ 之一同时发生

**说明** 全概率公式的主要用处在于它可以将一个复杂事件的概率计算问题 $P(A)$ ，分解为若干个简单事件的概率计算问题 $P(AB_i)$ ，最后应用条件概率求出最终结果.



**例：**有一批同一型号的产品，已知其中由一厂生产的占30%，二厂生产的占50%，三厂生产的占20%，又知这三个厂的产品次品率分别为2%，1%，1%，问从这批产品中任取一件是次品的概率是多少？

**解：**设 $A=\{ \text{任取一件产品为次品} \}$ ,

$B_i = \{ \text{任取一件为} i \text{厂生产的产品} \}, i = 1, 2, 3.$

由全概率公式得

$$\begin{aligned} P(A) &= P(A\Omega) = P\{A \cap (B_1 \cup B_2 \cup B_3)\} \\ &= P(AB_1 \cup AB_2 \cup AB_3) \\ &= 0.3 \times 0.02 + 0.5 \times 0.01 + 0.2 \times 0.01 = 0.013. \\ &= P(AB_1) + P(AB_2) + P(AB_3) \\ &= P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3) \end{aligned}$$





已知某种疾病的发病率为0.1%，该种疾病患者一个月以内的死亡率为90%；且知未患该种疾病的人一个月以内的死亡率为0.1%；现从人群中任意抽取一人，问此人在一个月內死亡的概率是多少？

解：设 $A=\{\text{某人在一个月內死亡}\}$ ，

$B=\{\text{某人患有该种疾病}\}$

则  $P(A) = P(AB) + P(A\bar{B})$

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

$$= 0.001 \times 0.9 + 0.999 \times 0.001 \approx 0.002$$

若已知此人在一个月內死亡，则此人是因该种疾病致死的概率为多少？

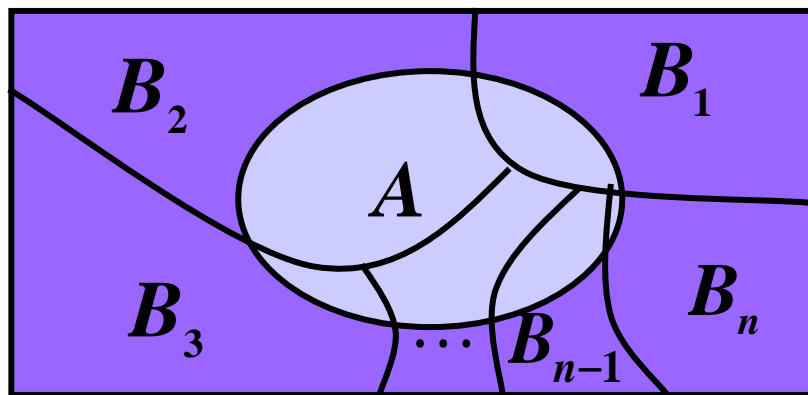
$$\begin{aligned} \text{则 } P(B|A) &= \frac{P(AB)}{P(A)} = \frac{P(B) \times P(A|B)}{P(B) \times P(A|B) + P(\bar{B}) \times P(A|\bar{B})} \\ &= \frac{0.001 \times 0.9}{0.001 \times 0.9 + 0.999 \times 0.001} = 0.47 \end{aligned}$$

# 贝叶斯公式 Bayes' Theorem

设 $B_1, B_2, \dots, B_n$ 构成样本空间的一个完备事件组， $A$ 为任意事件，则有

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

$\Omega$



$P(B_k)$ -----先验概率

$P(B_k/A)$ -----后验概率

**例** 玻璃杯论箱出售，每箱20只，其中每箱含0,1,2只次品的概率分别为0.8,0.1,0.1，一顾客欲购一箱玻璃杯，在购买时售货员随意取一箱，而顾客随机地查看4只，若无次品，则买下该箱玻璃杯，否则退回。试求：(1)顾客买下该箱的概率；  
(2)在顾客买下的一箱中，确实没有次品的概率。

解:设 $A=\{\text{顾客买下一箱玻璃杯}\}$ ,  $B_i=\{\text{每箱含}i\text{只次品}\}$ ,  $i=0,1,2$ .

$$\begin{aligned}P(A) &= P(AB_0) + P(AB_1) + P(AB_2) \\&= P(B_0)P(A|B_0) + P(B_1)P(A|B_1) + P(B_2)P(A|B_2) \\&= 0.8 \times 1 + 0.1 \times \frac{C_{19}^4}{C_{20}^4} + 0.1 \times \frac{C_{18}^4}{C_{20}^4} = \frac{448}{475}\end{aligned}$$

$$P(B_0|A) = \frac{P(AB_0)}{P(A)} = \frac{P(B_0)P(A|B_0)}{P(A)} = \frac{0.8 \times 1}{\frac{448}{475}} = \frac{95}{112}$$



在1, 2, 3号箱中装有同型号的合格产品分别为20件、12件、15件, 不合格品分别为5件、4件、5件。现任意打开一箱并从中任取一件检验, 由于检验的误差, 每个合格品被检验为不合格品的概率为0.04, 而不合格品被检验为合格品的概率为0.06, 求下列概率:

- (1) 取到一件产品被验定为合格品;
- (2) 若已知取到的一件产品被验定为合格品, 那么它确实是合格品。

答案:

- (1) 全概论公式求解得到: 0.75
- (2) 贝叶斯公式求解得到: 0.98

## 小结

条件概率  $P(B|A) = \frac{P(AB)}{P(A)}$   $\longrightarrow$  乘法定理

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

全概率公式

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \cdots + P(B_n)P(A|B_n)$$

贝叶斯公式

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

# 第五节

第五节

事件的独立性

# 事件的独立性 independence

## ■ 定义

设A,B为任意两个随机事件，如果

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

即事件B发生的可能性不受事件A的影响，则称事件B对于事件A独立.

显然，B对于A独立，则A对于B也独立，故称A与B相互独立.

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

# 事件的独立性判别

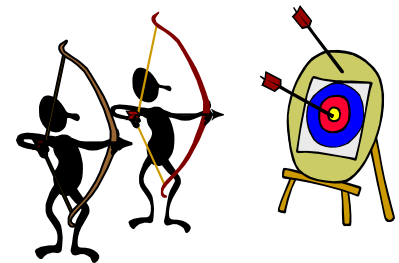
- 事件 A 与事件 B 独立的充分必要条件是

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

**注：**必然事件 $\Omega$ 及不可能事件 $\Phi$ 与任何事件A均独立。

- 实际问题中，事件的独立性可根据问题的实际意义来判断

如甲乙两人射击，“甲击中”与“乙击中”可以认为相互之间没有影响，即可以认为相互独立。





又如： 一批产品共 $n$ 件，从中抽取2件，设  
 $A_i = \{\text{第}i\text{件是合格品}\} \quad i=1,2$

若抽取是有放回的，则 $A_1$ 与 $A_2$ 独立.

因为第二次抽取的结果  
不受第一次抽取的影响.

若抽取是无放回的，则 $A_1$ 与 $A_2$ 不独立.

因为第二次抽取的结果受到第一次 抽取的影响.

■ **结论** 下列四组事件，有相同的独立性：

(1)  $A$ 与 $B$ ;      (2)  $A$ 与 $\bar{B}$ ;

(3)  $\bar{A}$ 与 $B$ ;      (4)  $\bar{A}$ 与 $\bar{B}$

证明：若 $A$ 、 $B$ 独立，则  $P(AB) = P(A) \cdot P(B)$

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

$$= 1 - [P(A) + P(B) - P(AB)]$$

$$= 1 - P(A) - P(B) + P(A)P(B)$$

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

所以， $\bar{A}$ 与 $\bar{B}$  独立。

## ■ 概念辨析

事件 A 与事件 B 独立

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

事件 A 与事件 B 互不相容

$$AB = \Phi \quad P(AB) = 0$$

二者之间没有必然联系

事件 A 与事件 B 为对立事件

$$AB = \Phi \quad A \cup B = \Omega \quad P(A) + P(B) = 1$$

**例：**甲乙二人向同一目标射击，甲击中目标的概率为0.6，乙击中目标的概率为0.5。试计算：

- 1) 两人都击中目标的概率；
- 2) 恰有一人击中目标的概率；
- 3) 目标被击中的概率。

**解：**设 $A = \{\text{甲击中目标}\}$ ,  $B = \{\text{乙击中目标}\}$ ，则 $A$ 与 $B$ 独立，且有  $P(A) = 0.6, P(B) = 0.5$

$$P(AB) = P(A)P(B) = 0.6 \times 0.5 = 0.3$$

$$P(\bar{A}B \cup A\bar{B}) = P(\bar{A})P(B) + P(A)P(\bar{B}) = 0.5$$

$$P(A \cup B) = P(A) + P(B) - P(AB) = 0.8$$

# 有限多个事件的独立性

**定义** 若三个事件A、B、C满足：

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

$$P(AC)=P(A)P(C),$$

$$P(BC)=P(B)P(C),$$

则称事件A、B、C**两两相互独立**；

若在此基础上还满足：

(2)  $P(ABC)=P(A)P(B)P(C),$

则称事件A、B、C**相互独立**。

注意

事件A，B，C相互独立与事件A，B，C两两独立不同，相互独立一定两两独立，但反之不一定。

# 定义

设 $A_1, A_2, \dots, A_n$ 为 $n$ 个事件。如果对于所有可能的组合  
 $1 \leq i < j < k < \dots \leq n$ 下列各式同时成立

[illegible]

那么称 $A_1, A_2, \dots, A_n$ 是相互独立的。

共有  $(2^n - n - 1)$  个等式

对满足相互独立的多个事件，有

(1) 若  $A_1, A_2, \dots, A_n$  相互独立，则

$$P(A_1 A_2 \cdots A_n) = P(A_1)P(A_2) \cdots P(A_n) = \prod_{i=1}^n P(A_i)$$

(2) 若  $A_1, A_2, \dots, A_n$  相互独立，则

$$P\left(\bigcup_{i=1}^n A_i\right) = 1 - P(\overline{A_1} \overline{A_2} \cdots \overline{A_n}) = 1 - \prod_{i=1}^n P(\overline{A_i})$$

**例：**某工人照看三台机床，一个小时内三台机床需要照看的概率分别为0.3,0.2,0.1。设各机床之间是否需要照看是相互独立的，求在一小时内：1) 没有一台机床需要照看的概率；2) 至少有一台不需要照看的概率；3) 至多有一台需要照看的概率。

**解：**设 $A_i=\{\text{第}i\text{台机床需要照看}\}$ ， $i=1,2,3$ ，则 $A_1, A_2, A_3$ 相互独立，且有  $P(A_1)=0.3, P(A_2)=0.2, P(A_3)=0.1$ ,

$$P(\overline{A_1}\overline{A_2}\overline{A_3}) = P(\overline{A_1})P(\overline{A_2})P(\overline{A_3}) = 0.7 \times 0.8 \times 0.9 = 0.504$$

$$P(\overline{A_1} \cup \overline{A_2} \cup \overline{A_3}) = 1 - P(A_1 A_2 A_3)$$

$$= 1 - P(A_1)P(A_2)P(A_3) = 1 - 0.3 \times 0.2 \times 0.1 = 0.994$$

$$P(A_1 \overline{A_2} \overline{A_3} \cup \overline{A_1} A_2 \overline{A_3} \cup \overline{A_1} \overline{A_2} A_3 \cup \overline{A_1} \overline{A_2} \overline{A_3})$$

$$= P(A_1 \overline{A_2} \overline{A_3}) + P(\overline{A_1} A_2 \overline{A_3}) + P(\overline{A_1} \overline{A_2} A_3) + P(\overline{A_1} \overline{A_2} \overline{A_3}) = 0.902$$



**例：** 若干人独立地向一游动目标射击，每人击中目标的概率都是0.6，问至少需要多少人，才能以0.99以上的概率击中目标？

**解：** 设至少需要 $n$ 个人，

设 $A_i = \{\text{第}i\text{个人击中目标}\}$ ， $i=1,2,\dots,n$ ，则 $A_1, A_2, \dots, A_n$  相互独立，且有  $P(A_i) = 0.6, i = 1, 2, \dots, n$

$$\begin{aligned} P(\text{击中目标}) &= P\left(\bigcup_{i=1}^n A_i\right) = 1 - P\left(\bigcap_{i=1}^n \bar{A}_i\right) \\ &= 1 - \prod_{i=1}^n P(\bar{A}_i) = 1 - 0.4^n > 0.99 \quad \text{即} \quad n > \frac{\ln 0.01}{\ln 0.4} \approx 5.026 \end{aligned}$$

所以至少需要6人。

**练习：**  $n$ 对夫妇排成一列，求对于每对夫妇而言妻子总是排在丈夫前方的概率。

$$\text{答案:} \left(\frac{1}{2}\right)^n$$

# 主要公式

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

**A、B无任何条件**

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

**A、B互斥**

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

**A、B无任何条件**

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

**A、B独立**

## 第二章 随机变量及其分布

- 离散型随机变量及其分布律
- 连续型随机变量及其分布
- 随机变量函数的分布

# 随机变量及其分布

## Random Variable and Distribution

在前面的学习中,我们用字母A、B、C...表示事件,并视之为样本空间 $\Omega$ 的子集;针对古典概型,主要研究了用排列组合手段计算事件的概率。

本章,将用随机变量表示随机事件,以便采用高等数学的方法描述、研究随机现象。

# 第一节

第一节

随机变量

# 随机变量 Random Variable

## ■基本思想

将样本空间数量化, 即用数值来表示试验的结果

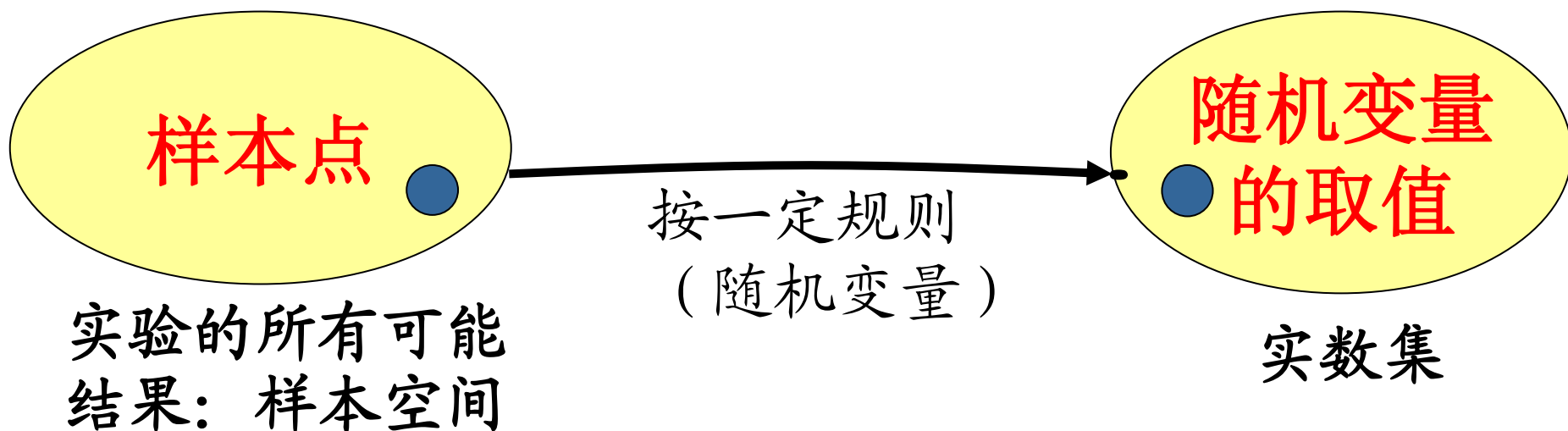
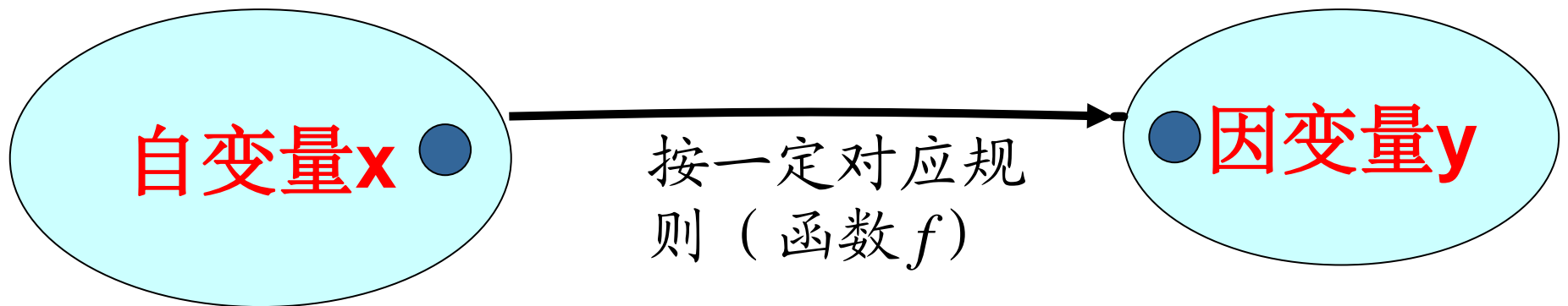
- 有些随机试验的结果可直接用数值来表示.

例如: 在掷骰子试验中, 结果可用1,2,3,4,5,6来表示

- 有些随机试验的结果不是用数量来表示, 但可数量化

例如: 掷硬币试验, 其结果是用汉字“正面”和“反面”来表示的  
可规定: 用 1 表示 “正面朝上” 用 0 表示 “反面朝上”





# 试验结果的数量化

**例：**设箱中有10个球，其中有2个红球，8个白球；从中任意抽取2个，观察抽球结果。

取球结果为：两个白球；两个红球；一红一白

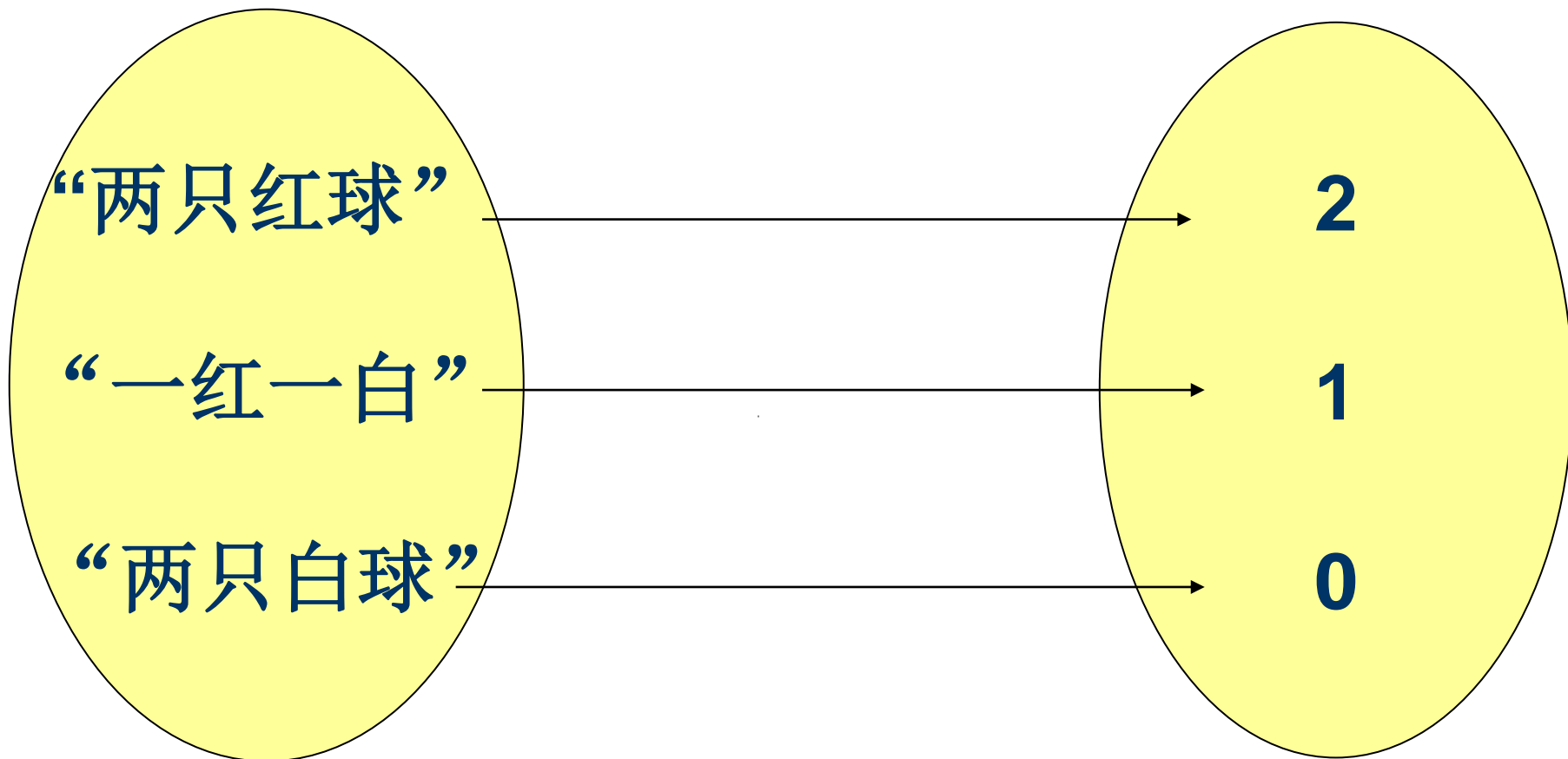
用**X**表示取得的红球数，则**X**的取值可为**0,1,2**. 此时：

“两只红球” 记为 $\{X=2\}$

“一红一白” 记为 $\{X=1\}$

“两只白球” 记为 $\{X=0\}$

**特点：**试验结果数量化了，试验结果与数建立了对应关系



实验的所有可能  
结果：样本空间

按一定规则：  
随机变量  $X$  表示  
取得的红球数

实数集：  
随机变量  $X$   
的取值

# 随机变量的定义

## ■ 随机变量

设随机试验的样本空间为  $\Omega$ ，如果对于每一个样本点  $\omega \in \Omega$ ，均有唯一的实数  $X(\omega)$  与之对应，称  $X = X(\omega)$  为样本空间  $\Omega$  上的随机变量, 简记为 *r.v.*

## ■ 随机变量的三个特征:

- 1) 它是一个变量
- 2) 它的取值随试验结果而改变
- 3) 随机变量在某一范围内取值，表示一个随机事件

# 随机变量的实例

## ■ 例

➤ 某个灯泡的使用寿命 $X$ 。

$X$ 的可能取值为  $[0, +\infty)$

➤ 某电话总机在一分钟内收到的呼叫次数 $Y$ 。

$Y$ 的可能取值为  $0, 1, 2, 3, \dots$ ,

➤ 检查产品质量时，是合格品还是不合格品。

令  $X = \begin{cases} 0, & \text{取到不合格品} \\ 1, & \text{取到合格品} \end{cases}$ ，则 $X$ 也是一个随机变量

# 随机变量的分类

随机变量

离散型

非离散型

连续型

混合型

(1) 离散型 随机变量所取的可能值是有限多个或无限可列个，叫做离散型随机变量.

实例1 观察掷一个骰子出现的点数.

随机变量 $X$  的可能值是: 1, 2, 3, 4, 5, 6.

(2) 连续型 随机变量所取的可能值可以连续地充满某个区间, 叫做连续型随机变量.

实例1 随机变量 $X$ 为“灯泡的寿命”.

则 $X$ 的取值范围为  $[0, +\infty)$ .

实例2 随机变量 $X$ 为“测量某零件尺寸时的测量误差”

则 $X$ 的取值范围为  $(a, b)$ .

# 离散型随机变量

**概率分布定义：** 设  $x_k$  ( $k=1,2, \dots$ ) 是离散型随机变量  $X$  所取的一切可能值，称

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

为离散型随机变量  $X$  的概率分布或分布律(分布列).

其中  $p_k$  ( $k=1,2, \dots$ ) 满足：

① 非负性：  $p_k \geq 0, \quad k = 1, 2, \dots$

② 规范性：  $\sum_k p_k = 1$

用这两条性质  
判断一个函数  
是否是分布律



# 离散型随机变量表示方法

## (1) 公式法

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

## (2) 列表法

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

或

$X$	$x_1$	$x_2$	$\cdots$	$x_n$	$\cdots$
$p_k$	$p_1$	$p_2$	$\cdots$	$p_n$	$\cdots$

**例1:** 设随机变量 $X$ 的分布律为:

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

试确定常数 $a$  .

**解:** 依据分布律的性质

$$\begin{cases} P(X = k) \geq 0, \\ \sum_k P(X = k) = 1 \end{cases}$$

$$e^\lambda = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$$

即  $a \geq 0$  ,  $\sum_{k=0}^{\infty} a \frac{\lambda^k}{k!} = a e^\lambda = 1$  从中解得  $a = e^{-\lambda}$

**例2：**某篮球运动员投篮的命中率是0.9，求他两次独立投篮，投中次数X的概率分布.

**解：** X可取值为0,1,2；

$$P(X=0)=(0.1)(0.1)=0.01$$

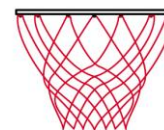
$$P(X=1)=2(0.9)(0.1)=0.18。$$

$$P(X=2)=(0.9)(0.9)=0.81$$

常表示为：

X	0	1	2
$p_k$	0.01	0.18	0.81

利用第一章  
的知识求解  
概率



**例3:** 若X的分布为:

$X$	0	1	2
$p_k$	0.01	0.18	0.81

试计算:  $P(X \geq 1.5)$ ,  $P(0 < X < 3)$ ,  $P(X \geq 1 | X < 2)$

**解:**  $P(X \geq 1.5) = P(X = 2) = 0.81$

$$P(0 < X < 3) = P(X = 1) + P(X = 2) = 0.18 + 0.81 = 0.99$$

$$P(X \geq 1 | X < 2) = \frac{P(X \geq 1, X < 2)}{P(X < 2)} = \frac{P(X = 1)}{P(X = 0) + P(X = 1)} = \frac{0.18}{0.19}$$

## 练习：

一个盒中装有**2**个白球，**3**个黑球，每次从中任取一个，直到取得白球为止。求取球次数的概率分布。

- (**1**) 每次取出的球不再放回；
- (**2**) 每次取出的球再放回。

# 几种常见的 离散型随机变量

➤ 0-1分布(两点分布)

➤ 二项分布

➤ 泊松分布

还有：几何分布，  
超几何分布，  
巴斯卡分布等。

# 0-1分布(两点分布)

**△定义：** 设随机变量  $X$  只可能取0与1两个值，它的分布律为

$X$	0	1
$p_k$	$1-p$	$p$

则称  $X$  服从 **(0-1) 分布或两点分布**.

**△背景：** 随机试验的结果只有两种可能，或试验的观察者只关心某个事件发生与否，都可以用两点分布来描述。



**说明：** 两点分布是最简单的一种分布,任何一个只有两种可能结果的随机现象,比如新生婴儿是男还是女、明天是否下雨、种籽是否发芽等,都属于两点分布.

**△注：** 0-1分布还可以用来计数。

# 伯努利试验和二项分布

**定义：** 设将试验独立重复进行 $n$ 次，每次试验中，事件A发生的概率均为 $p$ ，则称这 $n$ 次试验为 **$n$ 重伯努利试验**.

若以 $X$ 表示 $n$ 重伯努利试验中事件A发生的次数，则 $X$ 就是一个随机变量。

称 $X$ 服从参数为 $n$ ， $p$ 的**二项分布或伯努利分布**，记作 **$X \sim B(n, p)$** ，其分布律为：

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

**例1:** 一大批种子发芽率为90%，今从中任取10粒. 求播种后，（1）恰有8粒发芽的概率；（2）不小于8粒发芽的概率

**解:** 设随机变量X表示发芽的种子数，则有

$$\mathbf{X \sim B(10, 0.9)} \quad P(X = k) = C_{10}^k (0.9)^k (0.1)^{10-k}$$

$$(1) \quad P(X=8) = C_{10}^8 0.9^8 \times 0.1^2 \approx 0.1937$$

$$\begin{aligned} (2) \quad P(X \geq 8) &= P(X=8) + P(X=9) + P(X=10) \\ &= C_{10}^8 0.9^8 \times 0.1^2 + C_{10}^9 0.9^9 \times 0.1 + C_{10}^{10} 0.9^{10} \\ &\approx 0.9298 \end{aligned}$$

**例2：**已知100个产品中有5个次品，现从中有放回地取3次，每次任取1个，求在所取的3个中恰有2个次品的概率.

**解：**设随机变量 $X$ 表示所取的3个产品中的次品数，则

$$X \sim B(3, 0.05)$$

于是，所求概率为：

$$P(X=2)=C_3^2 (0.05)^2 (0.95) = 0.007125$$

**例2:** 已知100个产品中有5个次品，现从中有放回地取3次，每次任取1个，求在所取的3个中恰有2个次品的概率.

**请注意:** 若将本例中的“有放回”改为“无放回”，那么各次试验条件就不同了，此试验就不是伯努利试验. 此时, 只能用古典概型求解.

$$P(X=2)=\frac{C_{95}^1 C_5^2}{C_{100}^3} \approx 0.00618$$

例3：若某人做某事的成功率为1%，他重复努力400次，则至少成功一次的概率为多少？

解：设随机变量X表示成功次数，则  $X \sim B(400, 0.01)$

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) \\ &= 1 - 0.99^{400} \approx 0.9820 \end{aligned}$$

有百分之一的希望，就要做百分之百的努力

# 泊松分布 Poisson distribution

## ■ 定义

若随机变量  $X$  所有可能的取值为  $0, 1, 2, \dots$ , 而取各个值的概率为:

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

其中  $\lambda > 0$  为常数, 则称  $X$  服从参数为  $\lambda$  的泊松分布, 记作

$$X \sim P(\lambda) \text{ 或者 } X \sim \Pi(\lambda)$$

在生物学、医学、工业统计、保险科学及公用事业的排队等问题中，泊松分布是常见的.例如地震、火山爆发、特大洪水、交换台的电话呼唤次数等,都服从泊松分布.

地震



火山爆发



特大洪水



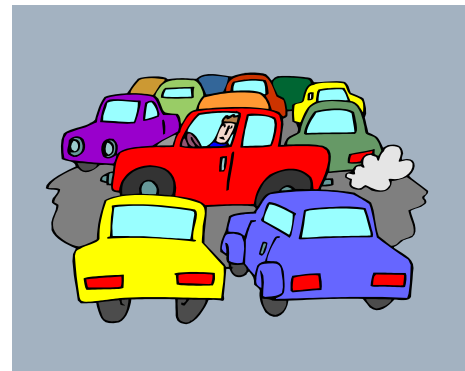
商场接待的顾客数



电话呼唤次数



交通事故次数





■ 在实际问题中，有很多r.v.都是服从或近似服从Poisson分布的.

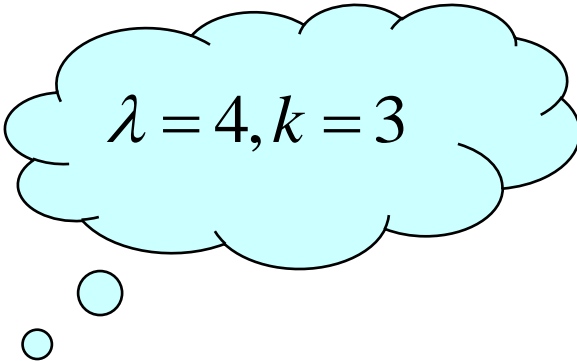
- 服务台在某时间段内接待的服务次数;
- 交换台在某时间段内接到呼叫的次数;
- 矿井在某段时间发生事故的次数;
- 显微镜下相同大小的方格内微生物的数目;
- 单位体积空气中含有某种微粒的数目;

体积相对小的物质在较大的空间内的稀疏分布，  
即稠密性问题，都可以看作泊松分布，其参数 $\lambda$ 可以由  
观测值的平均值求出。

**例4:** 已知某电话交换台每分钟接到的呼唤次数 $X$ 服从 $\lambda=4$ 的泊松分布, 分别求: (1) 每分钟内恰好接到3次呼唤的概率; (2) 每分钟不超过4次的概率

解

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


$$\lambda = 4, k = 3$$

$$P(X = 3) = \frac{4^3}{3!} e^{-4} = 0.19563$$

$$P(X \leq 4) = P(X = 0) + P(X = 1) + P(X = 2) \\ + P(X = 3) + P(X = 4)$$

注: 计算 $P(X \geq x)$ 型的概率, 也可以查表计算:  $\lambda=4, x=5$

$$P(X \leq 4) = 1 - P(X \geq 5) = 1 - 0.371163 = 0.628837$$

**例5:** 一家商店采用科学管理, 由该商店过去的销售记录知道, 某种商品每月的销售数可以用参数 $\lambda=5$ 的泊松分布来描述, 为了以95%以上的把握保证不脱销, 问商店在月底至少应进某种商品多少件?

**解:** 设商店在月底应进某种商品 $m$ 件,

设该商品每月的销售数为 $X$ , 则 $X \sim \Pi(\lambda)$ .

求满足  $P(X \leq m) > 0.95$  的最小的 $m$ .

销售数

进货数

求满足  $P(X \leq m) > 0.95$  的最小的  $m$ .

也即  $P(X \geq m+1) \leq 0.05$

或 
$$\sum_{k=m+1}^{\infty} \frac{e^{-5} 5^k}{k!} \leq 0.05$$

查泊松分布表得

$$\sum_{k=10}^{\infty} \frac{e^{-5} 5^k}{k!} \approx 0.032, \quad \sum_{k=9}^{\infty} \frac{e^{-5} 5^k}{k!} \approx 0.068$$

于是得  $m+1 \geq 10$ ,

$m \geq 9$  件

## 二项分布的泊松逼近

历史上，泊松分布是作为二项分布的近似，于1837年由法国数学家泊松引入的。

### 泊松定理

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

$$\lambda = np$$

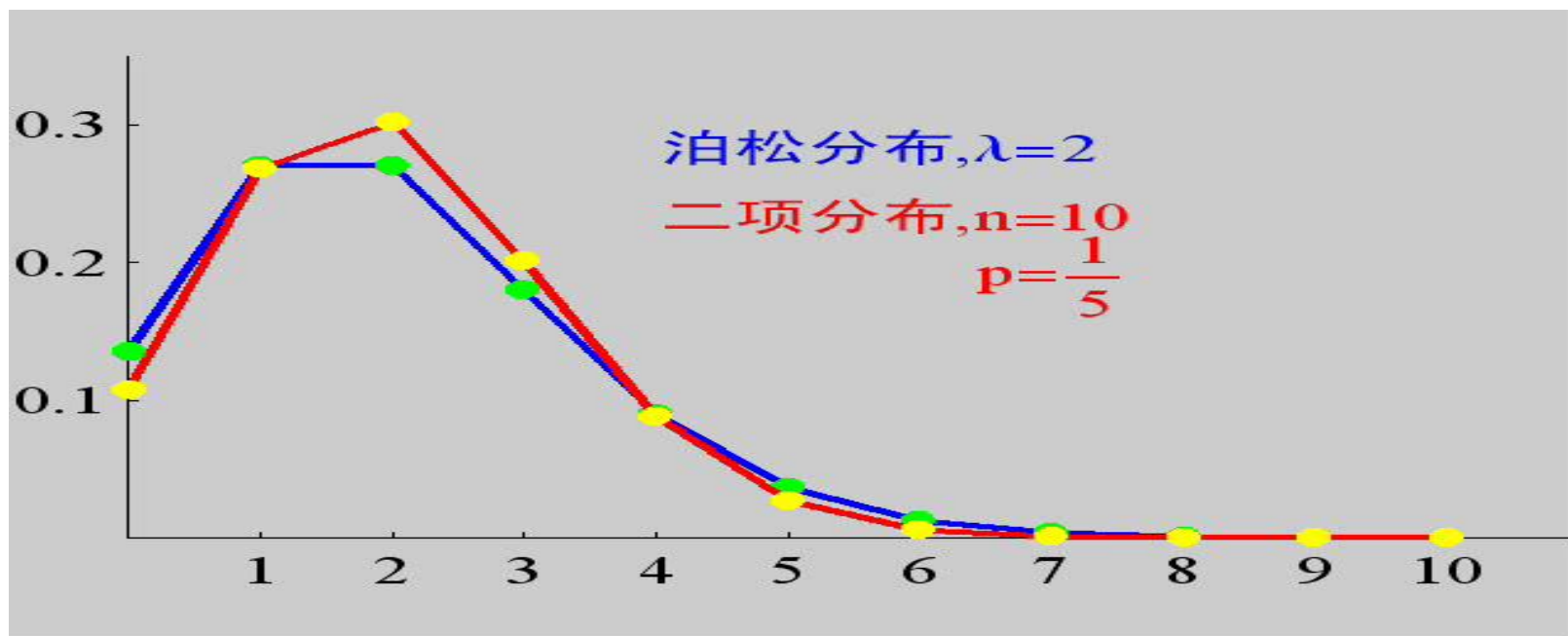
实际应用中：当 $n$ 较大， $p$ 较小， $np$ 适中时，即可用泊松公式近似替换二项概率公式。



泊松，S. -D.

# 二项分布的泊松逼近

二项分布  $np \rightarrow \lambda (n \rightarrow +\infty)$  泊松分布



**例6:** 独立射击5000次, 命中率为0.001, 求

(1) 最可能命中次数及相应的概率;

(2) 命中次数不少于1 次的概率.

**解:** 设随机变量 $X$  表示命中次数, 则

$$X \sim B(5000, 0.001)$$

(1) 最可能命中次数 $k = [(n + 1)p] = [(5000 + 1)0.001] = 5$

$$P(X = 5) = C_{5000}^5 (0.001)^5 (0.999)^{4995} \approx 0.1756$$

$$\approx \frac{5^5}{5!} e^{-5} = 0.1755$$

利用Poisson定理试试

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

**例：**独立射击**5000**次，命中率为**0.001**，求  
(2) 命中次数不少于**1** 次的概率。

**解：**由命中次数  $X \sim B(5000, 0.001)$ ，可得

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) = 1 - C_{5000}^0 (0.001)^0 (0.999)^{5000} \\ &= 0.9934. \end{aligned}$$

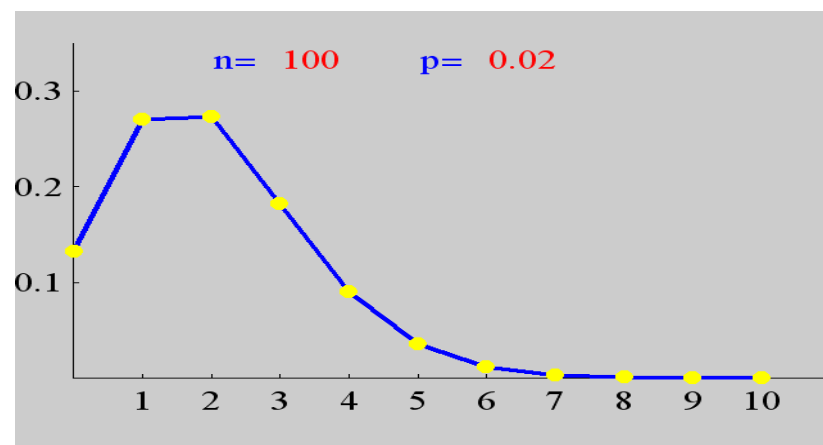
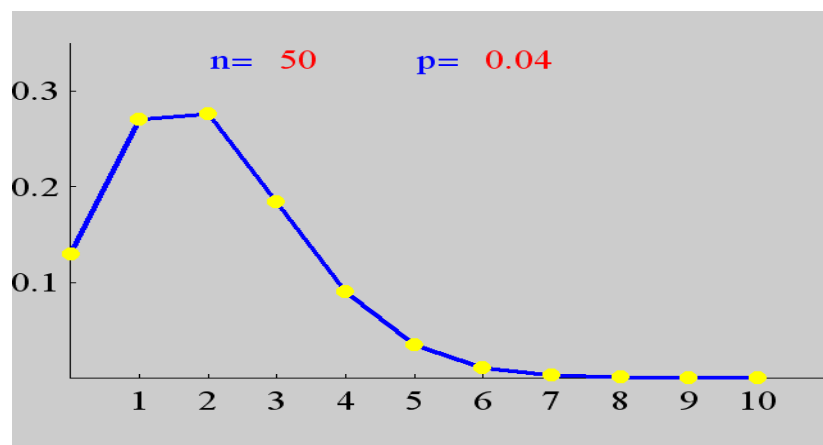
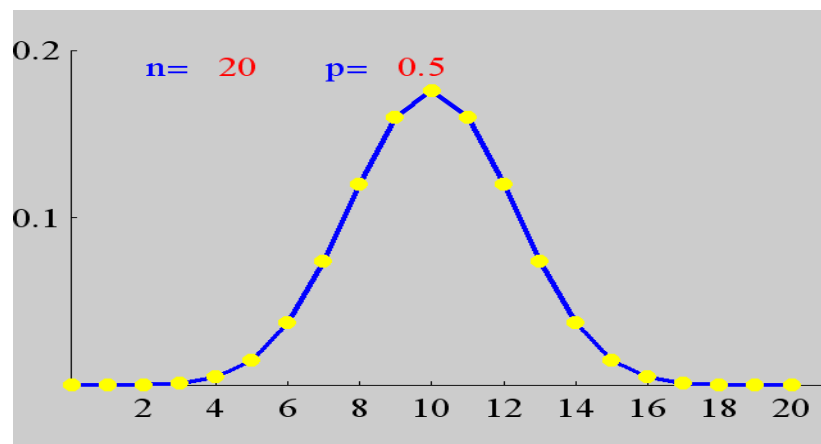
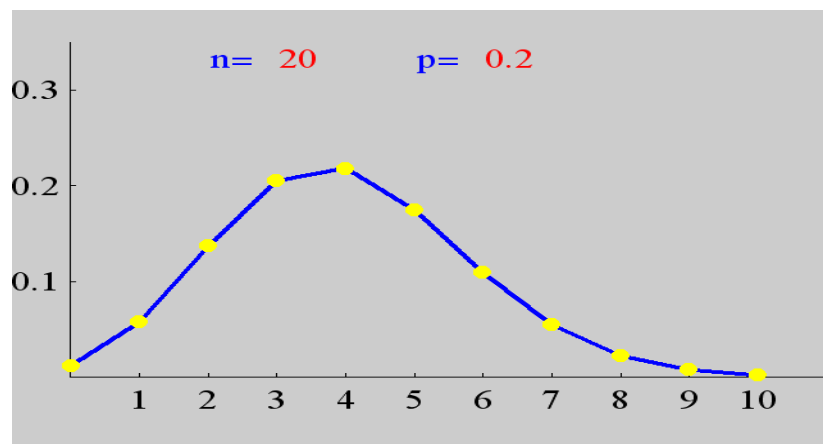
利用**Poisson**定理试试

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

$$\begin{aligned} P(X \geq 1) &= 1 - C_{5000}^0 (0.001)^0 (0.999)^{5000} \\ &\approx 1 - \frac{5^0}{0!} e^{-5} = 0.9933 \end{aligned}$$



# 二项分布的图形



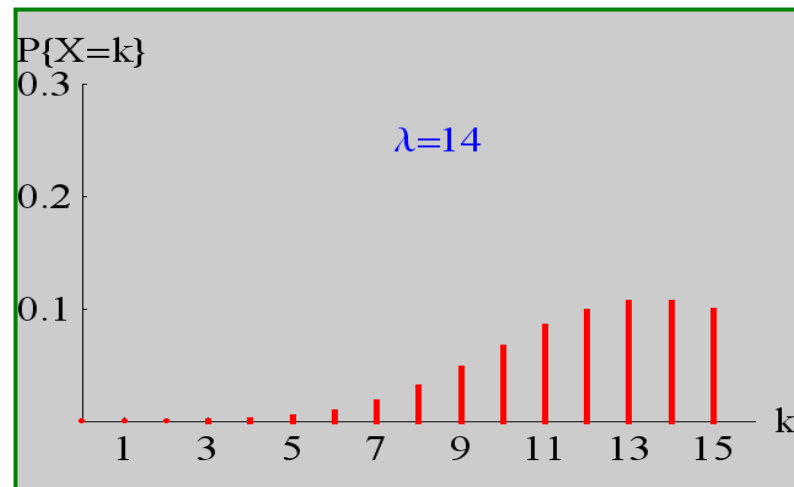
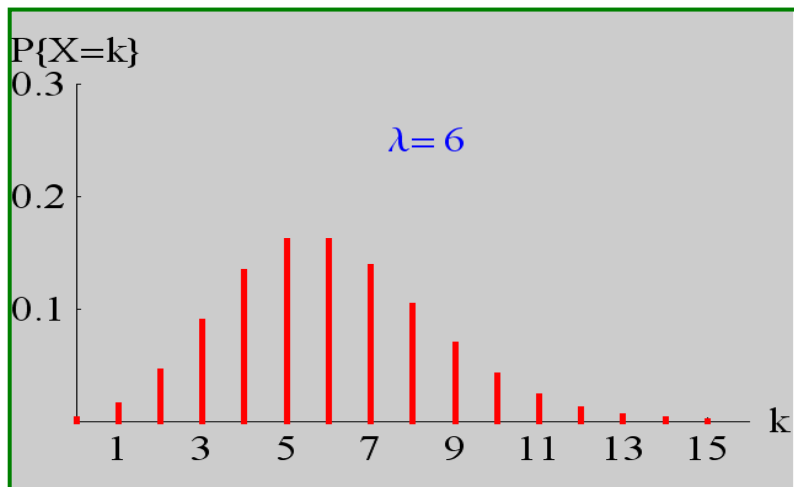
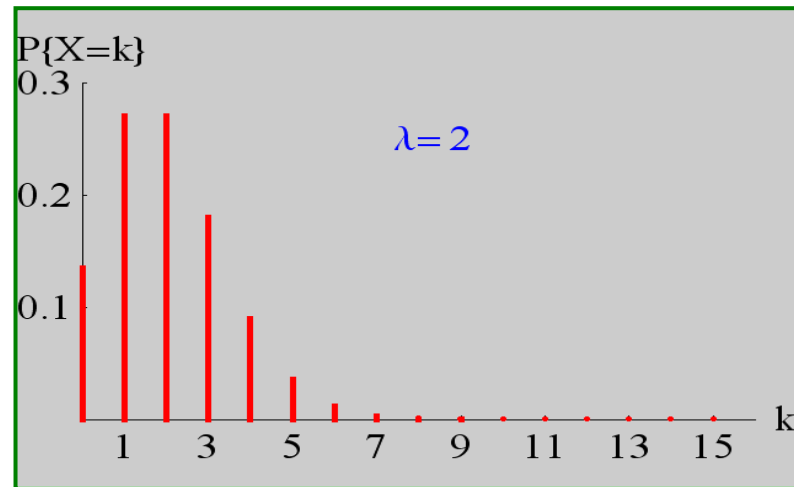
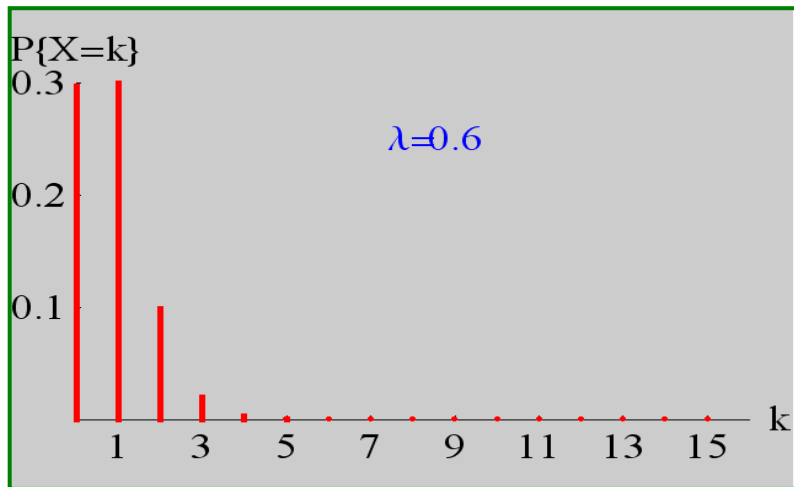
二项分布的最大可能值：一般地，若在 $k_0$ 处，概率 $P\{X=k\}$ 达到最大（称 $k_0$ 为随机变量 $X$ 的最可能次数）。则 $k_0$ 应满足

$$\begin{cases} \frac{P(X = k_0)}{P(X = k_0 + 1)} \geq 1 \\ \frac{P(X = k_0)}{P(X = k_0 - 1)} \geq 1 \end{cases}$$

解上述不等式得 $(n+1)p-1 \leq k_0 \leq (n+1)p$ ，因为 $k_0$ 必须为整数，所以

$$k_0 = \begin{cases} (n+1)p \text{ 和 } (n+1)p - 1, & \text{当 } (n+1)p \text{ 为整数时} \\ [(n+1)p], & \text{其他} \end{cases}$$

# 泊松分布的图形



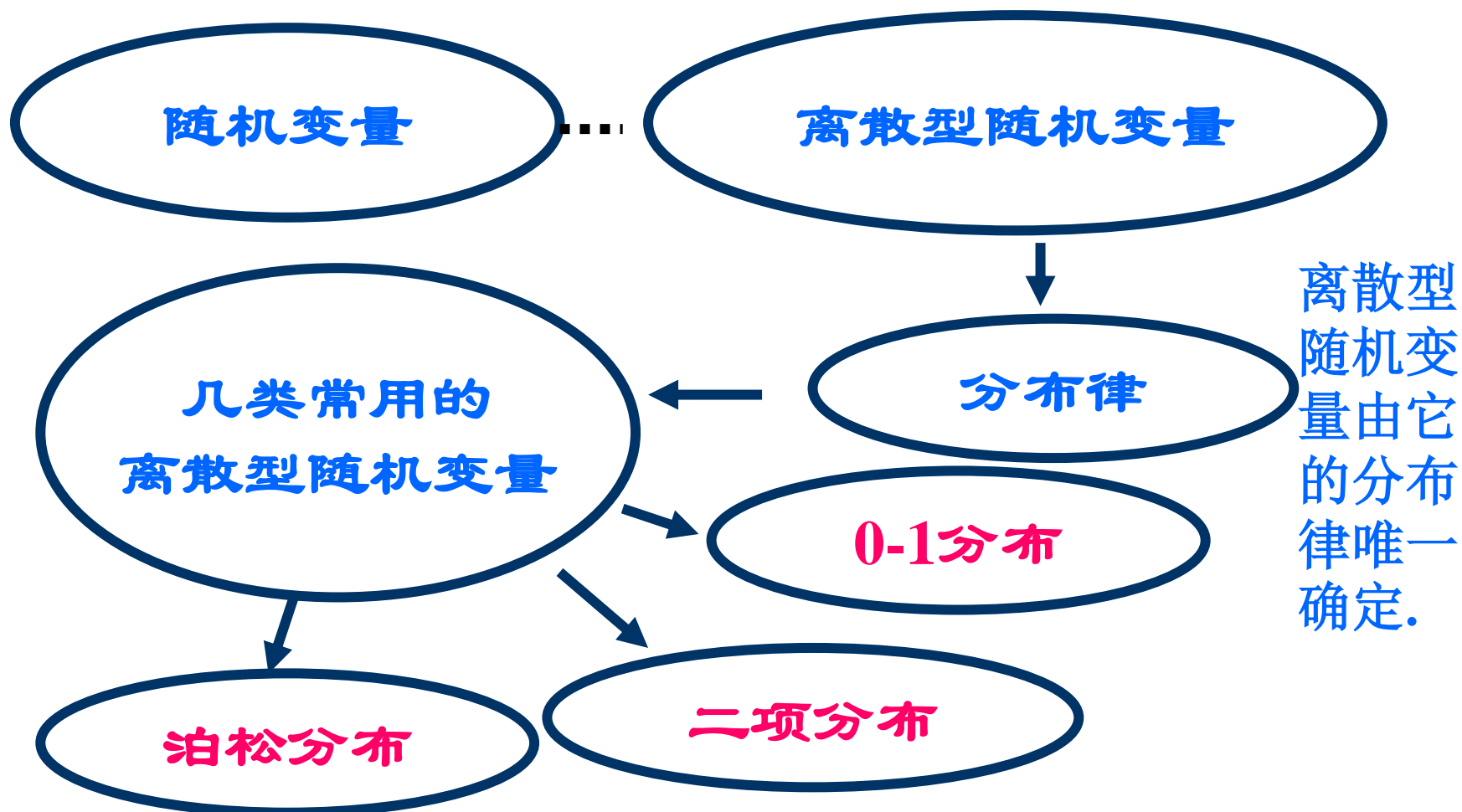
同样地，解如下不等式

$$\begin{cases} \frac{P(X = k_0)}{P(X = k_0 + 1)} \geq 1 \\ \frac{P(X = k_0)}{P(X = k_0 - 1)} \geq 1 \end{cases}$$

得  $\lambda - 1 \leq k_0 \leq \lambda$ 。因为  $k_0$  必须为整数，所以泊松分布的最可能次数为

$$k_0 = \begin{cases} \lambda \text{ 和 } \lambda - 1, & \text{当 } \lambda \text{ 为整数,} \\ [\lambda], & \text{其它,} \end{cases}$$

# 小 结



**例1：** 有**90**台同类型设备，各台设备的工作相互独立，发生故障的概率均为**0.01**，且一台设备的故障由一名设备维修人员能够处理。配备维修人员的方法有两种：  
一种是**3**名维修人员单独工作，每人负责**30**台设备的维修；另一种是**3**名维修人员联合工作，共同负责**90**台设备的维修。  
试比较两种配备维修人员方法的优劣。

例2：某商场每天的客流量服从参数为  $\lambda$  ( $\lambda > 0$ ) 的泊松分布。假定每位顾客在该商场消费的概率为  $p$  ( $0 < p < 1$ )，且他们在该商场消费是相互独立的。求：某天在该商场恰有  $k$  人消费的概率？

# 第四节

## 分布函数



# 随机变量的分布函数

## ■ 分布函数的定义

设 $X$ 为一随机变量,则对任意实数 $x$ ,  $(X \leq x)$  是一个随机事件, 称  $F(x) = P(X \leq x)$

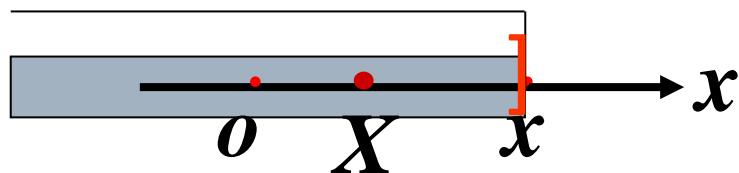
为随机变量 $X$ 的分布函数

定义域为  $(-\infty, +\infty)$ ; 值域为  $[0, 1]$



$F(x)$ 是一个普通的函数!

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



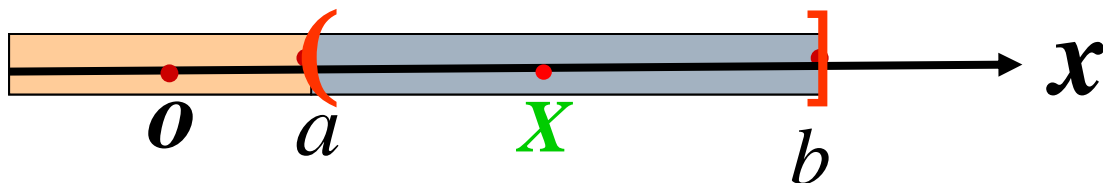
分布函数是一个普通的函数，  
正是通过它，我们可以用高等数  
学的工具来研究随机变量.

**请注意：**

(1) 在分布函数的定义中， $X$ 是随机变量， $x$ 是参变量.

(2) 对任意实数 $a < b$ ，随机点落在区间 $(a, b]$ 内的概率为：

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$$



因此，只要知道了随机变量 $X$ 的分布函数，  
它的统计特性就可以得到全面的描述。

# 分布函数的性质

1. 单调性:  $F(x)$  是单调不减函数

若  $x_1 < x_2$ , 则  $F(x_1) \leq F(x_2)$

2. 规范性:  $0 \leq F(x) \leq 1$

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

3. 连续性:  $F(x)$  处处右连续, 即不可能事件

$$F(x_0 + 0) = \lim_{x \rightarrow x_0 + 0} F(x) = F(x_0) \quad \text{必然事件}$$

**注:** 如果一个函数具有上述性质, 则一定是某个  $r.v.$   $X$  的分布函数。也就是说, 性质 1—3 是鉴别一个函数是否是某  $r.v.$  的分布函数的充分必要条件。

## 问一问

$F(x) = \frac{1}{1+x^2}$  是不是某一随机变量的分布函数?

不是 因为  $\lim_{x \rightarrow +\infty} F(x) = 0$

函数  $G(x) = \begin{cases} \frac{1}{1+x^2} & (x \leq 0) \\ 1 & (x > 0) \end{cases}$  可作为分布函数.

例：设随机变量  $X$  的分布律为

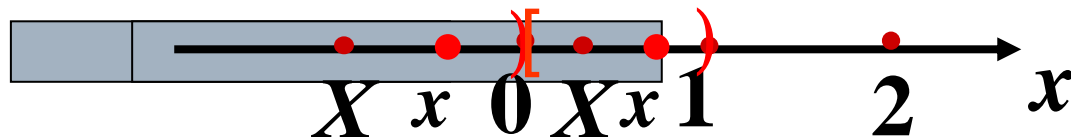
$X$	0	1	2
$p_k$	1/3	1/6	1/2

求  $X$  的分布函数  $F(x)$ .

解：  $F(x) = P(X \leq x)$

当  $x < 0$  时， $(X \leq x) = \Phi$ ，故  $F(x) = 0$

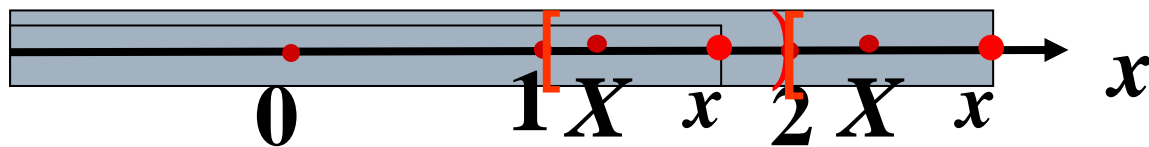
当  $0 \leq x < 1$  时， $F(x) = P(X \leq x) = P(X=0) = 1/3$



<b>X</b>	0	1	2
<b><math>P_k</math></b>	1/3	1/6	1/2

当  $1 \leq x < 2$  时,  $F(x) = P(X=0) + P(X=1) = 1/3 + 1/6 = 1/2$

当  $x \geq 2$  时,  $F(x) = P(X=0) + P(X=1) + P(X=2)$   
 $= 1/3 + 1/6 + 1/2 = 1$



$$\text{故 } F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{3}, & 0 \leq x < 1 \\ \frac{1}{2}, & 1 \leq x < 2 \\ 1, & x \geq 2 \end{cases}$$

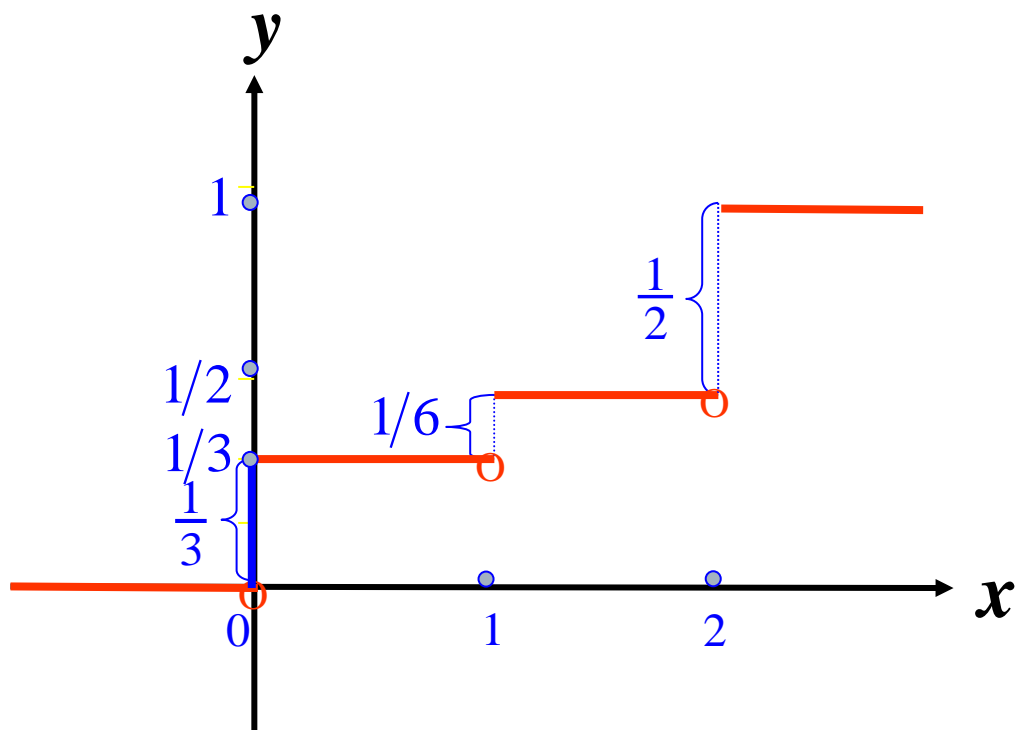
注意右连续

我们从图形上来看一下.

## $F(x)$ 的分布函数图

$X$	0	1	2
$p_k$	$1/3$	$1/6$	$1/2$

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{3}, & 0 \leq x < 1 \\ \frac{1}{2}, & 1 \leq x < 2 \\ 1, & x \geq 2 \end{cases}$$



离散型随机变量的分布函数是**阶梯函数**, **跳跃点**对应离散型随机变量的可能取值点, **跳跃高度**对应随机变量取对应值的概率;反之,也成立。



# 一般地

设离散型 r.v  $X$  的分布律是

$$P(X = x_k) = p_k, k = 1, 2, 3, \dots$$

则其分布函数

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

即  $F(x)$  是  $X$  取  $\leq x$  的诸值  $x_k$  的概率之和.



---

一个盒中装有**2**个白球，**3**个黑球，每次从中任取一个，直到取得白球为止。求取球次数的分布函数。

- (**1**) 每次取出的球不再放回；
  - (**2**) 每次取出的球再放回。
-



已知随机变量 $X$   
的分布函数为

$$F(x) = \begin{cases} a & x < -1 \\ 1/3 & -1 \leq x < 1 \\ 1/2 & 1 \leq x < 2 \\ b & x \geq 2 \end{cases}$$

求 $a, b$ , 并求 $X$ 的分布列

**解：**由分布函数的性质可得  $a=0, b=1$ ，且 $X$ 为一离散型随机变量，其分布列为

$X$	-1	1	2
$p_k$	1/3	1/6	1/2

# 第五节

第五节

连续型随机变量

◆离散型随机变量：取值有限或无限可列  
刻画方法：分布律

◆连续型随机变量：取值可充满一个区间  
刻画方法：概率密度函数

# 概率密度函数

■ 定义：设 $X$ 为一随机变量，若存在非负可积函数  $f(x)$ ，使对任意实数  $-\infty < x < +\infty$ ，有

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

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

连续型随机变量的分布  
函数在 $R$ 上连续

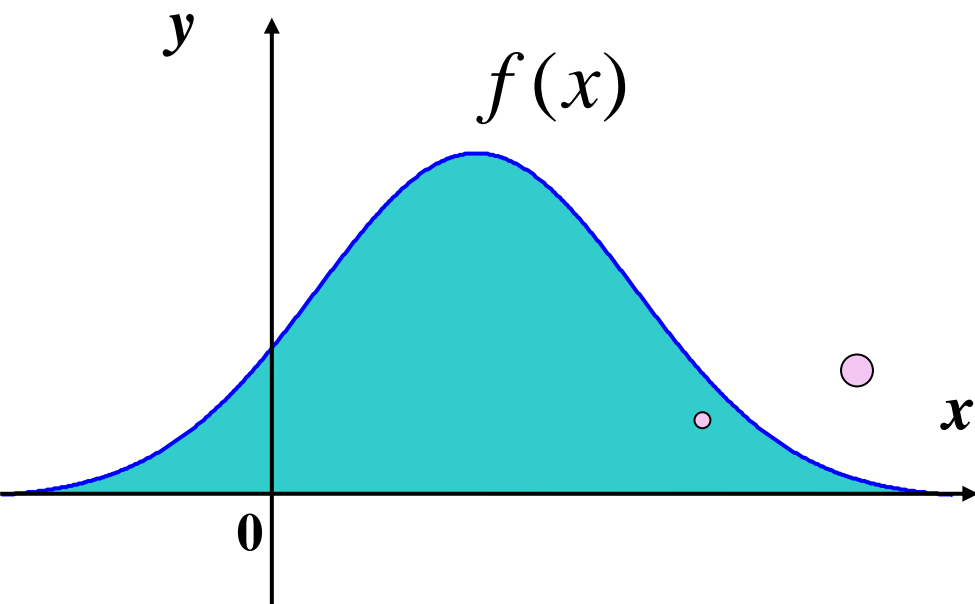
# 概率密度函数的性质

■ 非负性:

$$f(x) \geq 0$$

■ 规范性(归一性):

$$\int_{-\infty}^{+\infty} f(x) dx = 1$$



【注】这两条性质是判定一个函数  $f(x)$  是否为某  $r.v$   $X$  的概率密度的充要条件

阴影部分  
面积为1

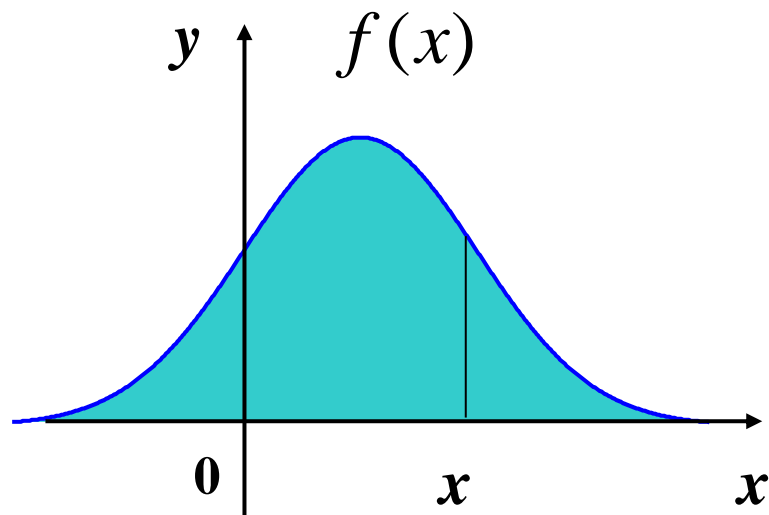
$$P(-\infty < X < +\infty) = 1$$

# 密度函数和分布函数的关系

## ■ 积分关系

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

**注：**分布函数 $F(x)$ 的非负性，单调性和规范性得到了很明显的体现。



## ■ 导数关系

若 $f(x)$ 在 $x$ 处连续，则  $F'(x) = f(x)$



# 密度在求概率上的两个结论

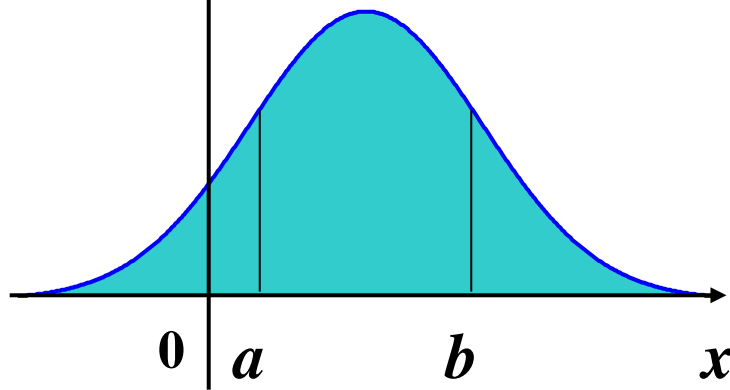
◆ 连续型 r.v.  $X$  取任意单点值  $a$  的概率均为 0, 即

$$P(X = a) = 0$$

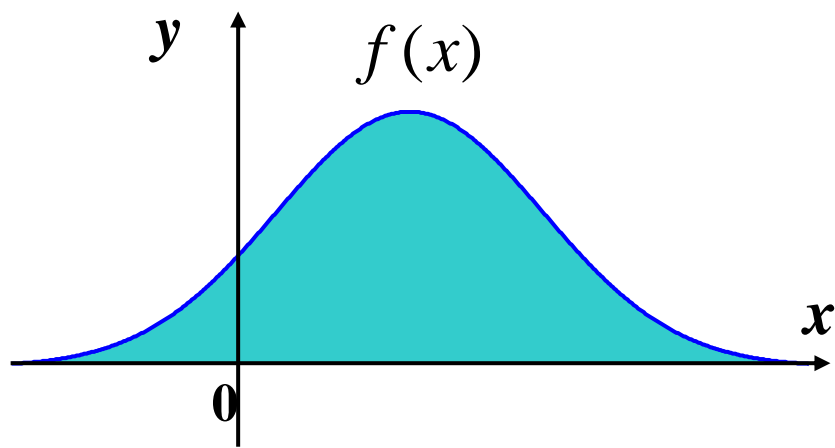
◆ 计算概率的公式: 设  $X$  是连续型随机变量,  $a < b$

$$P(a < X < b) = F(b) - F(a) = \int_a^b f(x) dx$$

$$\begin{aligned} & P(a < X < b) = \int_a^b f(x) dx \\ & = P(a < X \leq b) \\ & = P(a \leq X < b) \\ & = P(a < X < b) \end{aligned}$$



注意:



- 如果把概率理解为质量,  $f(x)$  相当于物理学中的线密度。
- 概率密度函数  $f(x)$  在点  $a$  处取值, 不是事件  $\{X=a\}$  的概率。但是, 该值越大,  $X$  在  $a$  点附近取值的概率越大。

# 连续型 r.v. 求概率的公式

◆ 设  $X$  是连续型随机变量,  $a < b$

$$P(a < X \leq b) = F(b) - F(a) = \int_a^b f(x) dx$$

$$P(X > a) = 1 - F(a) = \int_a^{+\infty} f(x) dx$$

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

无论等号  
在哪里

## 例1: 已知密度函数, 求其他

设随机变量 $X$ 具有概率密度

$$f(x) = \begin{cases} kx, & 0 \leq x < 3 \\ 2 - \frac{x}{2}, & 3 \leq x \leq 4 \\ 0, & \text{其它} \end{cases}$$

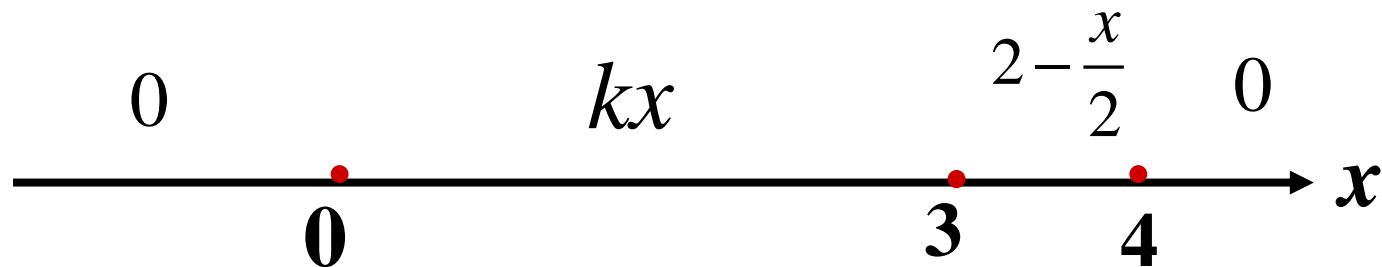
(1) 确定常数 $k$ ; (2) 求 $X$ 的分布函数 $F(x)$ ;

(3) 求 $P\left(1 < X \leq \frac{7}{2}\right)$

解: 
$$f(x) = \begin{cases} kx, & 0 \leq x < 3 \\ 2 - \frac{x}{2}, & 3 \leq x \leq 4 \\ 0, & \text{其它} \end{cases}$$

利用密度函数的性质求出  $k$   $\int_{-\infty}^{+\infty} f(x)dx = 1$

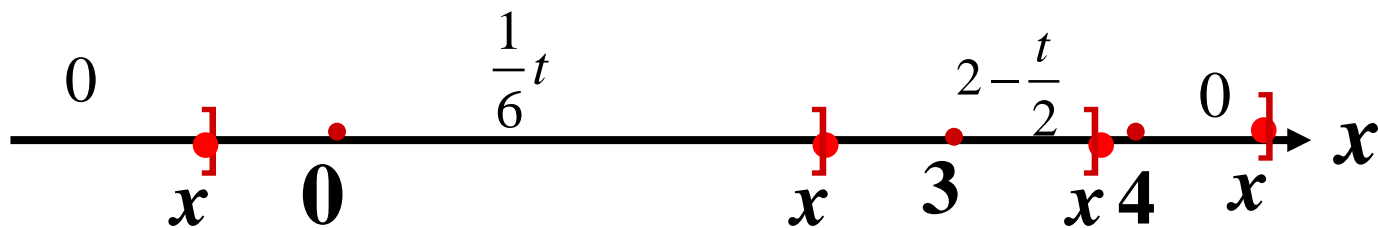
$$\int_{-\infty}^{+\infty} f(x)dx = \int_0^3 kx dx + \int_3^4 \left(2 - \frac{x}{2}\right) dx = 1 \Rightarrow k = \frac{1}{6}$$



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

(2) 分布函数为

$$F(x) = \begin{cases} 0, & x < 0 \\ \int_0^x \frac{t}{6} dt, & 0 \leq x < 3 \\ \int_0^3 \frac{t}{6} dt + \int_3^x \left(2 - \frac{t}{2}\right) dt, & 3 \leq x < 4 \\ \int_0^3 \frac{t}{6} dt + \int_3^4 \left(2 - \frac{t}{2}\right) dt, & x \geq 4 \end{cases}$$



即分布函数

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x^2}{12}, & 0 \leq x < 3 \\ -3 + 2x - \frac{x^2}{4}, & 3 \leq x < 4 \\ 1, & x \geq 4 \end{cases}$$

求概率  $P\left(1 < X \leq \frac{7}{2}\right)$ , 有两种方法

1. 已知密度

$$P(a < X \leq b) = \int_a^b f(x) dx$$

$$P\left(1 < X \leq \frac{7}{2}\right) = \int_1^{\frac{7}{2}} f(x) dx = \int_1^3 \frac{1}{6} x dx + \int_3^{\frac{7}{2}} \left(2 - \frac{x}{2}\right) dx = \frac{41}{48}$$

2. 已知分布

$$P(a < X \leq b) = F(b) - F(a)$$

$$P\left(1 < X \leq \frac{7}{2}\right) = F\left(\frac{7}{2}\right) - F(1) = \frac{41}{48}$$



## 例2：已知分布函数求其他

设连续型随机变量  $X$  的分布函数为

$$F(x) = \begin{cases} 0, & x \leq -a, \\ A + B \arcsin \frac{x}{a}, & -a < x \leq a, \\ 1, & x > a. \end{cases}$$

求：(1) 系数  $A, B$  的值；

(2)  $P\left(-a < X < \frac{a}{2}\right)$ ;

(3) 随机变量  $X$  的概率密度.

$$F(x) = \begin{cases} 0, & x \leq -a, \\ A + B \arcsin \frac{x}{a}, & -a < x \leq a, \\ 1, & x > a. \end{cases}$$

**解：**(1) 因为  $X$  是连续型随机变量，所以分布函数  $F(x)$  也连续，故有  $F(-a) = \lim_{x \rightarrow -a} F(x)$ ,

即

$$\begin{aligned} & F(a) = \lim_{x \rightarrow a} F(x), \\ & A + B \arcsin\left(\frac{-a}{a}\right) = A - \frac{\pi}{2}B = 0, \\ & A + B \arcsin\left(\frac{a}{a}\right) = A + \frac{\pi}{2}B = 1, \end{aligned} \Rightarrow \begin{cases} A = \frac{1}{2} \\ B = \frac{1}{\pi} \end{cases}$$

$$F(x) = \begin{cases} 0, & x \leq -a, \\ \frac{1}{2} + \frac{1}{\pi} \arcsin \frac{x}{a}, & -a < x \leq a, \\ 1, & x > a. \end{cases}$$

$$\begin{aligned} (2) \quad P\left(-a < X < \frac{a}{2}\right) &= F\left(\frac{a}{2}\right) - F(-a) \\ &= \frac{1}{2} + \frac{1}{\pi} \arcsin\left(\frac{a}{2a}\right) - 0 \\ &= \frac{1}{2} + \frac{1}{\pi} \times \frac{\pi}{6} = \frac{2}{3} \end{aligned}$$

$$F(x) = \begin{cases} 0, & x \leq -a, \\ \frac{1}{2} + \frac{1}{\pi} \arcsin \frac{x}{a}, & -a < x \leq a, \\ 1, & x > a. \end{cases}$$

(3) 随机变量  $X$  的概率密度为

$$f(x) = F'(x) = \begin{cases} 1/\pi\sqrt{a^2 - x^2}, & -a < x < a, \\ 0, & \text{其它.} \end{cases}$$

几种常见的

连续型随机变量

➤ 均匀分布

➤ 指数分布

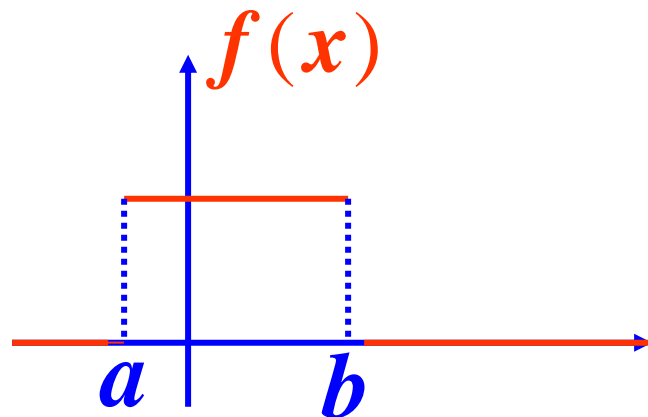
➤ 正态分布

# 均匀分布 Uniform Distribution

## ■ 定义

若连续型随机变量 $X$ 的概率密度为

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



则称 $X$ 在区间 $[a, b]$ 上服从均匀分布记为 $X \sim U[a, b]$ .

**注：**有时也记作  $X \sim U(a, b)$ .

# 均匀分布的适用范围

均匀分布常见于下列情形：

如在数值计算中，由于四舍五入，小数点后某一位小数引入的误差；

公交线路两辆公共汽车前后通过某汽车停车站的时间，即乘客的候车时间等。

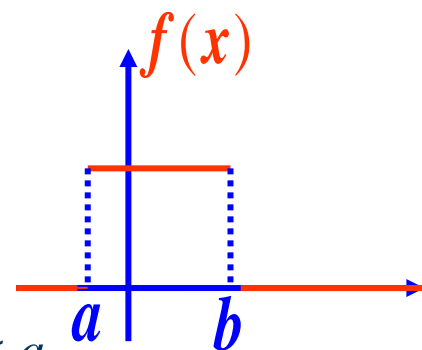




# 均匀分布的特点

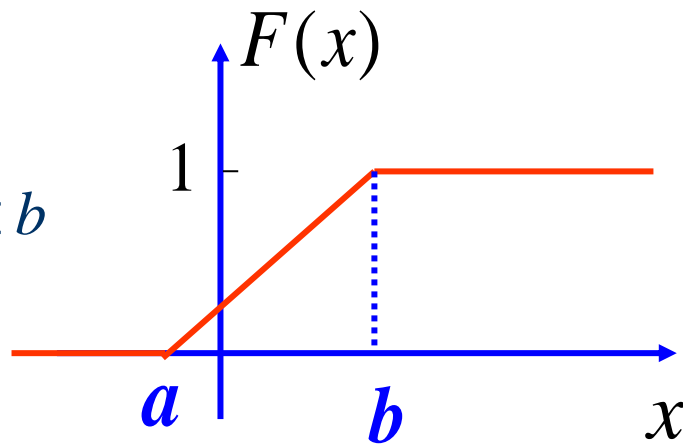
◆ 若  $X \sim U[a, b]$ , 则  $X$  的分布函数为

$$F(x) = \int_{-\infty}^x f(t)dt = \begin{cases} \int_{-\infty}^x 0dt, & -\infty < x < a \\ \int_{-\infty}^a 0dt + \int_a^x \frac{1}{b-a} dt, & a \leq x \leq b \\ \int_{-\infty}^a 0dt + \int_a^b \frac{1}{b-a} dt + \int_b^x 0dt, & x > b \end{cases}$$



分布函数的图为

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



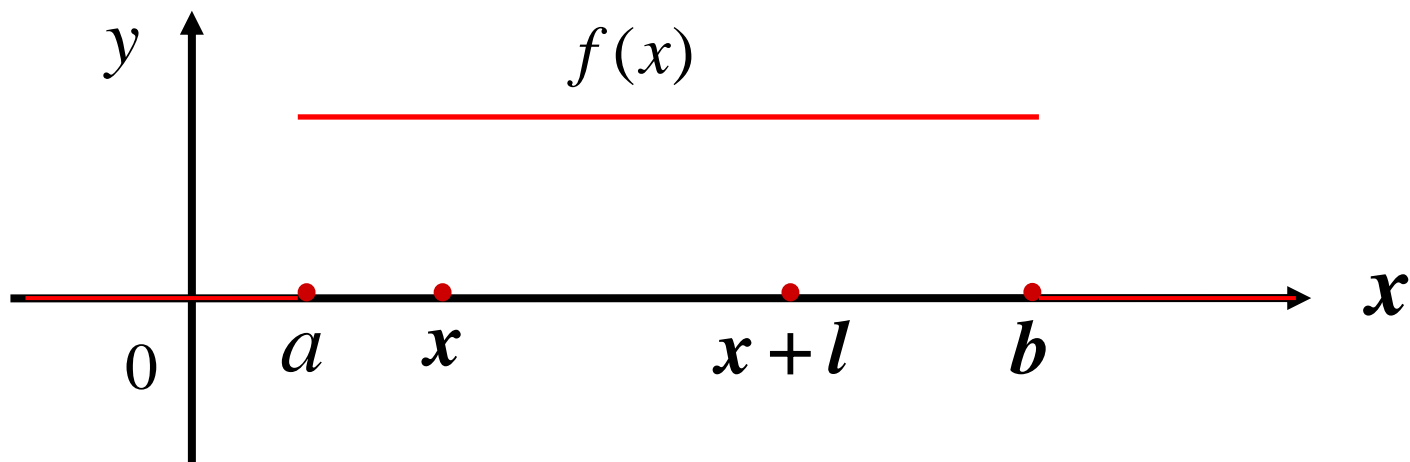
## 均匀分布的特点

若 $X \sim U[a, b]$ ，则 $X$ 只在 $[a, b]$ 内取值，且

- ◆  $X$ 落入 $[a, b]$ 任意子区间的概率仅与该子区间的长度成正比，而与子区间的位置无关.

当 $a \leq x < x + l \leq b$ ，有

$$P(x \leq X \leq x + l) = \int_x^{x+l} \frac{1}{b-a} dx = \frac{l}{b-a}$$



**例1:** 设  $r.v.\xi$  服从  $[-1, 5]$  上的均匀分布, 求方程

$$x^2 + 2\xi x + 1 = 0$$

有实根的概率。

**解:** 方程有实数根  $\iff 4\xi^2 - 4 \geq 0$  即  $|\xi| \geq 1$

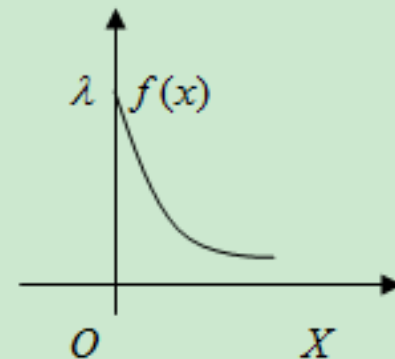
而  $\xi$  的密度函数为  $f(x) = \begin{cases} \frac{1}{6} & (-1 \leq x \leq 5) \\ 0 & \text{其它} \end{cases}$

$$\text{所求概率为 } P(|\xi| \geq 1) = \int_{-\infty}^{-1} f(x) dx + \int_1^{+\infty} f(x) dx = \frac{2}{3}$$

# 指数分布 Exponential Distribution

■ 定义: 若连续型随机变量 $X$ 的概率密度为

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



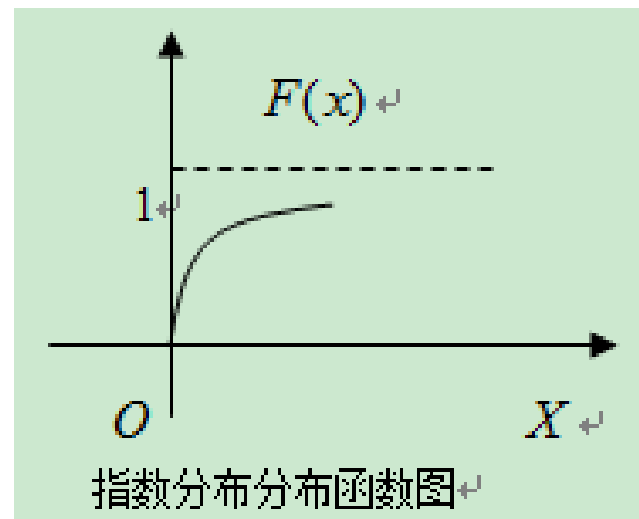
指数分布密度函数图

则称 $X$ 服从参数为 $\lambda$ 的指数分布, 记作:

$$X \sim E(\lambda)$$

■ 分布函数

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



指数分布分布函数图

# 指数分布的适用范围

指数分布常用于可靠性统计研究中。

如元件的寿命，随机服务系统的服务时间等。

**例2:** 设某类日光灯管的使用寿命 $X$ 服从参数为 $\lambda=1/2000$ 的指数分布(单位:小时),

(1)任取一只这种灯管, 求能正常使用1000小时以上的概率.

(2) 有一只这种灯管已经正常使用了2000 小时, 求还能使用1000小时以上的概率.

**解:** 由题意可知

$$f(x) = \begin{cases} \frac{1}{2000} e^{-\frac{1}{2000}x} & x > 0 \\ 0 & x \leq 0 \end{cases}, \quad F(x) = \begin{cases} 0 & x \leq 0 \\ 1 - e^{-\frac{1}{2000}x} & x > 0 \end{cases}$$

$$P(X > 1000) = \begin{cases} 1 - F(1000) \\ \int_{1000}^{+\infty} f(x) dx \end{cases} = e^{-\frac{1}{2}} \approx 0.607$$

(2) 有一只这种灯管已经正常使用了2000 小时，求还能使用1000小时以上的概率.

$$f(x) = \begin{cases} \frac{1}{2000} e^{-\frac{1}{2000}x} & x > 0 \\ 0 & x \leq 0 \end{cases}, \quad F(x) = \begin{cases} 0 & x \leq 0 \\ 1 - e^{-\frac{1}{2000}x} & x > 0 \end{cases}$$

$$\begin{aligned} P(X > 3000 | X > 2000) &= \frac{P(X > 3000, X > 2000)}{P(X > 2000)} \\ &= \frac{P(X > 3000)}{P(X > 2000)} = \frac{1 - F(3000)}{1 - F(2000)} = \frac{1 - (1 - e^{-3000/2000})}{1 - (1 - e^{-2000/2000})} \\ &= e^{-\frac{1}{2}} \approx 0.607 \end{aligned}$$

指数分布的重要性质：“无记忆性”。

## 指数分布的“无记忆性”

$X \sim E(\lambda), s > 0, t > 0$ , 有

$$P(X \geq s + t | X > s) = P(X \geq t)$$



# 第七节

第七节

正态分布

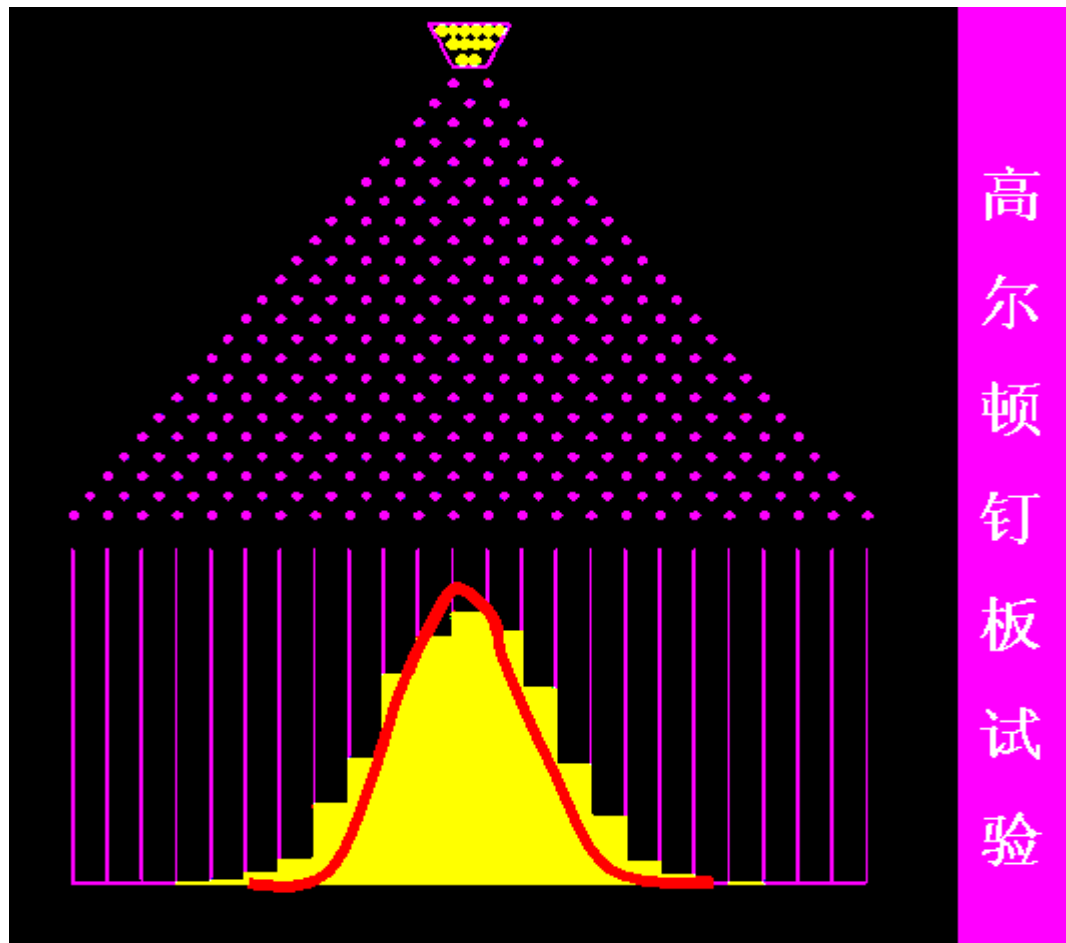
# 正态分布

## Normal Distribution

正态分布是实践中应用最为广泛，在理论上研究最多的分布之一，很多随机变量都服从或者近似服从正态分布，故它在概率统计中占有特别重要的地位。



正态分布是十九世纪初，由**高斯(Gauss)**给出并推广的一种分布，故也称高斯分布。



这条红色曲线近似我们将要介绍的正态分布的概率密度曲线。

# 正态分布

■ 若连续型随机变量X的概率密度为

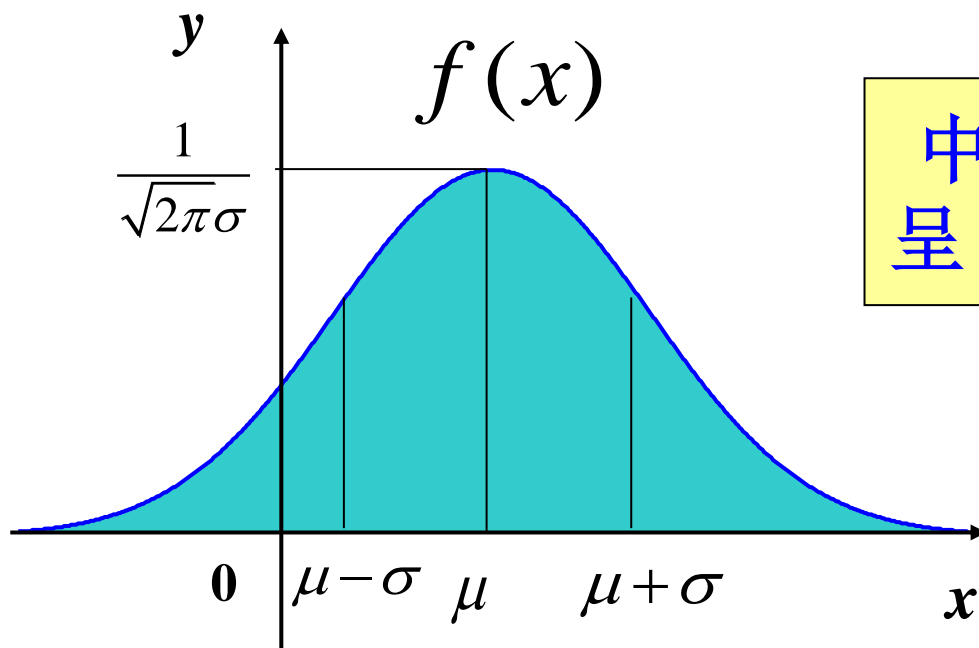
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty$$

其中 $\mu, \sigma (\sigma > 0)$ 是常数，则称X服从参数为 $\mu, \sigma$ 的正态分布(或高斯分布)，记作

$$X \sim N(\mu, \sigma^2)$$

注： $f(x)$ 所确定的曲线叫作正态曲线。

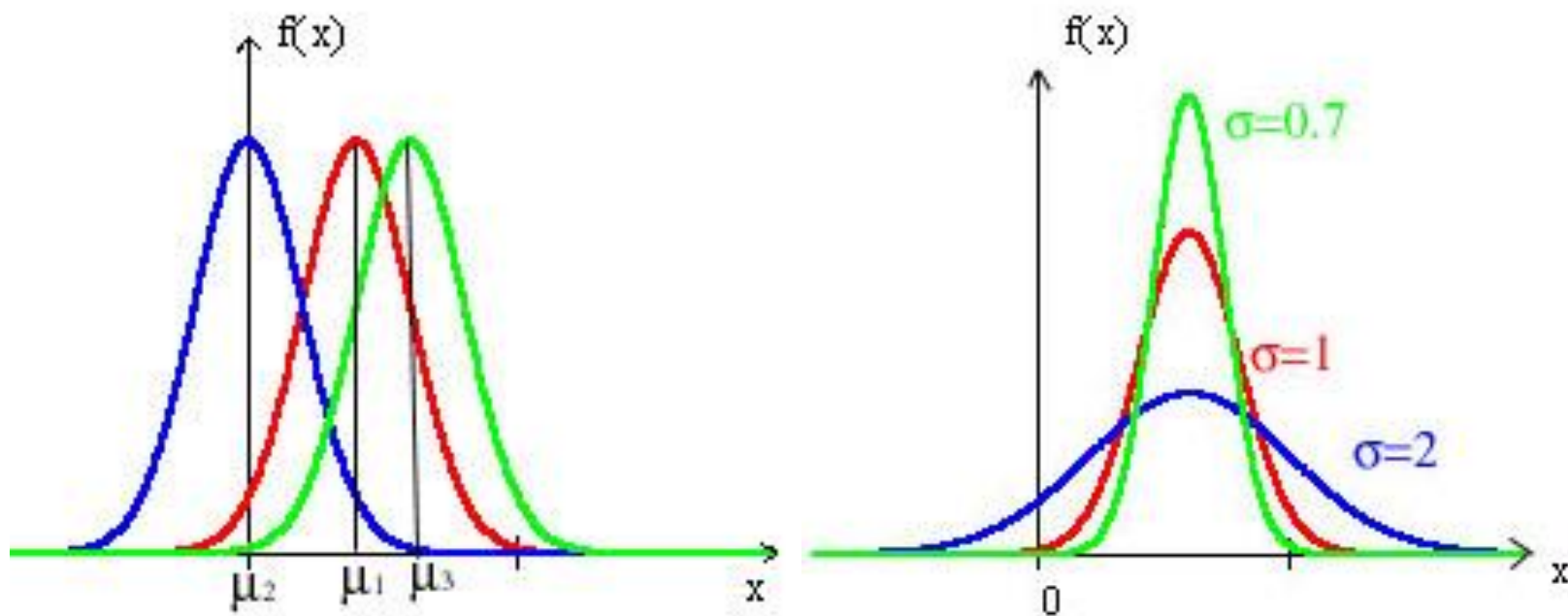
# 正态分布的密度函数的性质与图形



中间高,两边低  
呈“钟形”分布

- 对称性: 关于  $x = \mu$  对称
- 单调性:  $(-\infty, \mu)$  升,  $(\mu, +\infty)$  降
- 拐点:  $(\mu \pm \sigma, \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}})$ ;  $f_{\text{最大}}(\mu) = \frac{1}{\sqrt{2\pi\sigma}}$
- 渐进线:  $f(x)$  以  $x$  轴为渐近线

# 正态分布的图形特点



$\mu$  决定了图形的中心位置， $\sigma$  决定了图形中峰的陡峭程度。

# 正态分布的分布函数

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

计算随机变量落在某个区间内的概率时：

$$P(a < X \leq b) = F(b) - F(a) = \int_a^b f(x) dx$$

不容易求得。需要借助特殊的正态分布——标准正态分布

# 标准正态分布 $\mu=0, \sigma=1$

## Standard Normal distribution

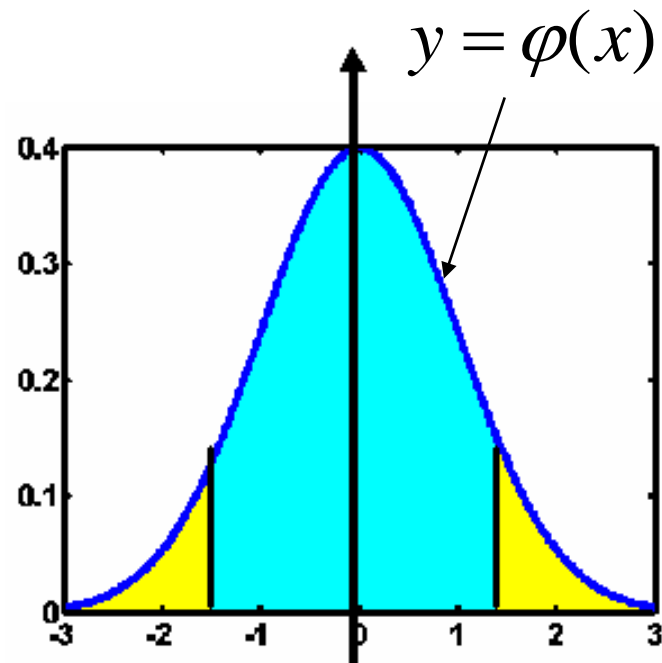
■ **定义：**  $X \sim N(0,1)$ 分布称为标准正态分布

■ **密度函数**

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

■ **分布函数**

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



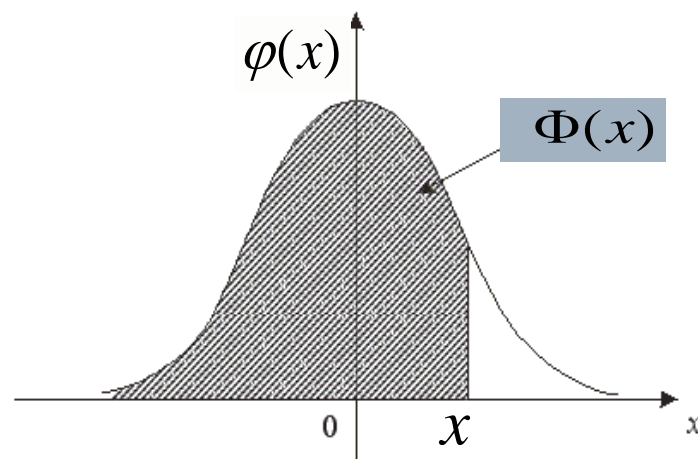


# 标准正态分布的性质

## ■ 分布函数

$$\Phi(x) = P\{X \leq x\}$$

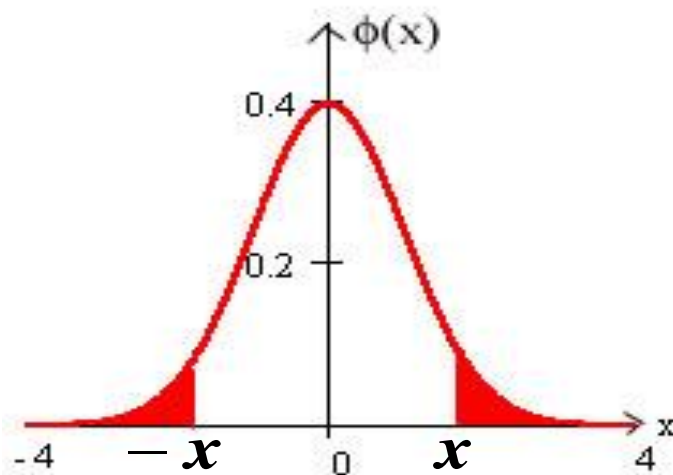
$$= \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$



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



$$\Phi(0) = 0.5 \quad \Phi(+\infty) = 1$$



# 标准正态分布的概率计算

- 公式：设  $X \sim N(0,1)$ ，则有

$$P(a \leq X \leq b) = \Phi(b) - \Phi(a)$$

$$P(X \leq b) = \Phi(b) \quad P(X \geq a) = 1 - \Phi(a)$$

- 查表  $x \geq 0$  时， $\Phi(x)$  的值可以查表。

$$x < 0 \text{ 时, } \Phi(-x) = 1 - \Phi(x)$$

P212

- 例如：  $X \sim N(0,1)$

$$P(1 \leq X \leq 2) = \Phi(2) - \Phi(1) = 0.9772 - 0.8413 = 0.1359$$

$$P(X \leq -1) = \Phi(-1) = 1 - \Phi(1) = 1 - 0.8413 = 0.1587$$

$$P(|X| \leq 1) = \Phi(1) - \Phi(-1) = 2\Phi(1) - 1 = 0.6826$$

# 一般正态分布的标准化

## ■ 定理


如果  $X \sim N(\mu, \sigma^2)$ , 则  $F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right)$

## ■ 概率计算 若 $X \sim N(\mu, \sigma^2)$

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

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

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



查标准正  
态分布表

**例1:** 设随机变量  $X \sim N(18, 2.5^2)$ , 求

①  $P(17 < X < 21)$

② 常数  $k$ , 使  $P(X < k) = 0.2236$

③ 最大的常数  $k$ , 使  $P(X > k) \geq 0.1814$

**解:**  $\mu = 18, \sigma = 2.5$

$$\begin{aligned} P(17 < X < 21) &= \Phi\left(\frac{21-18}{2.5}\right) - \Phi\left(\frac{17-18}{2.5}\right) = \Phi(1.2) - \Phi(-0.4) \\ &= \Phi(1.2) - [1 - \Phi(0.4)] = 0.5403 \end{aligned}$$

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

**例1:** 设随机变量  $X \sim N(18, 2.5^2)$ , 求

② 常数  $k$ , 使  $P(X < k) = 0.2236$

**解:**  $P(X < k) = \Phi\left(\frac{k-18}{2.5}\right) = 0.2236 \Rightarrow$  无法查表

$$\Phi\left(\frac{18-k}{2.5}\right) = 1 - \Phi\left(\frac{k-18}{2.5}\right) = 1 - 0.2236 = 0.7764$$
$$= \Phi(0.76)$$

即  $k = 18 - 0.76 \times 2.5 = 16.1$

**例1:** 设随机变量 $X \sim N(18, 2.5^2)$ , 求

③最大的常数 $k$ , 使  $P(X > k) \geq 0.1814$

**解:** 
$$P(X > k) = 1 - \Phi\left(\frac{k-18}{2.5}\right) \geq 0.1814$$
$$\Phi\left(\frac{k-18}{2.5}\right) \leq 1 - 0.1814 = 0.8186 = \Phi(0.91)$$

由  $\Phi(x)$  的单调性可得

$$\frac{k-18}{2.5} \leq 0.91 \quad \Rightarrow \quad k \leq 20.275$$

# 正态分布的实际应用

某单位招聘155人，按考试成绩录用，共有526人报名，假设报名者的考试成绩  $X \sim N(\mu, \sigma^2)$

已知90分以上的12人，60分以下的83人，若从高分到低分依次录取，某人成绩为78分，问此人能否被录取？

■ 分析 首先求出  $\mu$  和  $\sigma$

然后根据录取率或者分数线确定能否录取

解：成绩 $X$ 服从 $N(\mu, \sigma^2)$

$$\text{录取率为 } \frac{155}{526} \approx 0.2947$$

$$\begin{aligned} \text{又 } P(X > 90) &= \frac{12}{526} \approx 0.0228 \\ &= 1 - \Phi\left(\frac{90 - \mu}{\sigma}\right) \end{aligned} \quad \left. \vphantom{\begin{aligned} P(X > 90) &= \frac{12}{526} \approx 0.0228 \\ &= 1 - \Phi\left(\frac{90 - \mu}{\sigma}\right) \end{aligned}} \right\} \Rightarrow \frac{90 - \mu}{\sigma} \approx 2.0$$

$$\begin{aligned} \text{同理 } P(X < 60) &= \frac{83}{526} \approx 0.1588 \\ &= \Phi\left(\frac{60 - \mu}{\sigma}\right) = 1 - \Phi\left(\frac{\mu - 60}{\sigma}\right) \end{aligned} \quad \left. \vphantom{\begin{aligned} P(X < 60) &= \frac{83}{526} \approx 0.1588 \\ &= \Phi\left(\frac{60 - \mu}{\sigma}\right) = 1 - \Phi\left(\frac{\mu - 60}{\sigma}\right) \end{aligned}} \right\} \Rightarrow \frac{\mu - 60}{\sigma} \approx 1.0$$

$$\text{解得 } \mu = 70, \sigma = 10$$



解

.....

故  $X \sim N(70, 10^2)$

设录取的最低分为  $x$ , 则应有

$$P(X \geq x) = \frac{155}{526} \approx 0.2947$$

$$= 1 - \Phi\left(\frac{x-70}{10}\right)$$

$$\Phi\left(\frac{x-70}{10}\right) = 0.7053 \quad \frac{x-70}{10} \approx 0.54$$

$$x = 75.4$$

某人**78**分,  
可被录取。

# 3σ准则

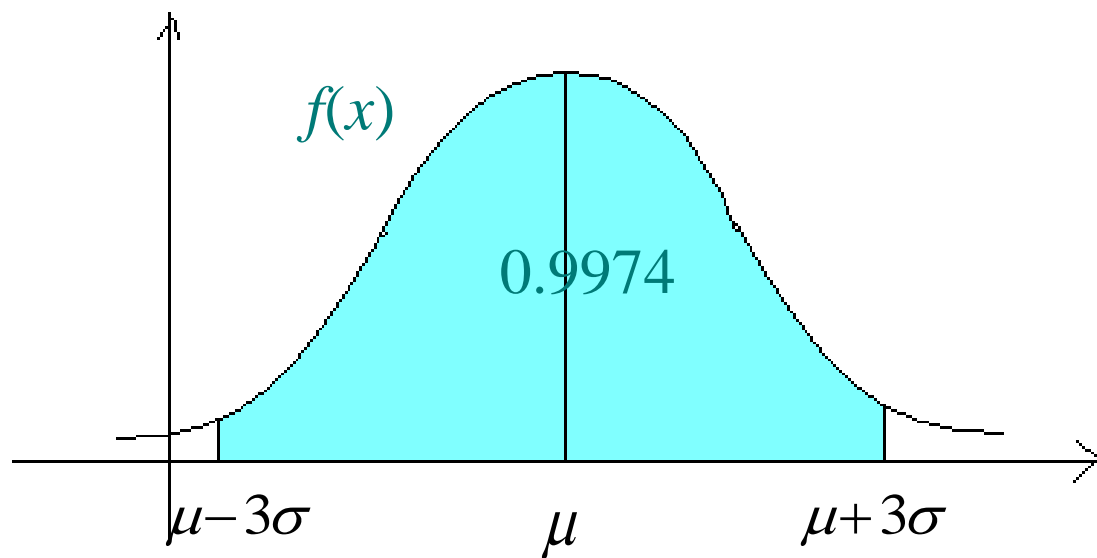
$$X \sim N(\mu, \sigma^2)$$

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

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

$$P(|X - \mu| \leq 3\sigma) = 0.9974$$

X的取值几乎都落入  
区间 $[\mu - 3\sigma, \mu + 3\sigma]$ 内



$$(|X - \mu| > 3\sigma)$$

是小概率事件

# 3 $\sigma$ 准则

在工程应用中，通常认为  $P(|X - \mu| \leq 3\sigma) \approx 1$ ，  
忽略  $(|X - \mu| > 3\sigma)$  的值。

如在质量控制中，常用标准指标值  $\pm 3\sigma$  作两条  
线，当生产过程的指标观察值落在两线之外时发出  
警报. 表明生产出现异常。

# 小结

几个常用的连续型随机变量

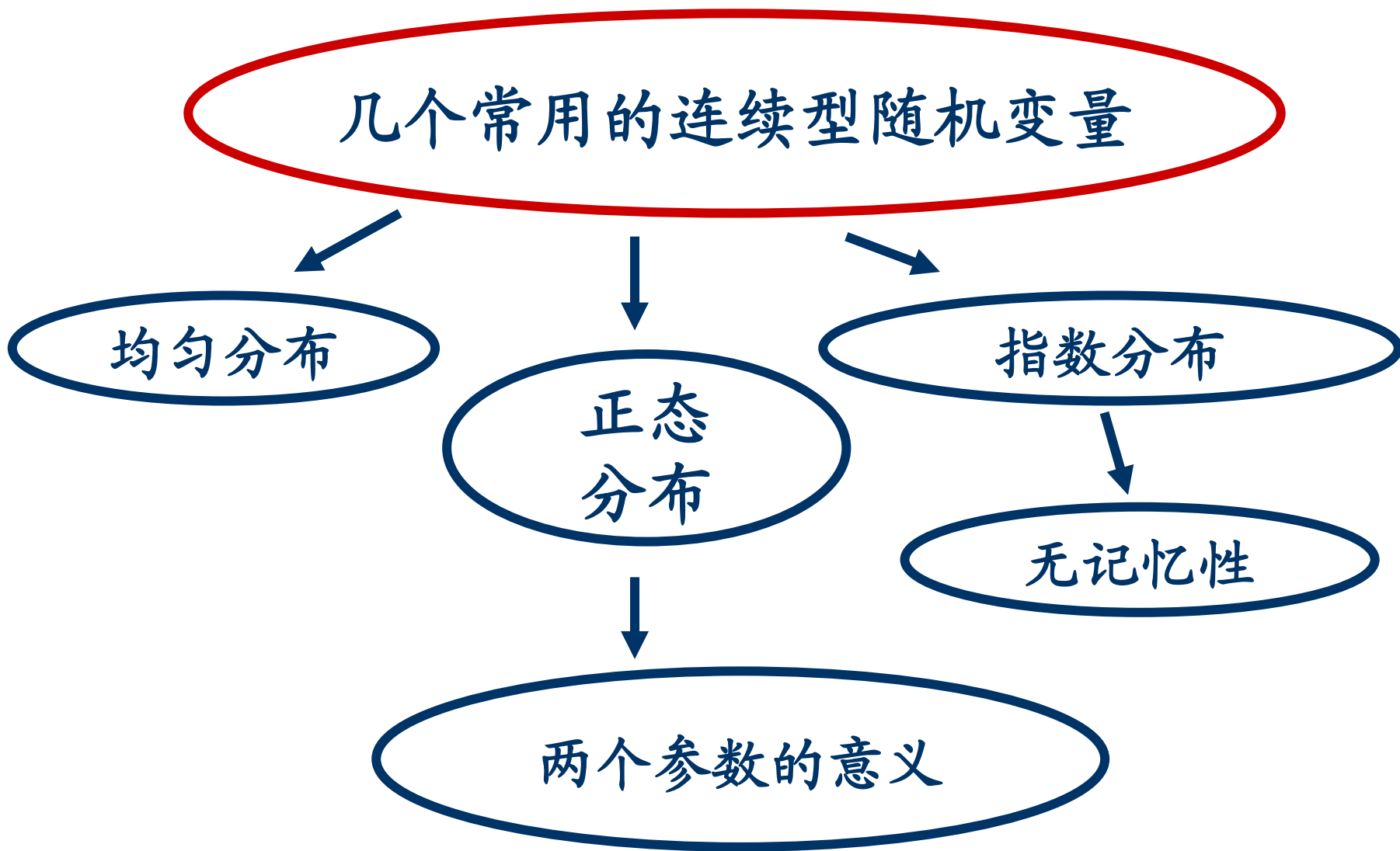
均匀分布

正态  
分布

指数分布

无记忆性

两个参数的意义



# 第八节

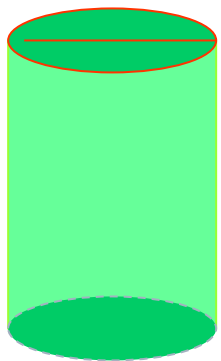
第八节

## 随机变量函数的分布

# 随机变量函数的分布

## ■ 背景

在许多实际问题中，人们常常对随机变量的函数更感兴趣.



比如，已知圆轴截面直径  $d$  的分布

求截面面积  $S = \frac{\pi d^2}{4}$  的分布.

设随机变量  $X$  的分布已知， $Y=g(X)$  (设  $g$  是连续函数)，如何由  $X$  的分布求出  $Y$  的分布？

这个问题无论在实践中还是在理论上都是非常重要的.

# 离散随机变量函数的分布

若 $X$ 为离散型随机变量, 其分布律为

$X$	$x_1$	$x_2$	$x_3$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$x_n$	$\cdot$	$\cdot$	$\cdot$
$p_k$	$p_1$	$p_2$	$p_3$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$p_n$	$\cdot$	$\cdot$	$\cdot$

则随机变量 $X$ 的函数  $Y = g(X)$  的分布律为

$Y$	$g(x_1)$	$g(x_2)$	$g(x_3)$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$g(x_n)$	$\cdot$	$\cdot$	$\cdot$
$p_k$	$p_1$	$p_2$	$p_3$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$p_n$	$\cdot$	$\cdot$	$\cdot$

如果 $g(x_i)$ 与 $g(x_j)$ 相同, 此时将两项合并, 对应概率相加.

**例：** 设随机变量 $X$ 的分布律为

$X$	-1	0	1	2
$p_k$	0.2	0.3	0.4	0.1

求 $Y=2X^2+1$ 的分布律.

**解：** 由题设可得如下表格

$Y=2X^2+1$	3	1	3	9
$X$	-1	0	1	2
$p_k$	0.2	0.3	0.4	0.1

所以， $Y=2X^2+1$ 的分布律为

$Y$	1	3	9
$p_k$	0.3	0.6	0.1



**例：** 设随机变量X的分布律为

$$P(X = k) = \frac{1}{2^k}, k = 1, 2, \dots$$

求随机变量  $Y = \cos\left(\frac{X}{2}\pi\right)$  的分布律。

**解：** 由题设可得如下表格

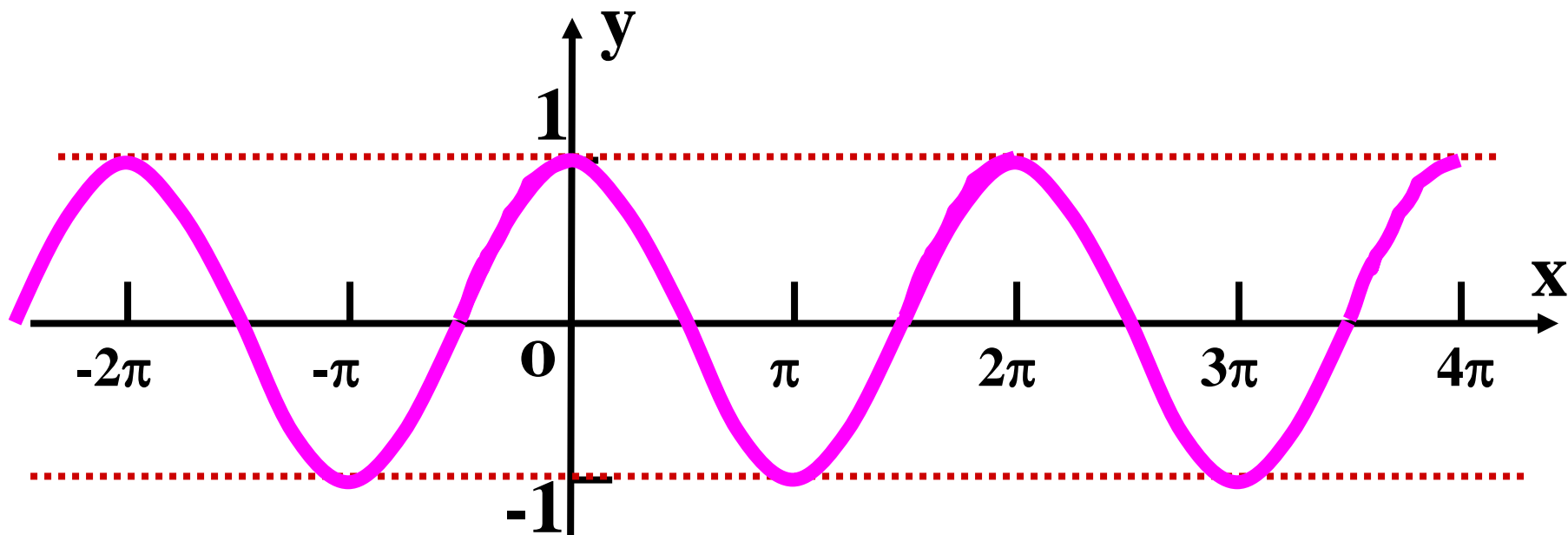
Y	0	-1	0	1	...	$\cos\left(\frac{n}{2}\pi\right)$	...
X	1	2	3	4	...	n	...
$p_k$	$\frac{1}{2}$	$\frac{1}{2^2}$	$\frac{1}{2^3}$	$\frac{1}{2^4}$	...	$\frac{1}{2^n}$	...

所以，Y的分布律为

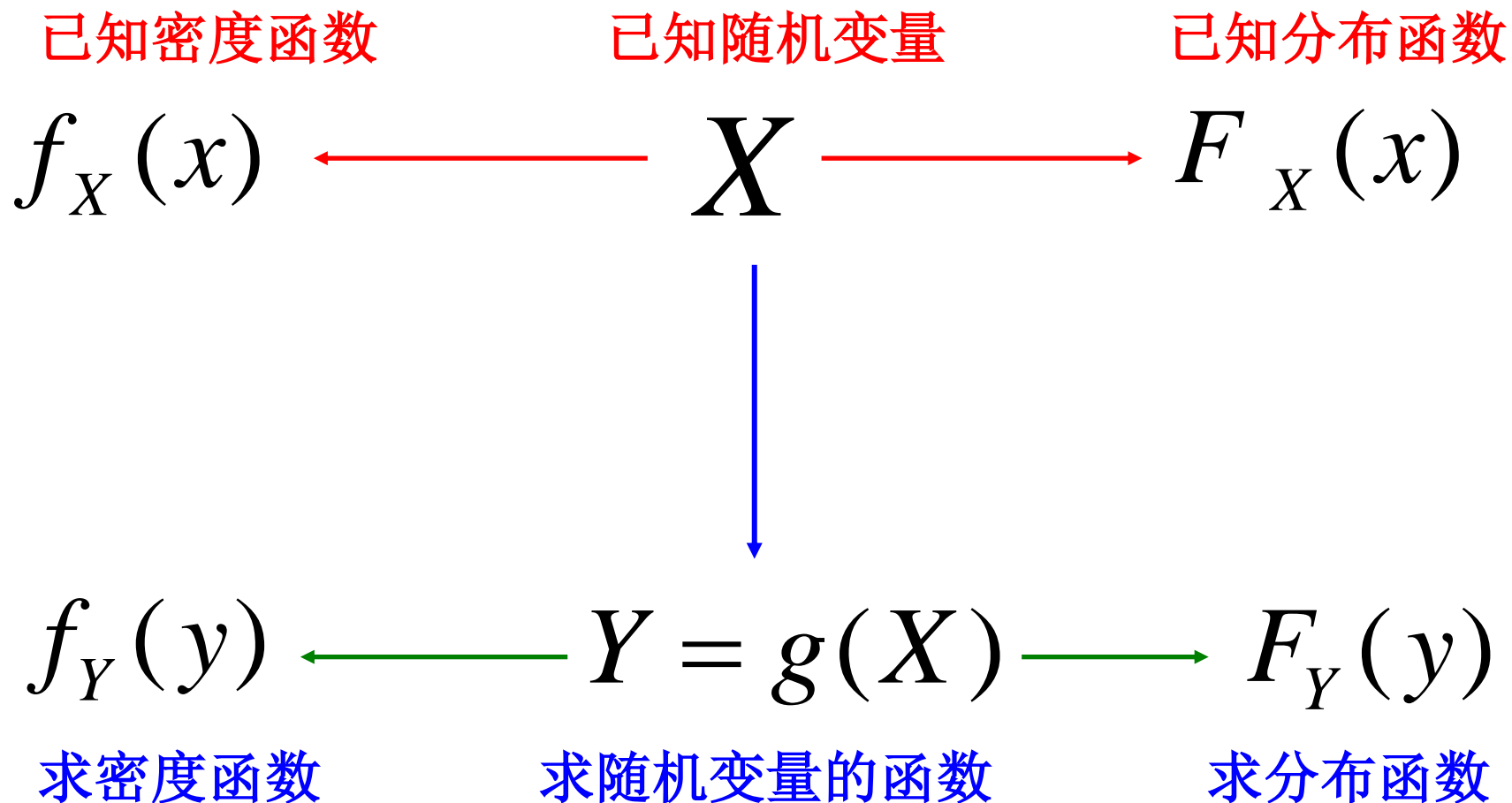
Y	-1	0	1
$p_k$	$\frac{4}{15}$	$\frac{2}{3}$	$\frac{1}{15}$

## 余弦曲线

---



# 连续型随机变量函数的分布



# 连续型随机变量函数的分布

设  $X$  为一个连续型随机变量，已知其分布函数  $F_X(x)$  和概率密度函数为  $f_X(x)$ ，求随机变量函数  $Y=g(X)$  的分布函数  $F_Y(y)$  和概率密度函数为  $f_Y(y)$ 。

## ■ 一般方法

分布函数  
法

### (1) 先求 $Y$ 的分布函数 $F_Y(y)$

$$\begin{aligned} F_Y(y) & \xlongequal{\text{根据分布函数的定义}} P(Y \leq y) = P(g(X) \leq y) \\ & = P\left(X \in \{x | g(x) \leq y\}\right) = \int_{\{x | g(x) \leq y\}} f_X(x) dx \end{aligned}$$

### (2) 对 $F_Y(y)$ 求导，得到 $f_Y(y)$      $f_Y(y) = F'_Y(y)$

**例：** 设随机变量X的密度函数为

$$f(x) = \begin{cases} \frac{x}{8}, & 0 < x < 4 \\ 0, & \text{其它} \end{cases}$$

求随机变量 $Y=2X+8$ 的概率密度。

**解：** (1) 先求 $Y=2X+8$ 的分布函数 $F_Y(y)$ .

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(2X + 8 \leq y) \\ &= P\left(X \leq \frac{y-8}{2}\right) = \int_{-\infty}^{\frac{y-8}{2}} f(x) dx \end{aligned}$$

(2) 求 $Y=2X+8$ 的概率密度

$$f_Y(y) = F'_Y(y) = \left( \int_{-\infty}^{\frac{y-8}{2}} f(x) dx \right)',$$

$$= f\left(\frac{y-8}{2}\right) \left(\frac{y-8}{2}\right)',$$

$$= \begin{cases} \frac{1}{8} \left(\frac{y-8}{2}\right) \cdot \frac{1}{2}, & 0 < \frac{y-8}{2} < 4 \\ 0, & \text{其它} \end{cases}$$

$$= \begin{cases} \frac{y-8}{32}, & 8 < y < 16 \\ 0, & \text{其它} \end{cases}$$

**例：** 设  $X$  具有概率密度  $f_X(x)$ , 求  $Y=X^2$  的概率密度.

**解：** 设  $Y$  和  $X$  的分布函数分别为  $F_Y(y)$  和  $F_X(x)$ ,

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y)$$

注意到  $Y=X^2 \geq 0$ , 故当  $y < 0$  时,  $F_Y(y) = 0$ .

当  $y \geq 0$  时,  $F_Y(y) = P(Y \leq y) = P(X^2 \leq y)$

$$\begin{aligned} &= P(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}) \left( = \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x) dx \right) \end{aligned}$$

求导可得:

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \begin{cases} \frac{1}{2\sqrt{y}} [f_X(\sqrt{y}) + f_X(-\sqrt{y})], & y \geq 0 \\ 0, & y < 0 \end{cases}$$

**练习：** 设  $X$  具有概率密度  $f_X(x) = \begin{cases} e^{-x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$  ,

求  $Y=e^X$  的概率密度.



## 书上P54的例2

设随机变量 $X$ 服从正态分布  $N(\mu, \sigma^2)$ ，求  $Y=aX+b$  的概率密度。

结论：  $Y \sim N(a\mu + b, (a\sigma)^2)$

## ■ 定理 正态分布的线性函数仍服从正态分布

设  $X \sim N(\mu, \sigma^2)$ ,  $Y = aX + b (a \neq 0)$ , 则

$$Y \sim N(a\mu + b, (a\sigma)^2)$$

## ■ 推论

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

下面给出一个定理，在满足定理条件时，可直接用它求出随机变量函数的概率密度。

**定理：** 设  $X$  是一个取值于区间  $[a, b]$ ，具有概率密度  $f(x)$  的连续型  $r.v$ ，又设  $y=g(x)$  处处可导，且对于任意  $x$ ，恒有  $g'(x) > 0$  或恒有  $g'(x) < 0$ ，则  $Y=g(X)$  是一个连续型  $r.v$ ，它的概率密度为

$$f_Y(y) = \begin{cases} f[h(y)] \left| \frac{dh(y)}{dy} \right|, & \alpha < y < \beta \\ 0, & \text{其它} \end{cases}$$

其中，  $\alpha = \min_{a \leq x \leq b} g(x), \beta = \max_{a \leq x \leq b} g(x),$

$x=h(y)$  是  $y=g(x)$  的反函数。

此定理的  
证明与前  
面的解题  
思路类似



设圆的半径 $X$ 服从区间 $(1, 2)$ 上的均匀分布，  
求圆面积的分布密度函数。

答案： 设 $Y=\pi X^2$

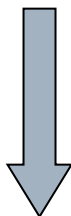
$$f_Y(y) = \begin{cases} \frac{1}{2\sqrt{\pi y}}, & \pi < y < 4\pi \\ 0, & \text{其它} \end{cases}$$

## 第三章      多维随机变量及其分布

- 二维随机变量及其分布
- 随机变量的独立性
- 两个随机变量函数的分布

从本讲起，我们开始第三章的学习。  
它是第二章内容的推广。

## 一维随机变量及其分布

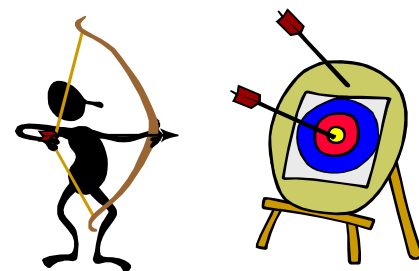


## 多维随机变量及其分布

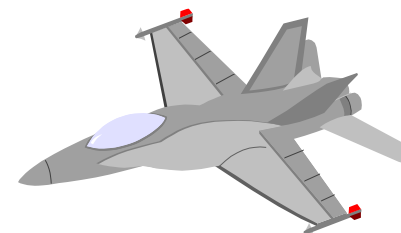
由于从二维推广到多维一般无实质性的困难，我们重点讨论二维随机变量。

到现在为止，我们只讨论了一维  $r.v.$  及其分布。  
但有些随机现象用一个随机变量来描述还不够，而  
需要用几个随机变量来描述。

在打靶时，命中点的位置是  
由一对  $r.v.$  (两个坐标) 来确定的。



飞机的重心在空中的位置是由  
三个  $r.v.$  (三个坐标) 来确定的等等。



钢的基本指标——含碳量  $X$ ，含硫量  $Y$  和硬度  $Z$   
三个  $r.v.$  来确定。



一般地,设 $E$ 是一个随机试验,它的样本空间是 $\Omega=\{\omega\}$ , 设  $X_1 = X_1(\omega), X_2 = X_2(\omega), \dots, X_n = X_n(\omega)$  定义在 $\Omega$ 上的随机变量, 由它们构成的一个 $n$ 维向量  $(X_1, X_2, \dots, X_n)$  叫做 **$n$ 维随机向量**或 **$n$ 维随机变量**。

设试验 $E$ 的样本空间为 $\Omega$ ,  $X=X(\omega)$ 与 $Y=Y(\omega)$ 是定义在 $\Omega$ 上的两个随机变量, 由它们构成的向量  $(X, Y)$  称为**二维随机变量**。

二维随机向量 $(X, Y)$ 的性质不仅与 $X$ 和 $Y$ 的性质有关, 而且还依赖于 $X$ 和 $Y$ 之间的相互关系。因此, 必须把 $(X, Y)$ 作为一个整体来看待, 加以研究。

# 第一节

## 二维离散型随机变量

## 二维离散型随机变量之联合分布律

■**定义**：如果二维随机变量 $(X, Y)$ 的每个分量都是离散型随机变量，则称 $(X, Y)$ 是二维离散型随机变量。

即：二维随机变量 $(X, Y)$ 所有可能取值是有限组或可列无限组 $(x_i, y_j)$ ,  $i, j=1, 2, \dots$ ，将 $(X, Y)$ 取每组值的概率

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

称为二维离散型随机变量 $(X, Y)$ 的联合分布律(或联合分布列)

**注意：** 记号  $(X = x_i, Y = y_j)$

表示随机事件  $(X = x_i)$  与  $(Y = y_j)$  的积事件

即：

$$(X = x_i, Y = y_j) = (X = x_i) \cap (Y = y_j)$$

# 性 质

一维离散型随机变量  $X$  的概率分布列:

$$P(X = x_k) = p_k, \\ k = 1, 2, \dots$$

$$\begin{cases} 0 \leq p_k \leq 1, & k = 1, 2, \dots \\ \sum_k p_k = 1 \end{cases}$$

二维离散型随机变量  $(X, Y)$  的联合概率分布:

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

$$\begin{cases} 0 \leq p_{ij} \leq 1, & i, j = 1, 2, \dots \\ \sum_i \sum_j p_{ij} = 1 \end{cases}$$

也可用表格来表示随机变量 $X$ 和 $Y$ 的联合分布律

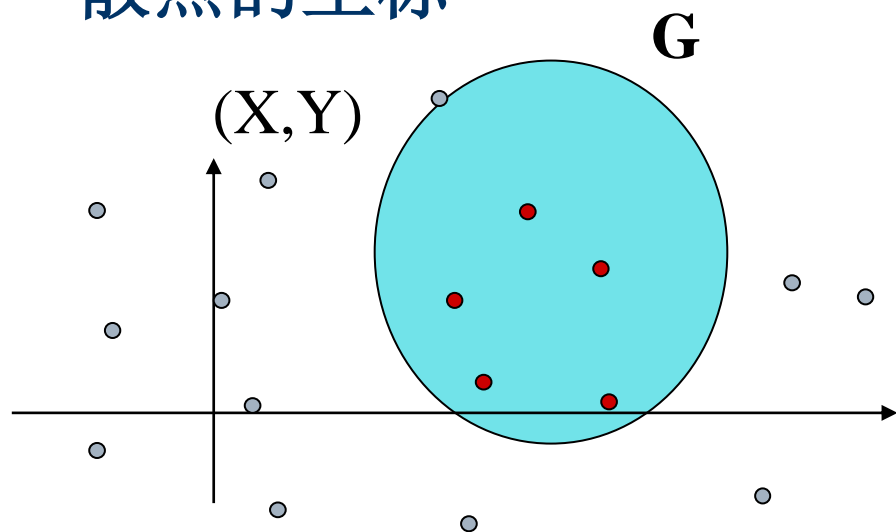
$X \backslash Y$	$y_1$	$y_2$	$\cdots$	$y_j$	$\cdots$
$x_1$	$p_{11}$	$p_{12}$	$\cdots$	$p_{1j}$	$\cdots$
$x_2$	$p_{21}$	$p_{22}$	$\cdots$	$p_{2j}$	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	
$x_i$	$p_{i1}$	$p_{i2}$	$\cdots$	$p_{ij}$	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	

# 随机变量的几何表示

一维随机变量 $X$ —— $\mathbf{R}^1$ 上的随机点坐标

二维随机变量 $(X,Y)$ —— $\mathbf{R}^2$ 上的随机点坐标

二维离散型随机变量 $(X, Y)$ 的取值可看作平面上的离散点的坐标



联合分布律完全反映了  
 $(X,Y)$ 的概率性质：设 $G$   
是一平面区域，则

$$P((X,Y) \in G) = \sum_{(x_i, y_j) \in G} p_{ij}$$

**例1:** 设随机变量  $X$  在 1,2,3,4 四个整数中等可能地取值, 另一个随机变量  $Y$  在  $1 \sim X$  中等可能地取一整数. 试求  $(X, Y)$  的联合分布律及  $P(X \leq 3, Y \leq 2)$

**解:**  $(X = i, Y = j)$  的取值情况是:  $i = 1, 2, 3, 4, j \leq i$ .

$$P(X = i, Y = j) = P(X = i)P(Y = j|X = i) = \frac{1}{4} \cdot \frac{1}{i},$$

于是,  $(X, Y)$  的联合分布律为

$$\begin{aligned} P(X \leq 3, Y \leq 2) \\ &= \frac{1}{4} + 0 + \frac{1}{8} + \frac{1}{8} + \frac{1}{12} + \frac{1}{12} \\ &= \frac{2}{3} \end{aligned}$$

$X \backslash Y$	1	2	3	4
1	1/4	0	0	0
2	1/8	1/8	0	0
3	1/12	1/12	1/12	0
4	1/16	1/16	1/16	1/16



**例2:** 一个袋中有三个球,依次标有数字 1, 2, 2,从中任取一个,不放回袋中,再任取一个,设每次取球时,各球被取到的可能性相等,以  $X, Y$  分别记第一次和第二次取到的球上标有的数字,求  $(X, Y)$  的联合分布律 **1 2 2**

**解:**  $(X, Y)$  的可能取值为  $(1, 2), (2, 1), (2, 2)$ .

$$P(X=1, Y=2) = \frac{1}{3} \cdot \frac{2}{2} = \frac{1}{3},$$

$$P(X=2, Y=1) = \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3},$$

$$P(X=2, Y=2) = \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}.$$

故  $(X, Y)$  的分布律为

		$Y$	
		1	2
$X$	1	0	1/3
	2	1/3	1/3

二维联合分布全面地反映了二维随机变量 $(X,Y)$ 的取值及其概率规律，是将 $(X,Y)$ 当做一个整体来考虑。

而单个随机变量 $X$ ， $Y$ 也具有自己的概率分布。

那么要问：二者之间有什么关系呢？

下面，我们就来探求这个问题。

## 二维离散型 r.v.——边缘分布律

如果二维离散型随机变量(X,Y)的联合分布律为

$$P(X = x_i, Y = y_j) = p_{ij} \quad i, j = 1, 2, 3, \dots$$

关于X的边缘分布律为:  $\left( \{X = x_i\} = \bigcup_j \{X = x_i, Y = y_j\} \right)$

$$P(X = x_i) = \sum_j P(X = x_i, Y = y_j) = \sum_j p_{ij} \quad (i = 1, 2, \dots)$$

关于Y的边缘分布律为:

$$P(Y = y_j) = \sum_i P(X = x_i, Y = y_j) = \sum_i p_{ij} \quad (j = 1, 2, \dots)$$

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

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

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

我们常将边缘分布律写在联合分布律表格的边缘上，由此得出边缘分布这个名词。

# 联合分布与边缘分布

$X \backslash Y$	$y_1$	$y_2$	$y_3$	$\cdots$	$P_X$
$x_1$	$p_{11}$	$p_{12}$	$p_{13}$	$\cdots$	$p_{1\cdot}$
$x_2$	$p_{21}$	$p_{22}$	$p_{23}$	$\cdots$	$p_{2\cdot}$
$x_3$	$p_{31}$	$p_{32}$	$p_{33}$	$\cdots$	$p_{3\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$P_Y$	$p_{\cdot 1}$	$p_{\cdot 2}$	$p_{\cdot 3}$	$\cdots$	<b>1</b>

关于X的边缘分布  $p_{i\cdot} = P(X = x_i) = \sum p_{ij}$

关于Y的边缘分布  $p_{\cdot j} = P(Y = y_j) = \sum_i p_{ij}$

**例:** 已知下列联合分布律,求其边缘分布律.

$X \backslash Y$		0	1
0	0	$\frac{16}{49}$	$\frac{12}{49}$
	1	$\frac{12}{49}$	$\frac{9}{49}$

解:

$Y \backslash X$	0	1	$p_{\cdot j} = P(Y = y_j)$
0	$\frac{16}{49}$	$\frac{12}{49}$	$\frac{4}{7}$
1	$\frac{12}{49}$	$\frac{9}{49}$	$\frac{3}{7}$
$p_{i \cdot} = P(X = x_i)$	$\frac{4}{7}$	$\frac{3}{7}$	1

注意: 联合分布  $\longleftrightarrow$  边缘分布

# 离散型随机变量——独立性的判别

若  $P(AB) = P(A)P(B)$ ，则称事件A与B相互独立。

■ 若  $(X, Y)$  是二维离散型随机变量，

$X$ 与 $Y$ 独立  $\iff P(X = x_i, Y = y_j) = P(X = x_i) \cdot P(Y = y_j)$

即：联合分布律等于边缘分布律的乘积

$\forall (x_i, y_j)$

逐个验证等式  $P(X = x_i, Y = y_j) = P(X = x_i) \cdot P(Y = y_j)$



# 离散型随机变量独立性的判别

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

$Y \backslash X$	0	1
0	$\frac{1}{2}$	$\frac{1}{4}$
1	$\frac{1}{4}$	0

$X$ 与 $Y$ 不独立

$Y \backslash X$	0	1
0	$\frac{1}{4}$	$\frac{1}{4}$
1	$\frac{1}{4}$	$\frac{1}{4}$

$X$ 与 $Y$ 独立

**例：** 设随机变量 $X_1$ 和 $X_2$ 的概率分布如下，而且  
 $P(X_1 X_2 = 0) = 1$

$X_1$	-1	0	1
$P$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

$X_2$	0	1
$P$	$\frac{1}{2}$	$\frac{1}{2}$

(1) 求 $X_1$ 和 $X_2$ 的联合分布列； (2) 问 $X_1$ 和 $X_2$ 是否独立？

$X_2 \backslash X_1$	-1	0	1
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	0	$\frac{1}{2}$	0

(2)  $X_1$ 和 $X_2$ 不独立

# 第二节

## 二维连续型随机变量

## 二维随机变量——联合分布函数

■**定义：**若 $(X,Y)$ 为二维随机变量，对于任意的实数 $x, y$ ，二元函数

$$\begin{aligned} F(x, y) &= P(X \leq x, Y \leq y) \\ &= P\{(X \leq x) \cap (Y \leq y)\} \\ -\infty &< x, y < +\infty \end{aligned}$$

一维随机变量  
 $X$ 的分布函数

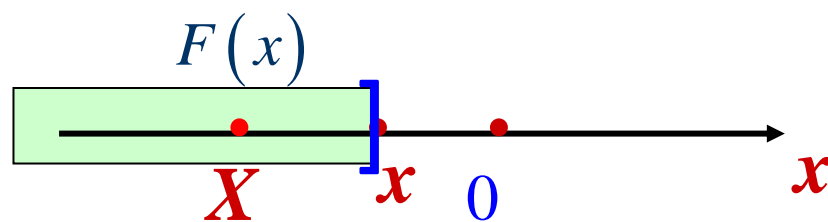
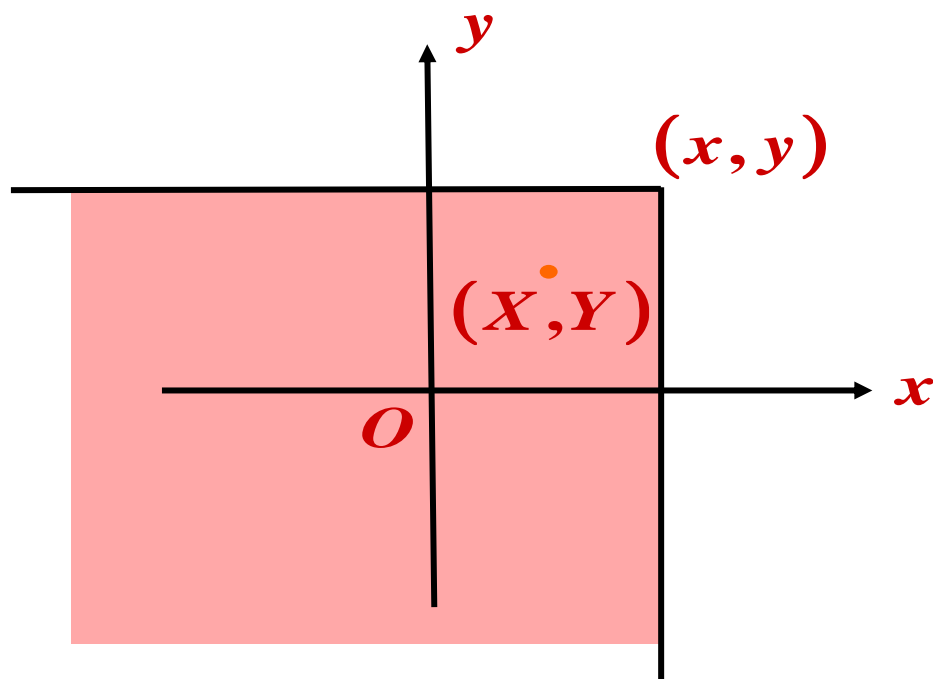
$$\begin{aligned} F(x) &= P(X \leq x) \\ -\infty &< x < \infty \end{aligned}$$

称为二维随机变量 $(X,Y)$ 的联合分布函数。

# 联合分布函数的几何解释

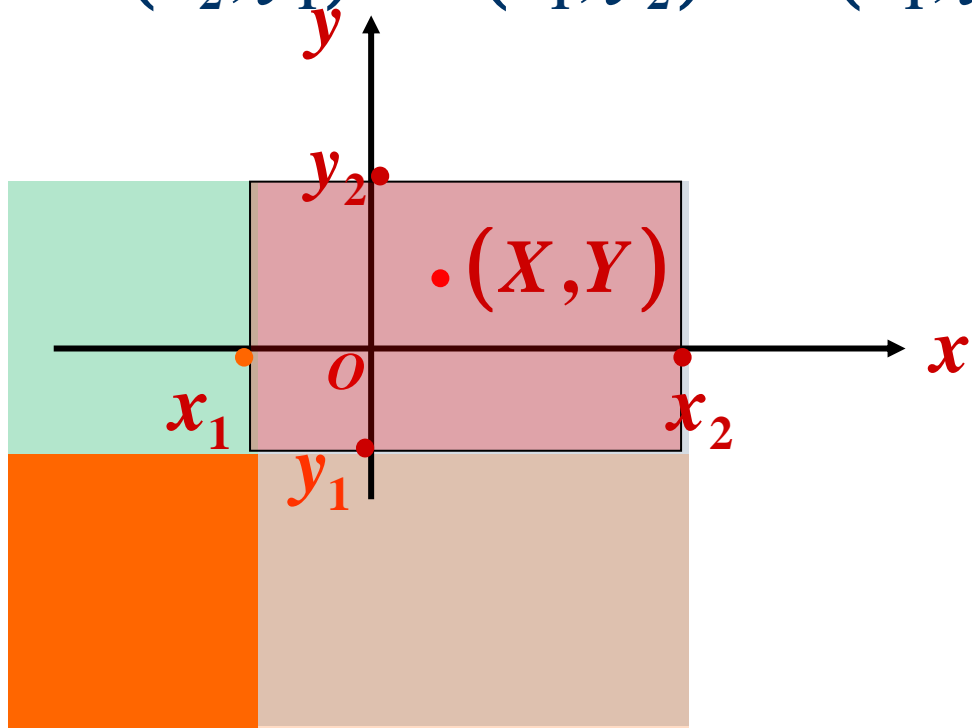
将二维随机变量 $(X,Y)$ 看成是平面上随机点的坐标

联合分布函数 $F(x,y)$ 在点 $(x,y)$ 处的函数值就是随机点 $(X,Y)$ 落在下面左图所示的,以点 $(x,y)$ 为顶点而位于该点左下方的无穷矩形域内的概率.



一维随机变量的分布函数

随机点 $(X,Y)$  落在矩形域  $x_1 < x \leq x_2, y_1 < y \leq y_2$   
内的概率为  $P(x_1 < X \leq x_2, y_1 < Y \leq y_2)$   
 $= F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1)$

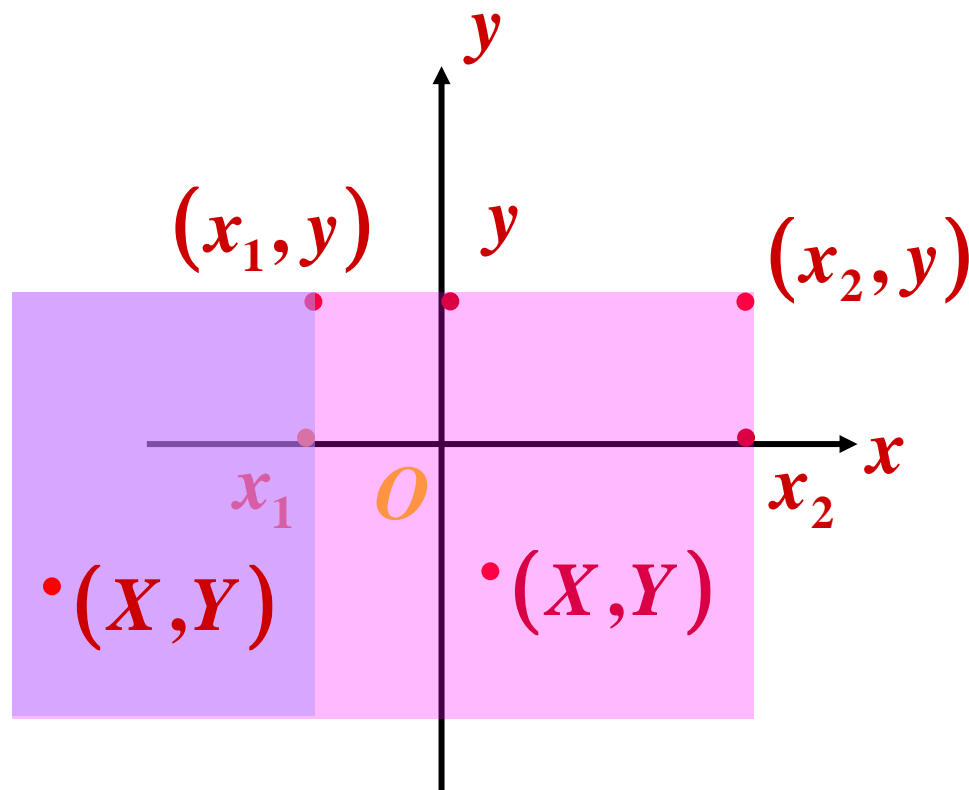


# 联合分布函数 $F(x, y)$ 的性质

1. 单调性:  $F(x, y)$  是关于变量  $x$  和  $y$  的不减函数;

对任意固定的  $y \in R$   
及  $\forall x_1, x_2 \in R$ , 当  $x_1 < x_2$   
时  $F(x_1, y) \leq F(x_2, y)$ ;

对任意固定的  $x \in R$   
及  $\forall y_1, y_2 \in R$ , 当  $y_1 < y_2$   
时  $F(x, y_1) \leq F(x, y_2)$ ;



# 联合分布函数 $F(x, y)$ 的性质

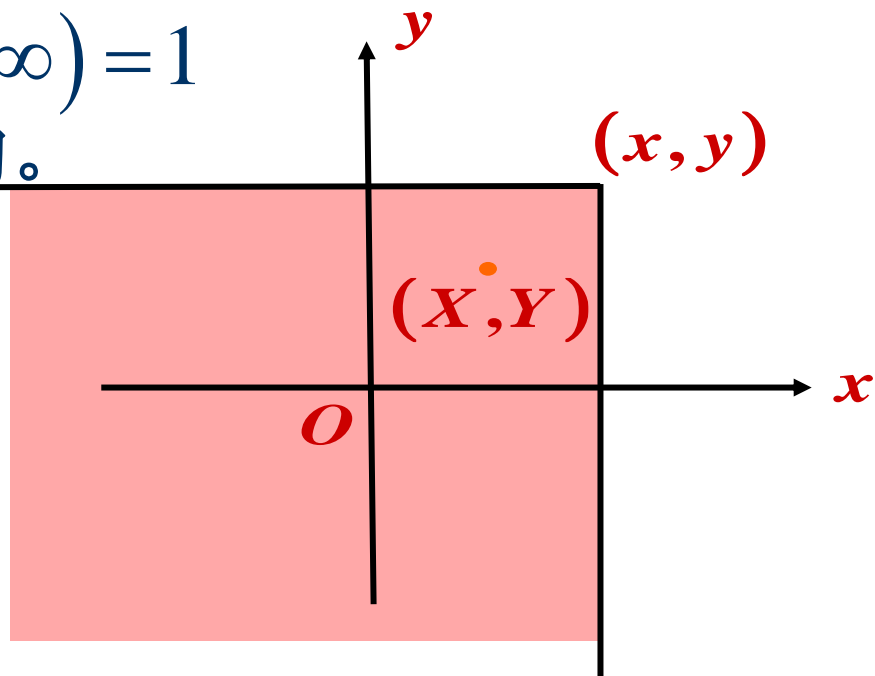
2. 规范性:  $0 \leq F(x, y) \leq 1$ , 且

对任意固定的  $x \in R$ ,  $F(x, -\infty) = 0$ ,

对任意固定的  $y \in R$ ,  $F(-\infty, y) = 0$ ,

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

3.  $F(x, y)$  关于  $x, y$  是右连续的。





## 二维离散型r.v.的联合概率分布与联合分布函数

设二维离散型随机向量  $(X, Y)$  的联合概率分布为

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

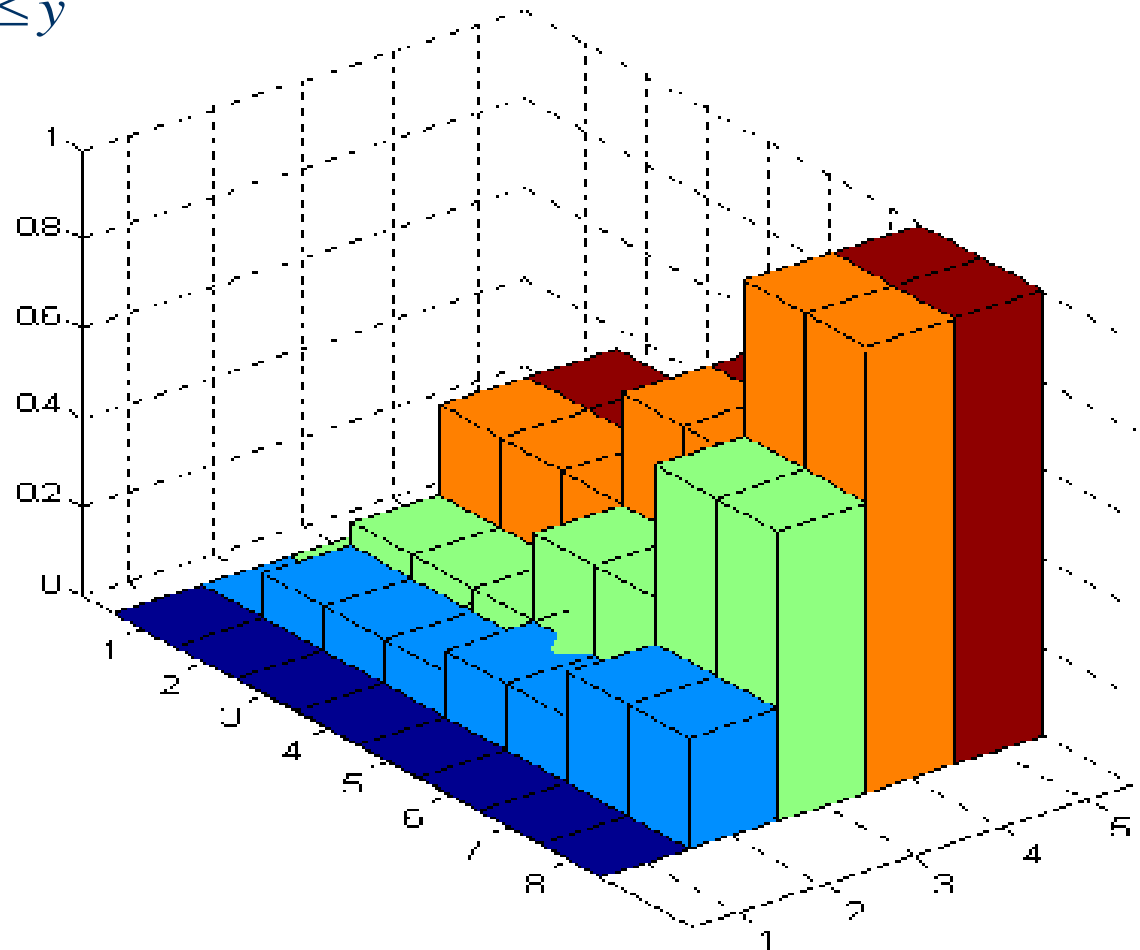
于是,  $(X, Y)$  的联合分布函数为

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

$$= P\left(\bigcup_{x_i \leq x, y_j \leq y} (X = x_i, Y = y_j)\right)$$

$$= \sum_{x_i \leq x} \sum_{y_j \leq y} p_{ij}$$

$$F(x, y) = \sum_{x_i \leq x} \sum_{y_j \leq y} p_{ij}$$



# 二维随机变量——边缘分布函数

二维随机变量  $(X,Y)$  作为一个整体,具有联合分布函数  $F(x,y)$ .

其分量  $X$  和  $Y$  都是随机变量,也有各自的分布函数,分别记作  $F_X(x)$  和  $F_Y(y)$ ,依次称为二维随机变量  $(X,Y)$  关于  $X$  和  $Y$  的边缘分布函数.

边缘分布函数完全由联合分布函数确定.

$$\begin{aligned} F_X(x) &= P(X \leq x) = P(X \leq x, Y < +\infty) = F(x, +\infty) \\ F_Y(y) &= P(Y \leq y) = P(X < +\infty, Y \leq y) = F(+\infty, y) \end{aligned}$$

**注意：联合和边缘都是相对的**

$X$ 与 $Y$ 的边缘分布函数实质上就是一维随机变量 $X$ 或 $Y$ 的分布函数，称其为边缘分布函数的是相对于 $(X, Y)$ 的联合分布而言的。

同样地， $(X, Y)$ 的联合分布函数 $F(x, y)$ 是相对于 $(X, Y)$ 的分量 $X$ 和 $Y$ 的分布而言的。

$$F_X(x) = F(x, +\infty)$$

$$F_Y(y) = F(+\infty, y)$$

**例：** 设 $(X,Y)$ 的联合分布函数为：

$$F(x, y) = \begin{cases} 1 - e^{-0.5x} - e^{-0.5y} + e^{-0.5(x+y)}, & x > 0, y > 0 \\ 0, & \text{其他} \end{cases}$$

求 $(X,Y)$  关于 $X$ 的边缘分布函数  $F_X(x)$ .

**解：**  $(X,Y)$ 关于 $X$ 的边缘分布函数为：

$$\begin{aligned} F_X(x) &= F(x, +\infty) = \lim_{y \rightarrow +\infty} F(x, y) \\ &= \begin{cases} \lim_{y \rightarrow +\infty} (1 - e^{-0.5x} - e^{-0.5y} + e^{-0.5(x+y)}), & x > 0 \\ 0, & x \leq 0 \end{cases} \\ &= \begin{cases} 1 - e^{-0.5x}, & x > 0 \\ 0, & x \leq 0 \end{cases} \end{aligned}$$

即 $X$ 服从参数  $\lambda = 0.5$  的指数分布.

# 联合分布函数——随机变量独立性的判断

若  $P(AB) = P(A)P(B)$ ，则称事件A与B相互独立。

◆ 随机变量独立的定义：

设  $X, Y$  是两个随即变量，对任意的  $x, y$ ，若

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

则称  $X$  与  $Y$  相互独立。

用联合分布函数与边缘分布函数表示上式，就是

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

【表明】两个r.v相互独立时，它们的联合分布函数等于两个边缘分布函数的乘积。

# 二维连续型随机变量

比较:

- ◆ 二维离散型随机变量: 取值有限组或无限可列组  
刻画方法: 联合分布律、联合分布函数
- ◆ 二维连续型随机变量: 取值可充满一个平面区域  
刻画方法: 联合概率密度函数、联合分布函数

# 二维连续型随机变量——联合密度函数

■**定义**：设二维随机向量 $(X, Y)$ 的联合分布函数为 $F(x, y)$ ，如果存在一个非负函数 $f(x, y)$ ，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 $(X, Y)$ 是二元连续型随机变量， $f(x, y)$ 称为二元随机变量 $(X, Y)$ 的联合概率密度函数，简称概率密度。

一元连续型随机变量 $X$ 有

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

$$-\infty < x < +\infty$$



# 二维连续型随机变量的性质

一维连续型随机变量  $X$  的概率密度:

$$\begin{cases} f(x) \geq 0, \\ \int_{-\infty}^{+\infty} f(x) dx = 1. \end{cases}$$

$$\begin{aligned} &P(a \leq X \leq b) \\ &= \int_a^b f(x) dx, \end{aligned}$$

二维连续型随机变量  $(X, Y)$  的联合概率密度:

$$\begin{cases} f(x, y) \geq 0 \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1 \end{cases}$$

$$\begin{aligned} &P((X, Y) \in G) \\ &= \iint_G f(x, y) dx dy. \end{aligned}$$

# 联合概率密度和联合分布函数的关系

## ■ 积分关系

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

## ■ 导数关系

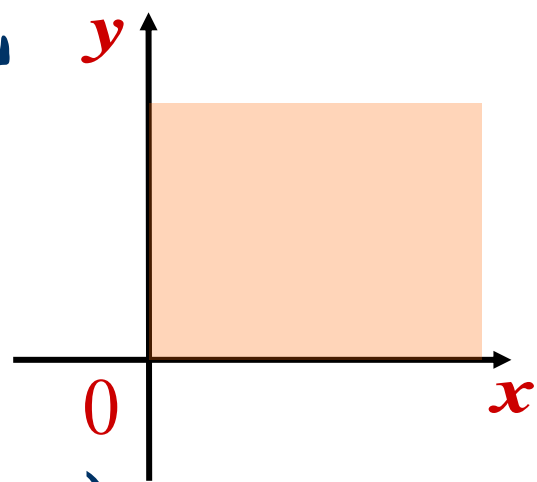
$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}$$

在  $f(x, y)$  的连续点处

**例1:** 已知 $f(x,y)$ , 求其他

**例:** 设 $(X,Y)$ 的联合概率密度是

$$f(x,y) = \begin{cases} Ae^{-(2x+y)}, & x > 0, y > 0, \\ 0, & \text{其它.} \end{cases}$$



求 (1) 常数 $A$ ; (2) 求联合分布函数  $F(x,y)$ ;  
(3) 求概率  $P(Y \leq X)$

**解:** (1) 由  $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y) dx dy = 1$

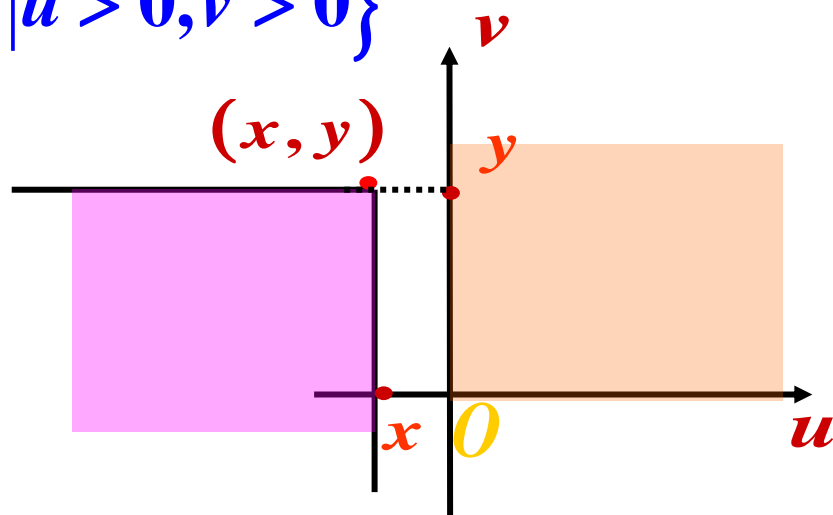
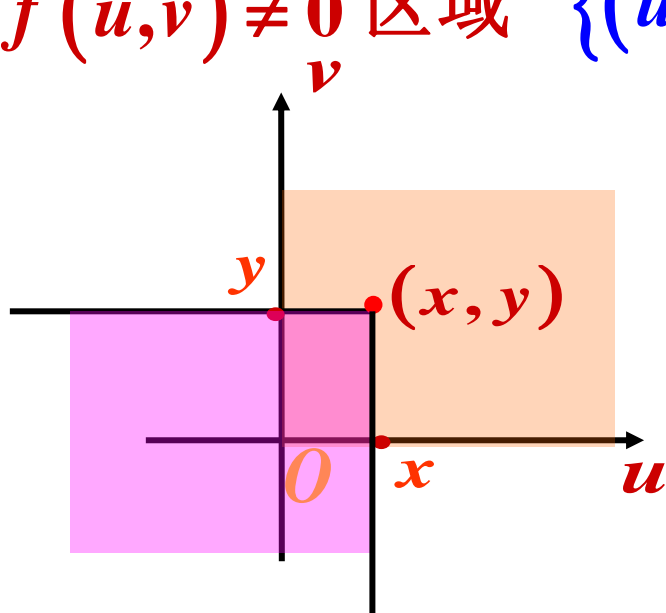
$$\int_0^{+\infty} \int_0^{+\infty} Ae^{-(2x+y)} dx dy = 1 \quad \longrightarrow \quad A=2$$

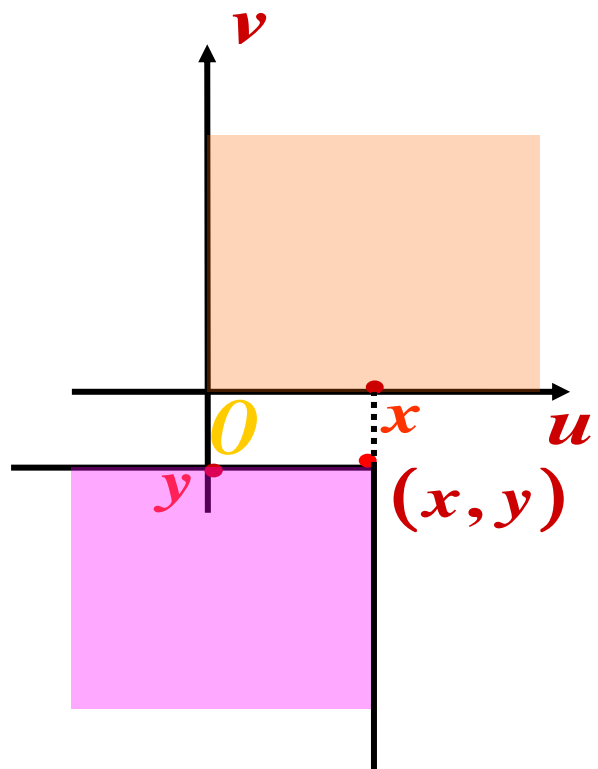
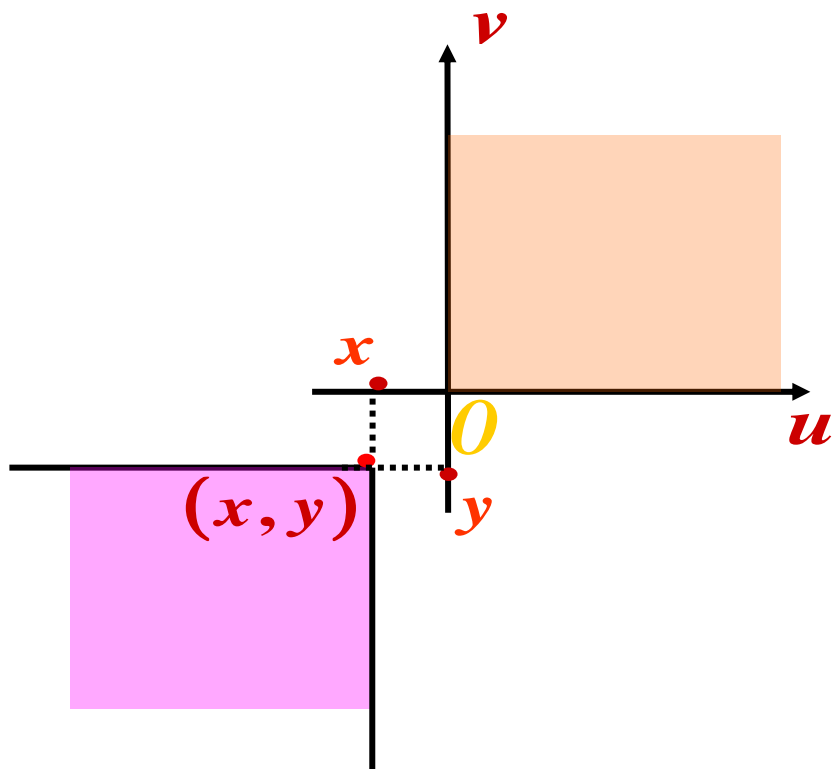
$$f(x, y) = \begin{cases} 2e^{-(2x+y)}, & x > 0, y > 0, \\ 0, & \text{其它.} \end{cases}$$

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

积分区域  $D = \{(u, v) | -\infty < u \leq x, -\infty < v \leq y\}$

$f(u, v) \neq 0$  区域  $\{(u, v) | u > 0, v > 0\}$





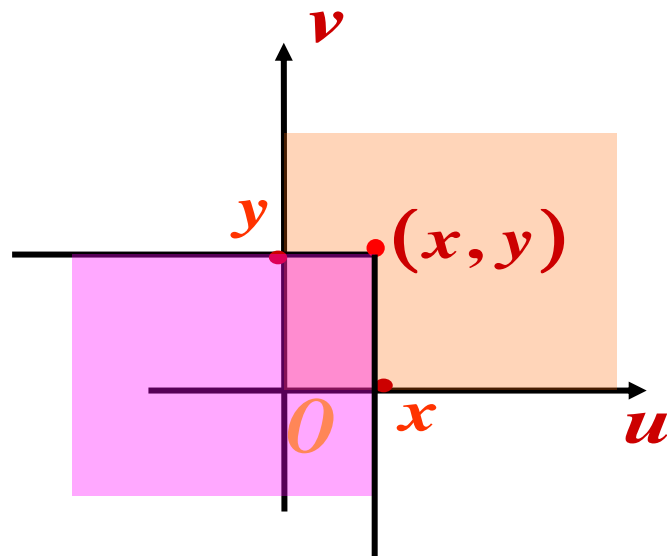
$$f(x, y) = \begin{cases} 2e^{-(2x+y)}, & x > 0, y > 0, \\ 0, & \text{其它.} \end{cases}$$

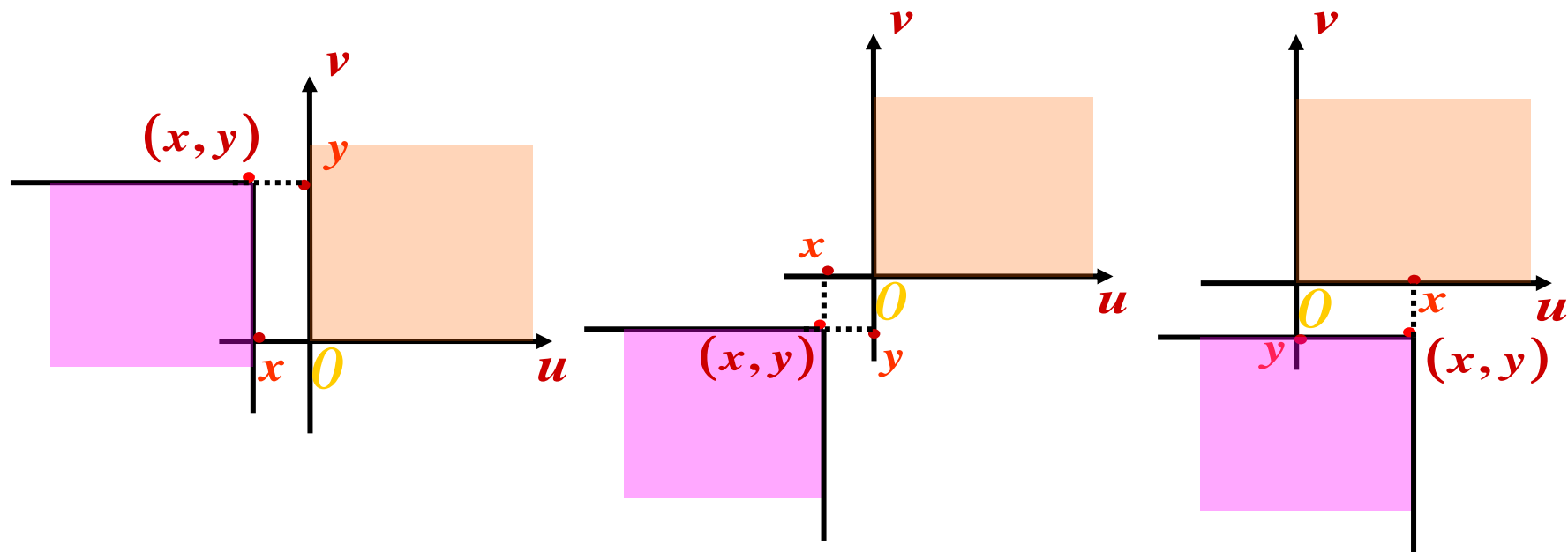
当  $x > 0, y > 0$  时,

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

$$= \int_0^x \int_0^y 2e^{-(2u+v)} du dv = 2 \int_0^y e^{-v} dv \cdot \int_0^x e^{-2u} du$$

$$= (1 - e^{-2x})(1 - e^{-y})$$



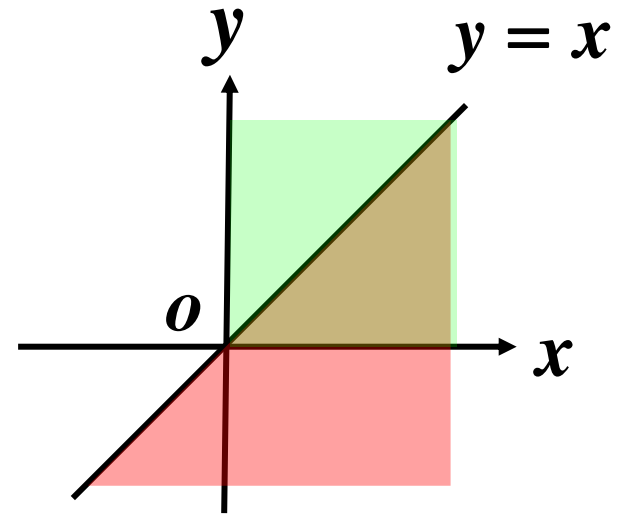


当  $x \leq 0$  或  $y \leq 0$  时,  $F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv = 0$

$$\text{故 } F(x, y) = \begin{cases} (1 - e^{-2x})(1 - e^{-y}), & x > 0, y > 0, \\ 0, & \text{其它.} \end{cases}$$

$$f(x, y) = \begin{cases} 2e^{-(2x+y)}, & x > 0, y > 0, \\ 0, & \text{其它.} \end{cases} \quad (3) \text{ 求 } P(Y \leq X)$$

$$\begin{aligned} P(Y \leq X) &= \iint_{y \leq x} f(x, y) dx dy \\ &= 2 \int_0^{+\infty} dx \int_0^x e^{-(2x+y)} dy \\ &= 2 \int_0^{+\infty} e^{-2x} dx \int_0^x e^{-y} dy \\ &= 2 \int_0^{+\infty} (e^{-2x} - e^{-3x}) dx = \frac{1}{3} \end{aligned}$$





## 例2：已知 $F(x,y)$ ，求其他

例：设随机变量 $(X,Y)$ 的联合分布函数为

$$F(x,y) = A \left( B + \arctan \frac{x}{2} \right) \left( C + \arctan \frac{y}{2} \right)$$

$$-\infty < x < +\infty, -\infty < y < +\infty$$

其中 $A, B, C$ 为常数，求


- (1) 确定常数  $A, B, C$ ；      (2)  $(X,Y)$ 的联合概率密度；  
(3)  $P(X > 2)$ 。

$$F(x, y) = A \left( B + \arctan \frac{x}{2} \right) \left( C + \arctan \frac{y}{2} \right)$$

**解:** (1)  $F(+\infty, +\infty) = A \left( B + \frac{\pi}{2} \right) \left( C + \frac{\pi}{2} \right) = 1$

$$F(-\infty, y) = A \left( B - \frac{\pi}{2} \right) \left( C + \arctan \frac{y}{2} \right) = 0$$

$$F(x, -\infty) = A \left( B + \arctan \frac{x}{2} \right) \left( C - \frac{\pi}{2} \right) = 0$$

  $B = \frac{\pi}{2}, C = \frac{\pi}{2}, A = \frac{1}{\pi^2}$

$$F(x, y) = \frac{1}{\pi^2} \left( \frac{\pi}{2} + \arctan \frac{x}{2} \right) \left( \frac{\pi}{2} + \arctan \frac{y}{2} \right)$$

$$(2) \quad f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} = \frac{1}{\pi^2} \frac{4}{(4 + x^2)(4 + y^2)}$$

$$-\infty < x < +\infty, -\infty < y < +\infty$$

$$(3) \quad P(X > 2)$$

常规方法  $P(X > 2) = \iint_{x>2} f(x, y) dx dy = \int_2^{+\infty} dx \int_{-\infty}^{+\infty} f(x, y) dy = \frac{1}{4}$

特殊方法：利用联合分布函数  $F(x, y)$

$$P(X > 2) = 1 - P(X \leq 2) = 1 - P(X \leq 2, Y < +\infty) = 1 - F(2, +\infty)$$

$$= 1 - \left( \frac{1}{2} + \frac{1}{\pi} \arctan \frac{2}{2} \right) = \frac{1}{4}$$

$(X, Y)$  作为一个整体, 有联合分布函数  $F(x, y)$ ;

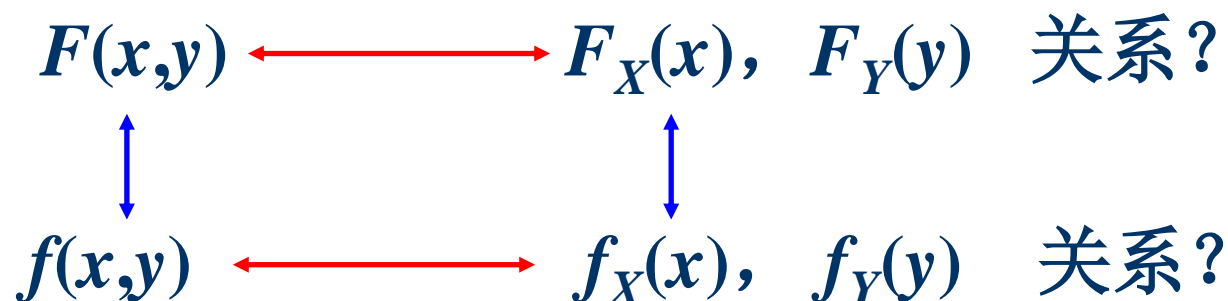
联合概率密度  $f(x, y)$ ;

$X$  作为一个部分, 有边缘分布函数  $F_X(x)$ ;

边缘概率密度  $f_X(x)$ ;

$Y$  作为一个部分, 有边缘分布函数  $F_Y(y)$ ;

边缘概率密度  $f_Y(y)$ ;



# 二维连续型随机变量——边缘密度函数

关于X的边缘分布函数为

$$F_X(x) = F(x, +\infty) = \int_{-\infty}^x \left[ \int_{-\infty}^{+\infty} f(u, v) dv \right] du$$

■关于X的边缘概率密度为

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

关于Y的边缘分布函数为

$$F_Y(y) = F(+\infty, y) = \int_{-\infty}^y \left[ \int_{-\infty}^{+\infty} f(u, v) du \right] dv$$

■关于Y的边缘概率密度为

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

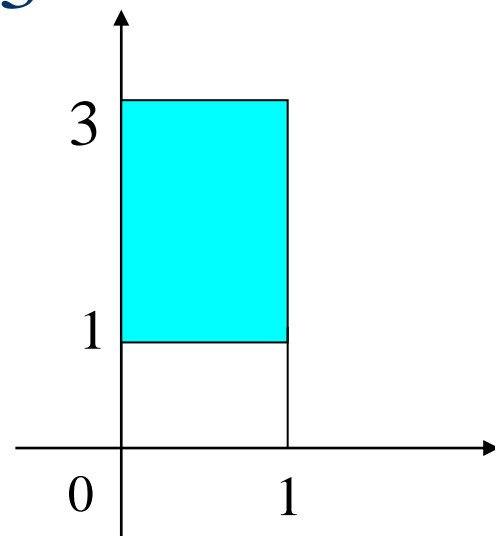
例1: 设(X, Y)的联合密度为

$$f(x, y) = \begin{cases} kxy & 0 \leq x \leq 1, 1 \leq y \leq 3 \\ 0 & \text{其它} \end{cases}$$

求  $k$  值和两个边缘概率密度

解: 由  $\int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} f(x, y) dy = 1$

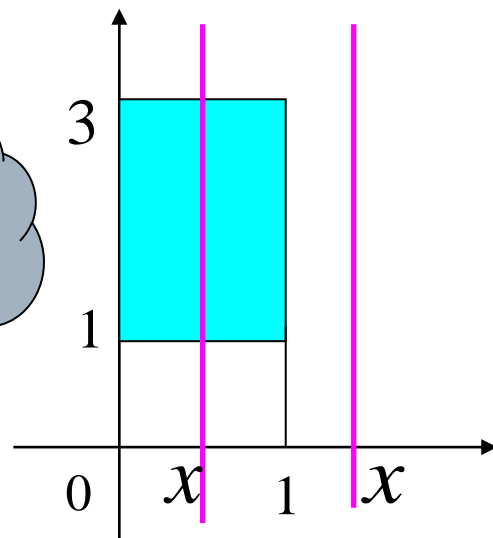
得  $k \int_1^3 y dy \cdot \int_0^1 x dx = 2k = 1 \rightarrow k = \frac{1}{2}$



$$f(x, y) = \begin{cases} \frac{1}{2}xy & 0 \leq x \leq 1, 1 \leq y \leq 3 \\ 0 & \text{其它} \end{cases}$$

解：关于X的边缘概率密度为

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



当  $0 \leq x \leq 1$  时  $f_X(x) = \int_1^3 \frac{1}{2}xy dy = 2x$

当  $x < 0$  或者  $x > 1$  时  $f_X(x) = \int_{-\infty}^{+\infty} 0 dy = 0$

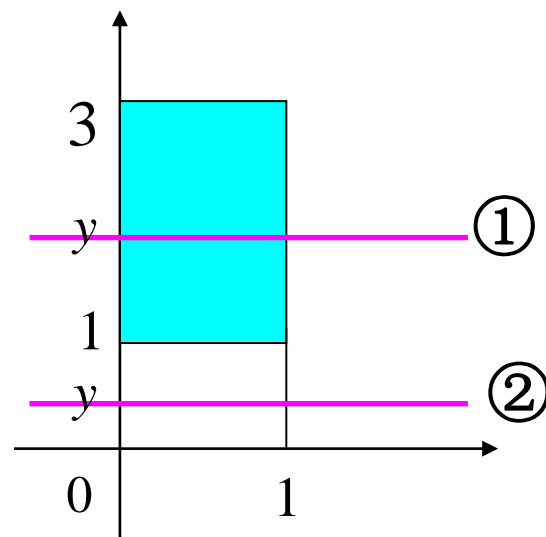
所以关于X的边缘概率密度为  $f_X(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{其它} \end{cases}$

$$f(x, y) = \begin{cases} \frac{1}{2}xy & 0 \leq x \leq 1, 1 \leq y \leq 3 \\ 0 & \text{其它} \end{cases}$$

关于Y的边缘概率密度为

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

$$= \begin{cases} \int_0^1 \frac{1}{2}xy dx & 1 \leq y \leq 3 \\ 0 & \text{其它} \end{cases} = \begin{cases} \frac{y}{4} & 1 \leq y \leq 3 \\ 0 & \text{其它} \end{cases}$$





**例2:** 设 $(X, Y)$ 的概率密度为

$$f(x, y) = \begin{cases} cy(2-x), & 0 \leq x \leq 1, 0 \leq y \leq x, \\ 0, & \text{其他.} \end{cases}$$

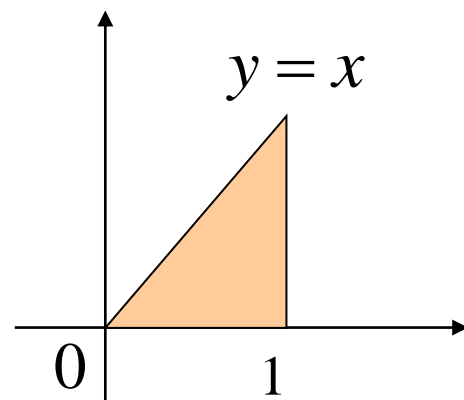
求 (1)  $c$  的值; (2) 边缘概率密度。

**解:** (1)  $1 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy$

$$= \int_0^1 \left[ \int_0^x cy(2-x) dy \right] dx$$

$$= c \int_0^1 [x^2(2-x)/2] dx = 5c/24$$

$$\Rightarrow c = 24/5$$



注意积分限

$$f(x, y) = \begin{cases} \frac{24}{5} y (2 - x), & 0 \leq x \leq 1, 0 \leq y \leq x, \\ 0, & \text{其他.} \end{cases}$$

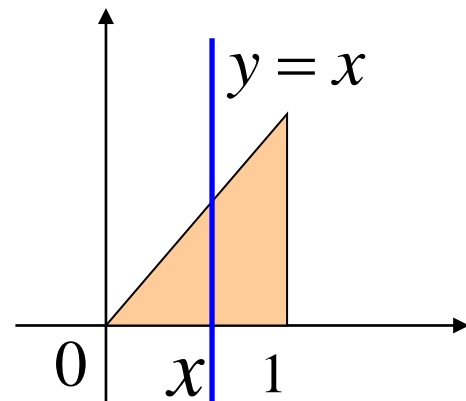
**解:** (2)  $f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy$

**注意积分限**  $= \int_0^x \frac{24}{5} y(2-x) dy$

$$= \frac{12}{5} x^2 (2-x), \quad 0 \leq x \leq 1$$

$$f_X(x) = \begin{cases} \frac{12}{5} x^2 (2-x), & 0 \leq x \leq 1, \\ 0, & \text{其他;} \end{cases}$$

**注意取值范围**



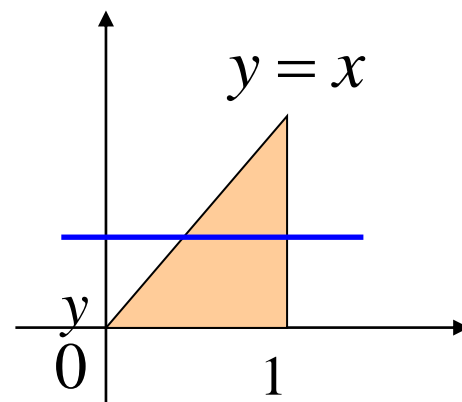
$$f(x, y) = \begin{cases} \frac{24}{5} y (2 - x), & 0 \leq x \leq 1, \quad 0 \leq y \leq x, \\ 0, & \text{其他.} \end{cases}$$

**解: (2)**  $f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx$

$$= \int_y^1 \frac{24}{5} y (2 - x) dx$$

$$= \frac{24}{5} y \left( \frac{3}{2} - 2y + \frac{y^2}{2} \right), \quad 0 \leq y \leq 1$$

$$f_Y(y) = \begin{cases} \frac{24}{5} y \left( \frac{3}{2} - 2y + \frac{y^2}{2} \right), & 0 \leq y \leq 1, \\ 0, & \text{其他.} \end{cases}$$



注意积分限

注意取值范围

# 连续型随机变量——独立性的判别

■ 若 $(X, Y)$ 是连续型随机变量,

$X$ 与 $Y$ 独立  $\longleftrightarrow f(x, y) = f_X(x) \cdot f_Y(y)$

定义

$\longleftrightarrow f(x, y) = g(x) \cdot h(y)$

形式  
上

**例3:** 甲乙两人约定中午12时30分在某地会面, 如果甲来到的时间在12:15到12:45之间是均匀分布; 乙独立地到达, 而且到达时间在12:00到13:00之间是均匀分布. 试求先到的人等待另一人到达的时间不超过5分钟的概率及甲先到的概率.

**解:** 设 $X$ 、 $Y$ 分别表示甲、乙到达时刻, 且 $X$ 与 $Y$ 相互独立。

以12时为起点, 以分为单位, 依题意,  $X \sim U(15, 45)$ ,  $Y \sim U(0, 60)$

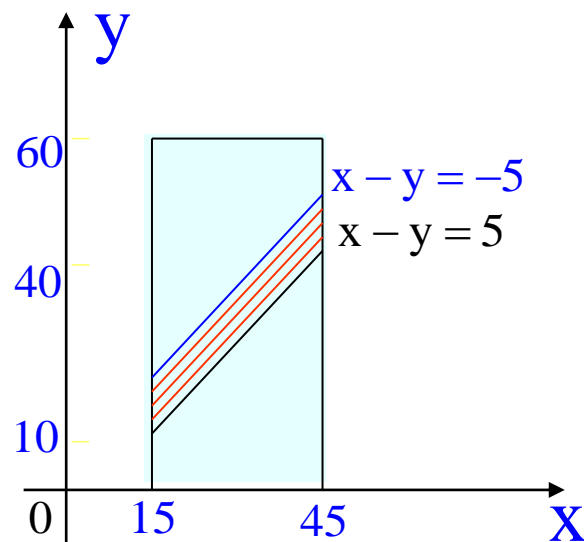
$$f_X(x) = \begin{cases} \frac{1}{30}, & 15 < x < 45 \\ 0, & \text{其它} \end{cases}, \quad f_Y(y) = \begin{cases} \frac{1}{60}, & 0 < y < 60 \\ 0, & \text{其它} \end{cases}$$

由独立性

$$f(x, y) = f_X(x) \cdot f_Y(y) = \begin{cases} \frac{1}{1800}, & 15 < x < 45, 0 < y < 60 \\ 0, & \text{其它} \end{cases}$$

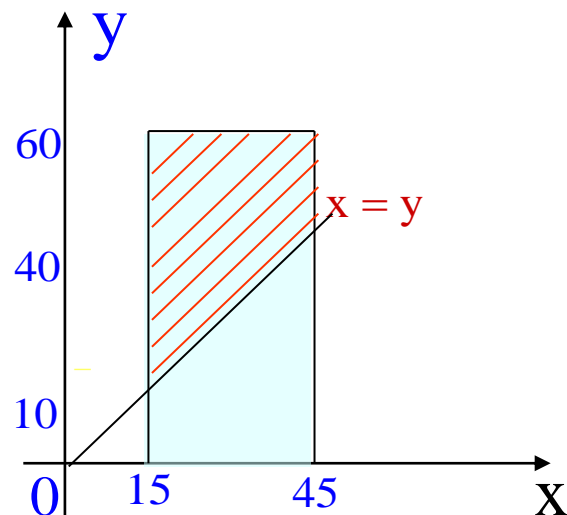
先到的人等待另一人到达的时间不超过5分钟的概率

$$\begin{aligned} P(|X-Y| \leq 5) &= P(-5 \leq X - Y \leq 5) \\ &= \iint_{-5 \leq x-y \leq 5} f(x, y) dx dy \\ &= \int_{15}^{45} \left[ \int_{x-5}^{x+5} \frac{1}{1800} dy \right] dx = 1/6 \end{aligned}$$



甲先到的概率

$$\begin{aligned} P(X < Y) &= \iint_{x < y} f(x, y) dx dy \\ &= \int_{15}^{45} \left[ \int_x^{60} \frac{1}{1800} dy \right] dx = 1/2 \end{aligned}$$



# 两个常用的二维连续型分布

◆ 二维均匀分布

◆ 二维正态分布

# 二维均匀分布

设 $G$ 是平面上的有界区域

二维随机变量

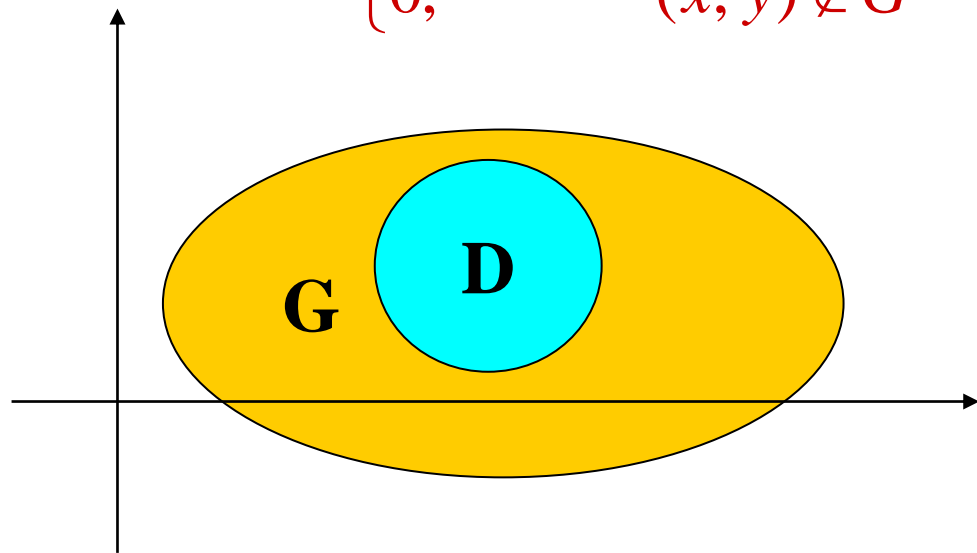
$(X, Y)$ 服从区域  
 $G$ 上的均匀分布

$(X, Y)$ 的联合概率密度为

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

对 $G$ 内任意区域 $D$ ，有

$$P((X, Y) \in D) = \frac{S_D}{S_G}$$







设 $(X, Y)$ 服从圆域  $x^2+y^2 \leq 4$  上的均匀分布,

1. 求 $(X, Y)$ 的联合概率密度;

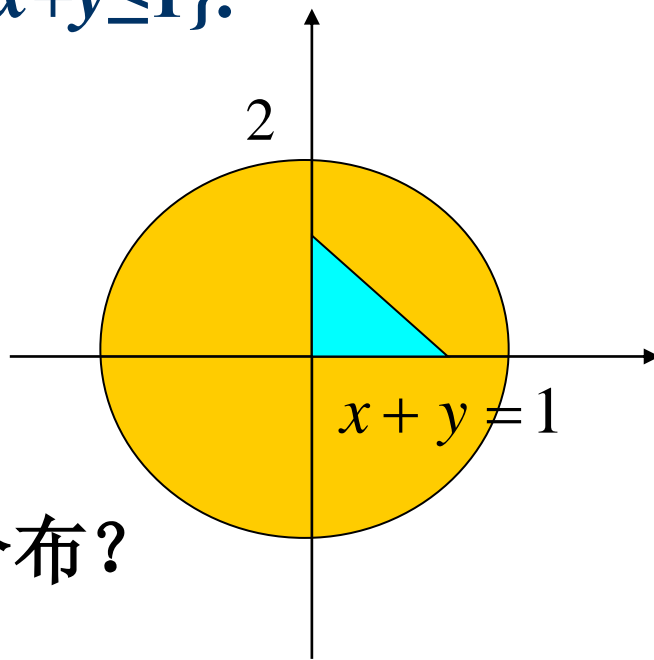
2. 计算 $P\{(X, Y) \in A\}$ , 其中

$$A = \{(x, y) \mid x, y \geq 0, \text{ 且 } x+y \leq 1\}.$$

答: 
$$f(x, y) = \begin{cases} \frac{1}{4\pi}, & x^2 + y^2 \leq 4 \\ 0, & x^2 + y^2 > 4 \end{cases}$$

$$P\{(X, Y) \in A\} = 0.5 / 4\pi = 1/8\pi$$

**X与Y的边缘密度? 是否为均匀分布?**



## 二维正态分布 $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

设二维随机变量(X,Y)的联合概率密度为

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

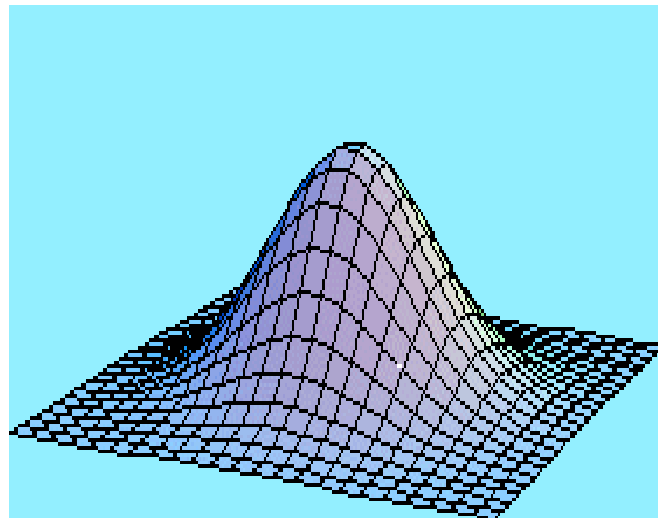
$(-\infty < x < +\infty, -\infty < y < +\infty)$

其中  $\mu_1, \mu_2, \sigma_1 > 0, \sigma_2 > 0, -1 < \rho < 1$  均为参数，则称 (X,Y)服从参数为  $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho$  的二维正态分布

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

二维正态分布  $(X, Y)$  的联合概率密度  $f(x, y)$  满足:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$$



**例：**二维正态分布  $(X, Y)$  的边缘概率密度满足：

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy,$$
$$= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}.$$

即  $X \sim N(\mu_1, \sigma_1^2)$

同理可得  $Y \sim N(\mu_2, \sigma_2^2)$

联合分布可以确定边缘分布，  
但边缘分布不能确定联合分布。<<

**结论：**

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

有  $X \sim N(\mu_1, \sigma_1^2),$

$Y \sim N(\mu_2, \sigma_2^2)$

不管  $\rho$  取何值

**例：** 设  $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ ，求证： $X$ 与 $Y$ 独立的充要条件为 $\rho = 0$ 。

**证明：**

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

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}},$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}.$$

$$X \text{与} Y \text{独立} \iff f(x, y) = f_X(x) \cdot f_Y(y) \iff \rho = 0$$

◆一个关于随机变量的独立性在统计中很有用的定理

**定理：** 设  $X_1, X_2, \dots, X_m$  与  $Y_1, Y_2, \dots, Y_n$  相互独立，

(1) 则  $X_{i_1}, X_{i_2}, \dots, X_{i_s}$  ( $i_1, i_2, \dots, i_s$  在  $1, 2, \dots, m$  中取不同的值) 与  $Y_{j_1}, Y_{j_2}, \dots, Y_{j_t}$  ( $j_1, j_2, \dots, j_t$  在  $1, 2, \dots, n$  中取不同的值) 相互独立.

(2) 如果  $g, h$  是连续函数，则  $g(X_1, X_2, \dots, X_m)$  与  $h(Y_1, Y_2, \dots, Y_n)$  相互独立.

## 说明

对于二维连续型r.v.来说,  $X$ 与 $Y$ 之间的关系的信息是包含在  $(X, Y)$  的联合概率密度函数之内的。

在第四章将指出: 对于二维正态分布而言, 参数  $\rho$  正好刻画了 $X$ 和 $Y$ 之间关系的密切程度。

因此, 仅由 $X$ 和 $Y$ 的边缘概率密度 (或边缘分布) 一般不能确定  $(X, Y)$  的联合概率密度函数 (或概率分布)。

# 小 结

二维离散型随机变量

二维连续型随机变量



## 二维离散型随机变量

$X \backslash Y$	$y_1$	$y_2$	$y_3$	$\cdots$	$P_X$
$x_1$	$p_{11}$	$p_{12}$	$p_{13}$	$\cdots$	$p_{1\cdot}$
$x_2$	$p_{21}$	$p_{22}$	$p_{23}$	$\cdots$	$p_{2\cdot}$
$x_3$	$p_{31}$	$p_{32}$	$p_{33}$	$\cdots$	$p_{3\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$P_Y$	$p_{\cdot 1}$	$p_{\cdot 2}$	$p_{\cdot 3}$	$\cdots$	<b>1</b>

## 二维连续型随机变量

$$(X,Y): F(x,y), f(x,y); \quad X: F_X(x), f_X(x);$$

$$Y: F_Y(y), f_Y(y);$$

$$F_X(x) = F(x, +\infty); \quad F_Y(y) = F(+\infty, y)$$

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

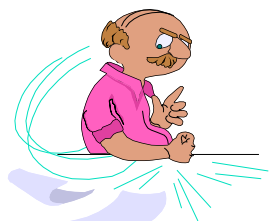
$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv; \quad f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}$$

$$F_X(x) = \int_{-\infty}^x f_X(t) dt; \quad f_X(x) = F'_X(x)$$

# 第六节

## 两个随机变量的函数的分布





在第二章中，我们讨论了一维  
随机变量函数的分布，现在我们进一步  
讨论：

当随机变量  $X, Y$  的联合分布已知时，如何  
求出它们的函数  $Z = g(X, Y)$  的分布？  
其中  $z = g(x, y)$  为连续函数。

# 两个随机变量的函数的分布

●  $Z=X+Y$  的分布

●  $Z_1=\max(X,Y)$  及  $Z_2=\min(X,Y)$  的分布

# $Z=X+Y$ 的分布

## ◆ 离散型随机变量和的分布

设 $(X,Y)$  是二维离散型随机变量, 其联合分布列为

$$P(X = a_i, Y = b_j) = p_{ij}, \quad (i = 1, 2, \dots; j = 1, 2, \dots)$$

则  $Z = g(X, Y)$  是一维的离散型随机变量, 其分布列为:

$$P(Z = g(a_i, b_j)) = p_{ij}, \quad (i = 1, 2, \dots; j = 1, 2, \dots)$$

步骤: (1) 确定 $Z=X+Y$  的取值;

(2) 确定 $Z=X+Y$  取相应值的概率。

例1：设 $(X,Y)$ 的联合分布列为

$X \backslash Y$	-1	0	1
0	0.1	0.2	0.3
1	0.2	0	0.2

分别求出 (1)  $Z=X+Y$ ; (2)  $W=X^2+Y-2$ 的分布列

$X \backslash Y$	-1	0	1
0	0.1 (-1)	0.2 (0)	0.3 (1)
1	0.2 (0)	0 (1)	0.2 (2)

则 $Z$ 的分布律:

$Z$	-1	0	1	2
$P$	0.1	0.4	0.3	0.2

$X \backslash Y$	-1	0	1
0	0.1	0.2	0.3
1	0.2	0	0.2

求  $W = X^2 + Y - 2$   
的分布列

$X \backslash Y$	-1	0	1
0	0.1 (-3)	0.2 (-2)	0.3 (-1)
1	0.2 (-2)	0 (-1)	0.2 (0)

则  $W$  的分布律:

$W$	-3	-2	-1	0
$P$	0.1	0.4	0.3	0.2



## 例2: 泊松分布的可加性

若  $X$  和  $Y$  相互独立, 它们分别服从参数为  $\lambda_1, \lambda_2$  的泊松分布, 证明  $Z=X+Y$  服从参数为  $\lambda_1 + \lambda_2$  的泊松分布。

解: 依题意

$$P(X = i) = \frac{\lambda_1^i}{i!} e^{-\lambda_1}, \quad P(Y = j) = \frac{\lambda_2^j}{j!} e^{-\lambda_2},$$

$$i = 0, 1, 2, \dots$$

$$j = 0, 1, 2, \dots$$

于是 
$$P(Z = r) = P(X + Y = r) = \sum_{i=0}^r P(X = i, Y = r - i)$$

**X与Y  
独立**

$$\bullet \quad \bullet \quad \bullet = \sum_{i=0}^r P(X = i) P(Y = r - i)$$

$$\begin{aligned}
 P(Z = r) &= \sum_{i=0}^r P(X = i) P(Y = r - i) \\
 &= \sum_{i=0}^r \frac{\lambda_1^i}{i!} e^{-\lambda_1} \cdot \frac{\lambda_2^{r-i}}{(r-i)!} e^{-\lambda_2} \\
 &= \frac{e^{-(\lambda_1 + \lambda_2)}}{r!} \sum_{i=0}^r \frac{r!}{i!(r-i)!} \lambda_1^i \lambda_2^{r-i} \\
 &= \frac{(\lambda_1 + \lambda_2)^r}{r!} e^{-(\lambda_1 + \lambda_2)}, \quad r = 0, 1, 2, \dots
 \end{aligned}$$

即Z服从参数为  $\lambda_1 + \lambda_2$  的泊松分布.

同理： 二项分布的可加性

设  $X \sim B(n_1, p)$ ,  $Y \sim B(n_2, p)$ , 且  $X$  和  $Y$  相互独立,

则  $Z = X + Y \sim B(n_1 + n_2, p)$ 。

# $Z=X+Y$ 的分布

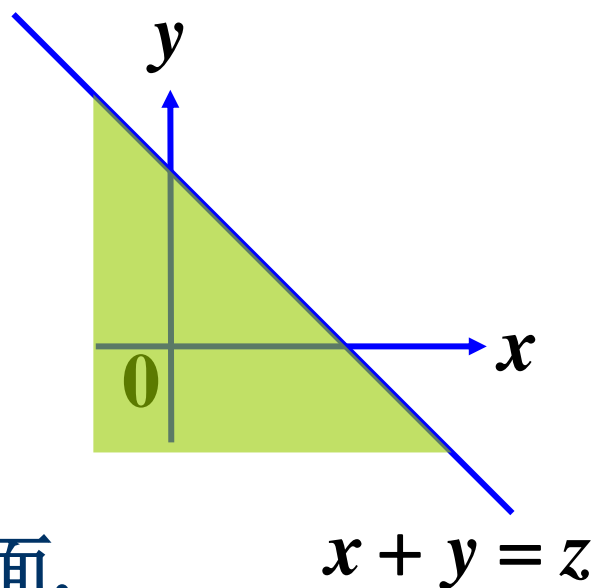
## ◆ 连续型随机变量和的分布

设 $(X, Y)$  是二维连续型随机变量, 其联合概率密度为 $f(x, y)$ , 则 $Z=X+Y$  的概率密度?

先求 $Z=X+Y$ 的分布函数:

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = P(X + Y \leq z) \\ &= \iint_{x+y \leq z} f(x, y) dx dy \end{aligned}$$

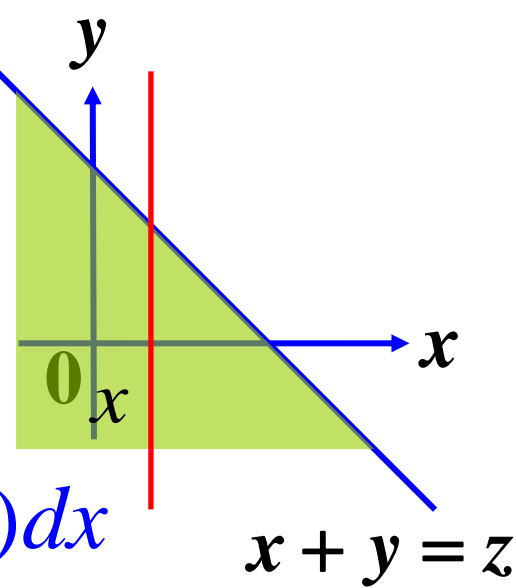
它是直线  $x+y=z$  及其左下方的半平面.



$$F_Z(z) = \iint_{x+y \leq z} f(x, y) dx dy$$

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

对  $z$  求导:  $f_Z(z) = F'_Z(z) = \int_{-\infty}^{+\infty} f(x, z-x) dx$



由  $X$  和  $Y$  的对称性,  $f_Z(z)$  又可写成

$$f_Z(z) = F'_Z(z) = \int_{-\infty}^{+\infty} f(z-y, y) dy$$

以上两式即是  $Z=X+Y$  的概率密度的一般公式.

# $Z=X+Y$ 的分布

## ◆ 连续型随机变量和的分布

设 $(X, Y)$  是二维连续型随机变量, 其联合概率密度为 $f(x, y)$ , 则 $Z=X+Y$  的概率密度?

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

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

特别地：当  $X$  和  $Y$  独立， $(X,Y)$  关于  $X, Y$  的边缘概率密度分别为  $f_X(x), f_Y(y)$ ，有： $f(x, y) = f_X(x)f_Y(y)$

则上述两式化为：

卷积公式

{

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

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

一般情况下

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

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

**X与Y独立时**

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

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



**例1:** 若  $X$  和  $Y$  独立, 具有相同的概率密度, 求:  $Z=X+Y$  的概率密度.

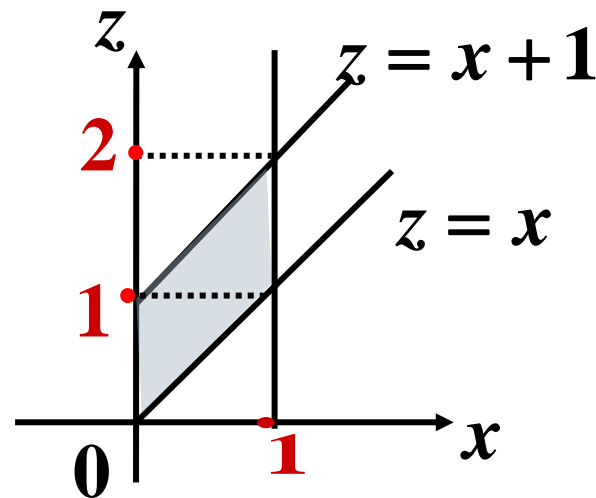
$$f_X(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0, & \text{其他} \end{cases}, f_Y(y) = \begin{cases} 1, & 0 \leq y \leq 1 \\ 0, & \text{其他} \end{cases}$$

**解:** 由卷积公式

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

为确定积分限, 先找出使被积函数不为 0 的区域

$$\begin{cases} 0 \leq x \leq 1 \\ 0 \leq z-x \leq 1 \end{cases} \quad \text{即} \quad \begin{cases} 0 \leq x \leq 1 \\ x \leq z \leq x+1 \end{cases}$$



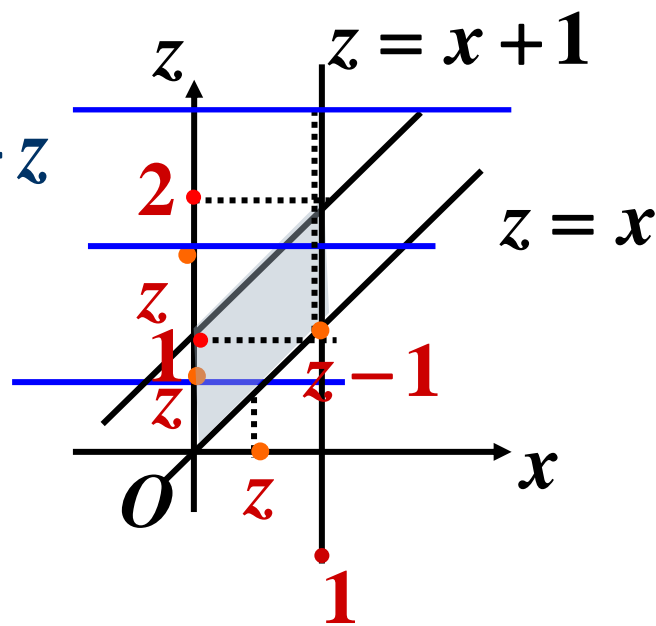
$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x) dx = \int_0^1 1 \cdot f_Y(z-x) dx$$

故当  $z < 0$  或  $z \geq 2$  时,  $f_Z(z) = 0$

当  $0 \leq z < 1$  时,  $f_Z(z) = \int_0^z dx = z$

当  $1 \leq z < 2$  时,  $f_Z(z) = \int_{z-1}^1 dx = 2 - z$

于是

$$f_Z(z) = \begin{cases} z, & 0 \leq z < 1, \\ 2 - z, & 1 \leq z < 2, \\ 0, & \text{其它.} \end{cases}$$


### P79页例3

若 $X$ 和 $Y$ 独立,  $X \sim N(0,1)$ ,  $Y \sim N(0,1)$ , 则:

$$Z=X+Y \sim N(0,2).$$

若 $X$ 和 $Y$ 独立,  $X \sim N(\mu_1, \sigma_1^2)$ ,  $Y \sim N(\mu_2, \sigma_2^2)$ ,  
结论又如何呢?

$$Z = X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$



正态分布的可加性

更一般地：

有限个独立正态变量的线性组合仍然服从正态分布. 即

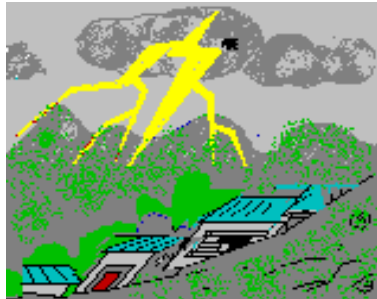
$X_i \sim N(\mu_i, \sigma_i^2), i = 1, 2, \dots, n$  且相互独立, 则:

$$\sum_{i=1}^n a_i X_i \sim N\left(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \sigma_i^2\right)$$

## $Z_1 = \max(X, Y)$ 及 $Z_2 = \min(X, Y)$ 的分布

设  $X, Y$  是两个相互独立的随机变量，分布函数分别为  $F_X(x)$  和  $F_Y(y)$ ，求  $Z_1 = \max(X, Y)$  及  $Z_2 = \min(X, Y)$  的分布函数。

桥梁或铸件所承受的最大应力、洪峰的高度、地震的震级等都用极值分布来描述。故研究最值分布有重要意义。



➤  $Z_1 = \max(X, Y)$  的分布函数

$$F_{Z_1}(z) = P(Z_1 \leq z)$$

$$= P(X \leq z, Y \leq z)$$

$$= P(X \leq z)P(Y \leq z)$$

$$= F_X(z)F_Y(z)$$

$Z_1 = \max(X, Y)$  的分布函数为:

$$F_{Z_1}(z) = F_X(z)F_Y(z)$$

$$Z_1 \leq z \Leftrightarrow \begin{cases} X \leq z \\ Y \leq z \end{cases}$$

$X$  和  $Y$  相互独立

➤  $Z_2 = \min(X, Y)$  的分布函数

$$F_{Z_2}(z) = P(Z_2 \leq z)$$

$$= 1 - P(Z_2 > z) = 1 - P(X > z, Y > z)$$

$$= 1 - P(X > z)P(Y > z)$$

$$= 1 - [1 - F_X(z)][1 - F_Y(z)]$$

$Z_2 = \min(X, Y)$  的分布函数为:

$$F_{Z_2}(z) = 1 - [1 - F_X(z)][1 - F_Y(z)]$$

$$Z_2 > z \Leftrightarrow \begin{cases} X > z \\ Y > z \end{cases}$$

$X$  和  $Y$  相互独立

# 一般结论

设  $X_1, \dots, X_n$  是  $n$  个相互独立的随机变量, 它们的分布函数分别为  $F_{X_i}(x)$ ,  $i = 1, 2, \dots, n$

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

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

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

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



特别地，当 $X_1, \dots, X_n$ 相互独立且具有相同分布函数 $F(x)$ 时，有

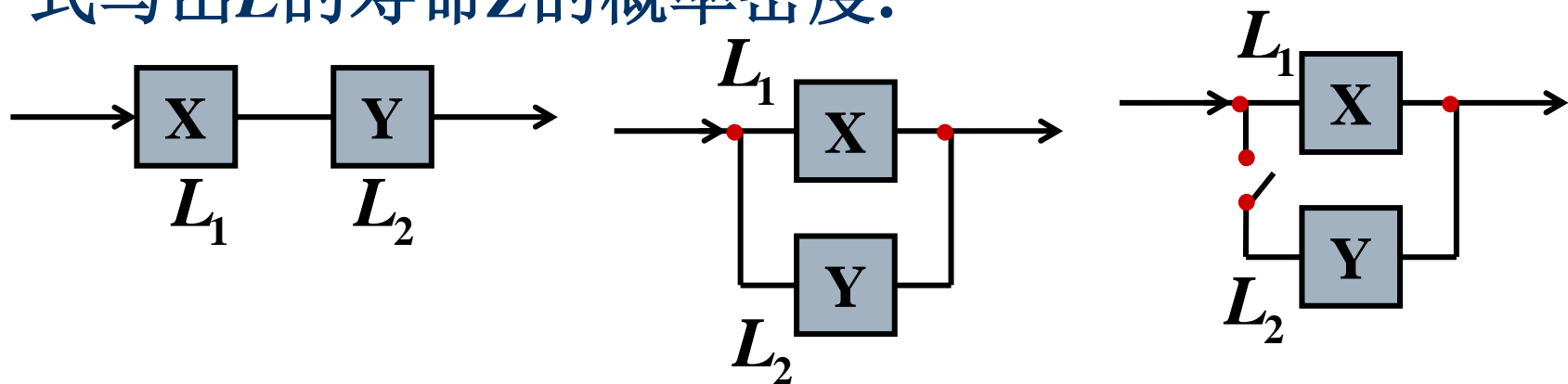
$$F_{\max}(y) = [F(y)]^n$$

$$F_{\min}(z) = 1 - [1 - F(z)]^n$$

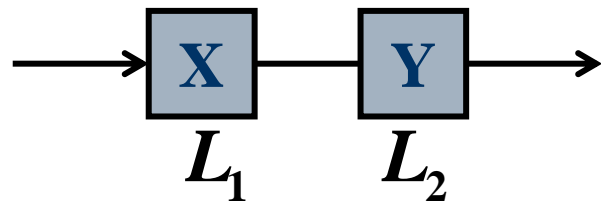
**例1:** 设系统  $L$  由两个相互独立的子系统  $L_1$  和  $L_2$  连接而成,连接的方式分别为 (i) 串联, (ii) 并联, (iii) 备用 (当系统  $L_1$  损坏时, 系统  $L_2$  开始工作), 如下图所示. 设  $L_1$  和  $L_2$  的寿命分别为  $X$  和  $Y$ , 已知它们的概率密度分别为

$$f_X(x) = \begin{cases} \alpha e^{-\alpha x}, & x > 0 \\ 0, & x \leq 0 \end{cases}, \quad f_Y(y) = \begin{cases} \beta e^{-\beta y}, & y > 0 \\ 0, & y \leq 0 \end{cases}$$

其中  $\alpha > 0, \beta > 0$  且  $\alpha \neq \beta$ . 试分别就以上三种连接方式写出  $L$  的寿命  $Z$  的概率密度.



**解：** (i) 串联的情况



由于当系统 $L_1$ 和 $L_2$ 中有一个损坏时,系统 $L$ 就停止工作,所以此时 $L$ 的寿命为  **$Z = \min(X, Y)$**

因为 $X$ 的概率密度和分布函数为:

$$f_X(x) = \begin{cases} \alpha e^{-\alpha x}, & x > 0 \\ 0, & x \leq 0 \end{cases}, F_X(x) = \begin{cases} 1 - e^{-\alpha x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

同理 $Y$ 的概率密度和分布函数为:

$$f_Y(y) = \begin{cases} \beta e^{-\beta y}, & y > 0 \\ 0, & y \leq 0 \end{cases}, F_Y(y) = \begin{cases} 1 - e^{-\beta y}, & y > 0 \\ 0, & y \leq 0 \end{cases}$$

$$F_X(x) = \begin{cases} 1 - e^{-\alpha x}, & x > 0 \\ 0, & x \leq 0 \end{cases}, F_Y(y) = \begin{cases} 1 - e^{-\beta y}, & y > 0 \\ 0, & y \leq 0 \end{cases}$$

于是  $Z = \min(X, Y)$  的分布函数为:

$$F_{\min}(z) = 1 - [1 - F_X(z)][1 - F_Y(z)] = \begin{cases} 1 - e^{-(\alpha + \beta)z}, & z > 0 \\ 0, & z \leq 0 \end{cases}$$

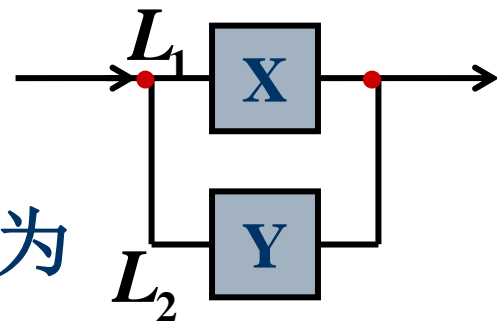
$Z = \min(X, Y)$  的概率密度为:

$$f_{\min}(z) = F'_{\min}(z) = \begin{cases} (\alpha + \beta)e^{-(\alpha + \beta)z}, & z > 0 \\ 0, & z \leq 0 \end{cases}$$

## (ii) 并联的情况

由于当且仅当系统 $L_1$ 和 $L_2$ 都损坏时，系统 $L$ 才停止工作，所以此时 $L$ 的寿命为

$$Z = \max(X, Y)$$



故  $Z = \max(X, Y)$  的分布函数为

$$F_{\max}(z) = F_X(z)F_Y(z) = \begin{cases} (1 - e^{-\alpha z})(1 - e^{-\beta z}), & z > 0, \\ 0, & z \leq 0, \end{cases}$$

于是  $Z = \max(X, Y)$  的概率密度为

$$f_{\max}(z) = F'_{\max}(z) = \begin{cases} \alpha e^{-\alpha z} + \beta e^{-\beta z} - (\alpha + \beta)e^{-(\alpha + \beta)z}, & z > 0 \\ 0, & z \leq 0 \end{cases}$$

$$\mathbf{Z} = \mathbf{X} + \mathbf{Y}$$

当  $z > 0$  时, 有

$$\begin{aligned} f_Z(z) &= \int_0^z \alpha e^{-\alpha x} \cdot \beta e^{-\beta(z-x)} dx \\ &= \frac{\alpha\beta}{\alpha - \beta} (e^{-\beta z} - e^{-\alpha z}). \end{aligned}$$

**例2:** 设相互独立的两个随机变量  $X, Y$  具有同一分布律, 且  $X$  的分布律为

$X$	0	1
$P$	0.5	0.5

试求:  $Z = \max(X, Y)$  的分布律.

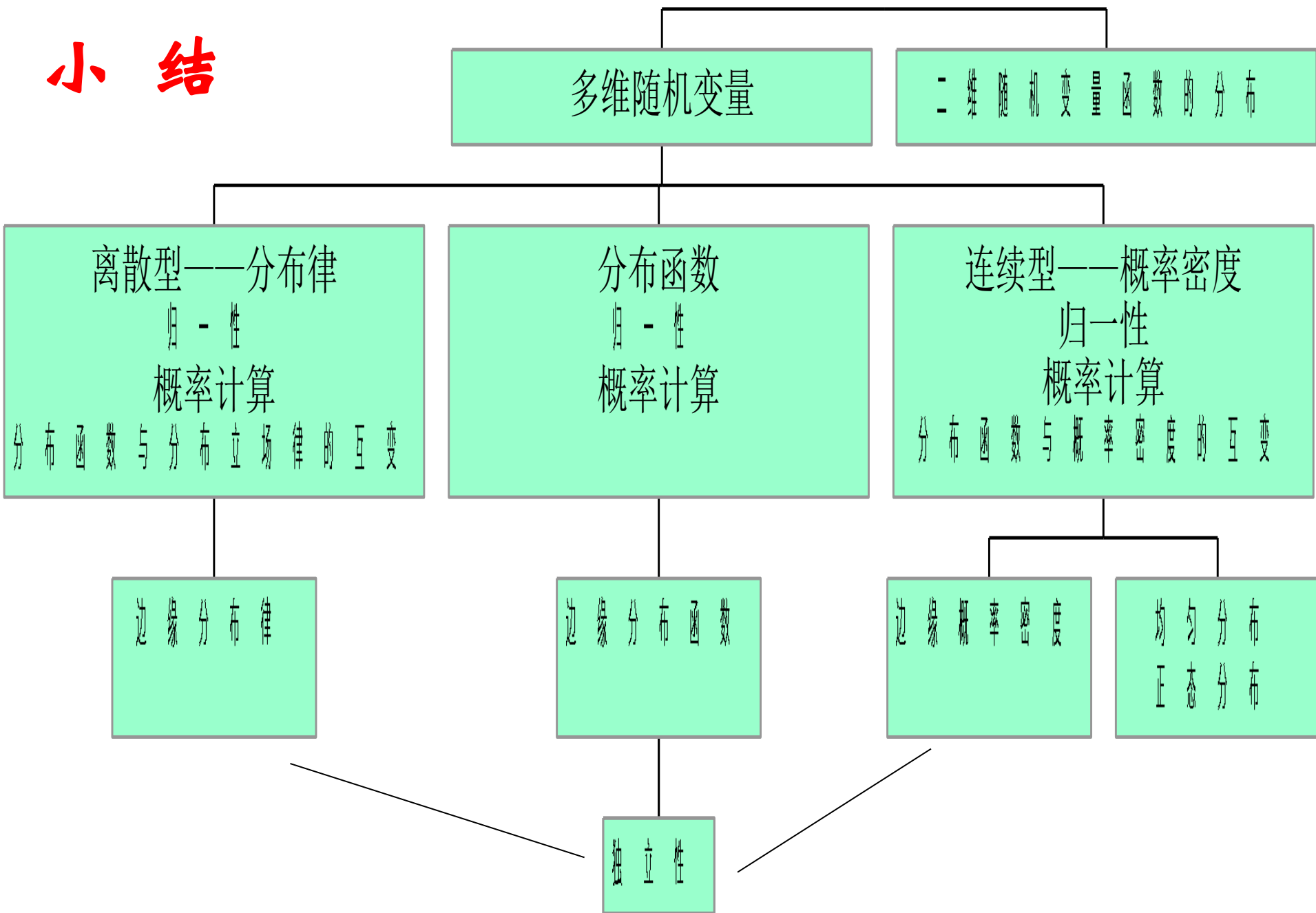
**解:**  $Z$  的取值为 0, 1, 且  $P(Z=0) = P(\max(X, Y)=0)$   
 $= P(X=0, Y=0)$   
 $= P(X=0)P(Y=0)$   
 $= 0.5 \cdot 0.5 = 0.25$

**X与Y独立**

所以: 分布律为

$Z$	0	1
$P$	0.25	0.75

# 小结





## 第四章 随机变量的数字特征

- ◆ 数学期望
- ◆ 方差
- ◆ 协方差与相关系数

前面讨论了随机变量及其分布。如果我们知道了随机变量  $X$  的概率分布，那么关于  $X$  的全部概率特征也就知道了。

然而，在实际问题中，概率分布是较难确定的。且有时在实际应用中，我们并不需要知道随机变量的所有性质，只要知道其一些数字特征就够了。

因此，在对随机变量的研究中，确定随机变量的某些数字特征是非常重要的。

最常用的数字特征是：

**数学期望、方差、协方差和相关系数**

# 数学期望的引例

例1 某高校调查了10位从本校毕业刚满一年的学生，得知他们月收入情况为

月收入 $X$ (元)	2000	2500	3000
人数	3	6	1

- (1) 求这10位毕业生的平均月收入；
- (2) 对于在校的马上就要毕业的大四学生来说，不考虑其它因素，只根据这组数据，他们对自己未来的月收入可以有怎样的预期？

解 将表格改动如下：

月收入 $X$ (元)	2000	2500	3000
比例	3/10	6/10	1/10

于是，毕业生的平均月收入（在校生的预期收入）为

$$2000 \times \frac{3}{10} + 2500 \times \frac{6}{10} + 3000 \times \frac{1}{10} = 2400$$

表格第二行的数据在这里有两种含义：

当用于求毕业生的平均收入时，表示某档收入出现的频率；

当它用于求在校生的预期收入，其含义则是在校生今后获得这一收入的概率，即对在校生来说，他今后月收入为2000、2500、3000元的概率分别为3/10, 6/10, 1/10. 因此，上述算式具备两种含义：

- (1)** 对已经发生的几种结果的加权平均（平均收入）；
- (2)** 对尚未发生的结果的一种预估（预期收入）。

# 数学期望 $E(X)$ Mathematical Expectation

## ◆ 一维离散型随机变量

定义：设离散型随机变量 $X$ 的概率分布为

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

若级数  $\sum_k x_k p_k$  绝对收敛，则称此级数为随机变量 $X$ 的数学期望(简称期望)或均值，记作 $E(X)$ ，即

$$E(X) = x_1 p_1 + x_2 p_2 + \dots + x_k p_k + \dots = \sum_k x_k p_k$$

若级数  $\sum_k x_k p_k$  不绝对收敛，则称 $X$ 的期望 $E(X)$ 不存在。

## 数学期望的计算

**例：**一批产品中有一、二、三等及废品4种，相应比例分别为60%，20%，13%，7%，若各等级的产值分别为10元，5.8元，4元及0元，求这批产品的平均产值。

**解：**设一个产品的产值为 $X$ 元，则 $X$ 的分布律为：

<b>X</b>	0	4	5.8	10
<b>P</b>	0.07	0.13	0.2	0.6

$$E(X) = x_1 p_1 + x_2 p_2 + x_3 p_3 + x_4 p_4$$

$$= 0 + 4 \times 0.13 + 5.8 \times 0.2 + 10 \times 0.6 = 7.68$$

**应用：**谁的技术比较好？

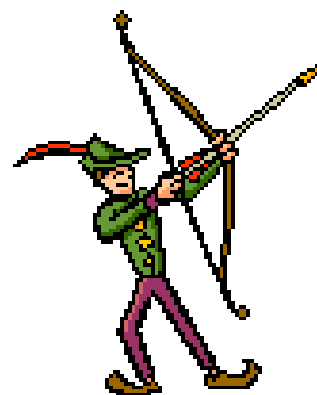
甲乙两个射手，他们射击的分布律分别为

甲射手

击中环数	8	9	10
概率	0.3	0.1	0.6

乙射手

击中环数	8	9	10
概率	0.2	0.5	0.3



试问哪个射手技术较好？

甲射手的平均击中环数为9.3环；

乙射手的平均击中环数为9.1环；

**甲射手的技术比较好。**

# 常见分布的数学期望

◆ 0-1分布: 

<b>X</b>	0	1
<b>P</b>	$1-p$	$p$

 $E(X) = 1 \times p + 0 \times (1-p) = p$

◆ 二项分布: 若  $X \sim B(n, p)$ , 则

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

$$E(X) = np$$

◆ 泊松分布: 若  $X \sim \pi(\lambda)$ , 则

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

$$E(X) = \lambda$$



# 随机变量函数的数学期望

**定理：** 设 $Y=g(X)$ 是随机变量 $X$ 的函数，若 $X$ 为离散型随机变量，其分布律为  $P(X = x_k) = p_k, \quad k = 1, 2, \dots$  则 $Y$ 的数学期望为

$$E(Y) = E[g(X)] = \sum_k g(x_k) p_k$$

$$E(X) = x_1 p_1 + x_2 p_2 + \dots + x_k p_k + \dots = \sum_k x_k p_k$$

该公式的重要性在于：当我们求  $E[g(X)]$  时，不必求  $g(X)$  的分布，而只需知道  $X$  的分布足矣。这对求  $g(X)$  的期望带来了极大方便。

# 数学期望的计算

例：设随机变量 $X$ 的分布律为：

$X$	-1	0	1	2
$P$	0.1	0.2	0.3	0.4

$$E(X) = -1 \cdot 0.1 + 0 \cdot 0.2 + 1 \cdot 0.3 + 2 \cdot 0.4 = 1$$

$$E(X^2 + 1) = 2 \cdot 0.1 + 1 \cdot 0.2 + 2 \cdot 0.3 + 5 \cdot 0.4 = 3$$

$$E\left(\frac{1}{X+2}\right) = 1 \cdot 0.1 + \frac{1}{2} \cdot 0.2 + \frac{1}{3} \cdot 0.3 + \frac{1}{4} \cdot 0.4 = 0.4$$

# 数学期望

## ◆ 二维离散型随机变量

**定理：** 设 $(X,Y)$ 是二维离散型随机变量，其分布律为：

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

则  $Z=g(X,Y)$ 的数学期望为：

$$E(Z) = E[g(X, Y)] = \sum_{i,j} g(x_i, y_j) p_{ij}$$

**例1:** 设二维离散型随机变量  $(X, Y)$  的概率分布为:

X \ Y	1	2
	1	2
1	$1/8$	$1/4$
2	$1/2$	$1/8$

$$E(XY) = (1 \cdot 1) \cdot \frac{1}{8} + (1 \cdot 2) \cdot \frac{1}{4} + (2 \cdot 1) \cdot \frac{1}{2} + (2 \cdot 2) \cdot \frac{1}{8} = \frac{17}{8}$$

$$E\left(\frac{1}{X}\right) = 1 \cdot \frac{3}{8} + \frac{1}{2} \cdot \frac{5}{8} = \frac{11}{16}$$

利用一维X的边缘分布。

# 数学期望

## ◆ 一维连续型随机变量

**定义：** 设连续型随机变量 $X$ 的概率密度为 $f(x)$ , 若积

分  $\int_{-\infty}^{+\infty} |x| f(x) dx$  有限, 则称

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx \quad \text{为} X \text{的数学期望。}$$

如果积分  $\int_{-\infty}^{+\infty} |x| f(x) dx$  发散, 则称 $X$ 的数学期

望不存在。

**即：** 连续型随机变量的数学期望是一个绝对收敛的积分值。

# 数学期望的计算

**例：**已知随机变量X的密度函数为

$$f(x) = \begin{cases} \frac{1}{\pi\sqrt{1-x^2}} & |x| < 1 \\ 0 & |x| \geq 1 \end{cases} \quad \text{求X的数学期望。}$$

**解：**  $E(X) = \int_{-\infty}^{+\infty} xf(x)dx$

$$= \int_{-\infty}^{-1} x \cdot 0 dx + \int_{-1}^1 x \frac{1}{\pi\sqrt{1-x^2}} dx + \int_1^{+\infty} x \cdot 0 dx = 0$$

定积分的对称性

# 常见分布的数学期望

◆ 均匀分布：若  $X \sim U[a, b]$ ，则

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{其它} \end{cases} \quad E(X) = \frac{a+b}{2}$$

◆ 指数分布：若  $X \sim E(\lambda)$ ，则

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (\lambda > 0 \text{ 为常数}) \quad E(X) = \frac{1}{\lambda}$$

◆ 正态分布：若  $X \sim N(\mu, \sigma^2)$ ，则

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

◆ 柯西分布：（P89例4）  $E(X)$  不存在

# 随机变量函数的数学期望

**定理：** 设 $Y=g(X)$ 是随机变量 $X$ 的函数，若 $X$ 为连续型随机变量，其密度函数为 $f(x)$ ，则 $Y$ 的数学期望为：

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

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



**例：**已知  $X \sim U(0, 2\pi)$ ，求  $Y = \sin X$  的数学期望。

**解：** $X$  的概率密度为  $f(x) = \begin{cases} \frac{1}{2\pi}, & 0 \leq x \leq 2\pi; \\ 0, & \text{其它。} \end{cases}$

$$\begin{aligned} E(Y) &= E(\sin X) = \int_{-\infty}^{+\infty} \sin x f(x) dx \\ &= \int_0^{2\pi} \frac{1}{2\pi} \sin x dx = 0 \end{aligned}$$

# 数学期望

## ◆ 二维连续型随机变量

**定理：** 设  $Z=g(X,Y)$ ，若  $(X,Y)$  为二维连续型随机变量，其密度函数为  $f(x,y)$ ，则  $Z$  的数学期望为：

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

**例：** 设随机变量 $X$ 和 $Y$ 相互独立， 概率密度分别为

$$f_X(x) = \begin{cases} 4e^{-4x}, & x > 0 \\ 0, & \text{其它} \end{cases}, \quad f_Y(y) = \begin{cases} 2e^{-2y}, & y > 0 \\ 0, & \text{其它} \end{cases} \quad \text{求} E(XY).$$

**解：**  $X$  和  $Y$  相互独立， 故：  $f(x, y) = f_X(x) \cdot f_Y(y)$

$$\begin{aligned} E(XY) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xyf(x, y)dxdy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xyf_X(x)f_Y(y)dxdy \\ &= \left( \int_{-\infty}^{+\infty} xf_X(x)dx \right) \cdot \left( \int_{-\infty}^{+\infty} yf_Y(y)dy \right) = \left( \int_0^{+\infty} 4xe^{-4x}dx \right) \cdot \left( \int_0^{+\infty} 2ye^{-2y}dy \right) \\ &= \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8} \end{aligned}$$

◆ **X**:一维离散型

$$E(X) = \sum_k x_k p_k$$

$$E[g(X)] = \sum_k g(x_k) p_k$$

◆ **X**:一维连续型

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

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

◆ **(X,Y)**:二维离散型

$$E[g(X,Y)]$$

$$= \sum_{i,j} g(x_i, y_j) p_{ij}$$

◆ **(X,Y)**:二维连续型

$$E[g(X,Y)]$$

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

# 数学期望的性质

◆ 设 $C$ 是常数, 则  $E(C)=C$ ;

◆ 设 $C$ 是常数, 则  $E(CX)=CE(X)$ ;

◆  $E(X+Y) = E(X)+E(Y)$ ;

推广:  $E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i)$ ;

◆ 设  $X, Y$  相互独立, 则  $E(XY)=E(X) \cdot E(Y)$ ;

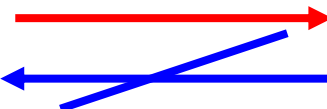
推广:  $E\left(\prod_{i=1}^n X_i\right) = \prod_{i=1}^n E(X_i)$

请注意:

$E(XY)=E(X)E(Y)$   
不一定能推出 $X, Y$   
独立

$X_i$ 相互独立

## ① P93的例4

$X, Y$  相互独立   $E(XY) = E(X) \cdot E(Y)$

## ② P93的例5

本题是将 $X$ 分解成数个随机变量之和  $X = \sum_{i=1}^n X_i$

然后利用数学期望的性质  $E(X) = E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i)$

此方法具有一定的意义。

# 数学期望在医学上的一个应用

考虑用验血的方法在人群中普查某种疾病。集体做法是每10个人一组，把这10个人的血液样本混合起来进行化验。如果结果为阴性，则10个人只需化验1次；若结果为阳性，则需对10个人在逐个化验，总计化验11次。假定人群中这种病的患病率是10%，且每人患病与否是相互独立的。试问：这种分组化验的方法与通常的逐一化验方法相比，是否能减少化验次数？

**分析：** 设随机抽取的10人组所需的化验次数为 $X$

我们需要计算 $X$ 的数学期望，然后与10比较

先求出化验次数 $X$ 的分布律。

化验次数 $X$ 的可能取值为1, 11

注意求  $X$  期望值的步骤!

$(X=1)$  = “10人都是阴性”

$$P(X=1) = (1-0.1)^{10} = 0.9^{10}$$

$(X=11)$  = “至少1人阳性”

$$P(X=11) = 1 - 0.9^{10}$$

$$E(X) = 0.9^{10} \times 1 + (1 - 0.9^{10}) \times 11 = 7.513 < 10$$

**结论：** 分组化验法的次数少于逐一化验法的次数



## 问题的进一步讨论

### 1. 概率 $p$ 对是否分组的影响

若 $p=0.2$ ，则

$$E(X) = 0.8^{10} \times 1 + (1 - 0.8^{10}) \times 11 = 9.9262$$

当 $p > 0.2057$ 时， $E(X) > 10$

### 2. 概率 $p$ 对每组人数 $n$ 的影响

当 $p=0.1$ 时，为使

$$E(X) = 0.9^n \times 1 + (1 - 0.9^n) \times 11 < 10$$

$$\longrightarrow n < 21.86$$

当 $p=0.2$ 时，可得出 $n < 10.32$ ，才能保证 $EX < 10$ .

## 小 结

本讲介绍了随机变量数学期望的概念、性质及计算，给出了几种常用随机变量的数学期望，介绍了求随机变量函数数学期望的方法。

◆ **X**:一维离散型

$$E(X) = \sum_k x_k p_k$$
$$E[g(X)] = \sum_k g(x_k) p_k$$

◆ **X**:一维连续型

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

◆ **(X,Y)**:二维离散型

$$E[g(X,Y)]$$
$$= \sum_{i,j} g(x_i, y_j) p_{ij}$$

◆ **(X,Y)**:二维连续型

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

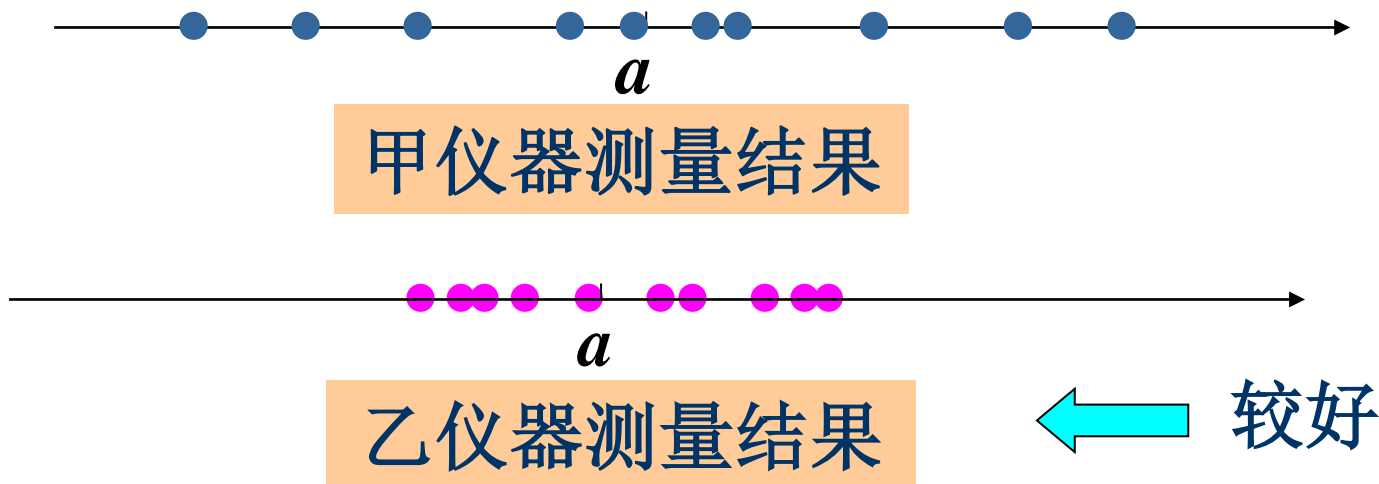
# 第三节

第三节

方差

# 方差的引例

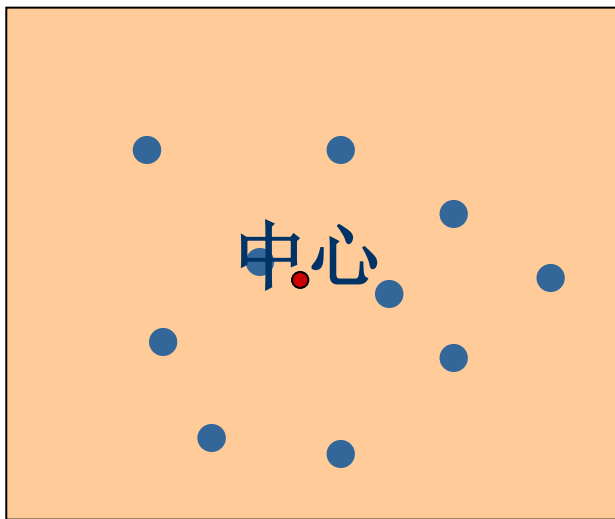
例如，某零件的真实长度为 $a$ ，现用甲、乙两台仪器各测量10次，将测量结果 $X$ 用坐标上的点表示如图：



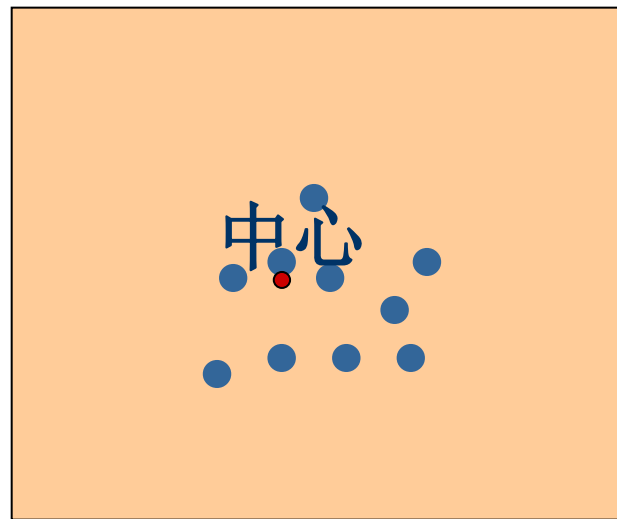
因为乙仪器的测量结果集中在均值附近



又如，甲、乙两门大炮同时向一目标射击10次，其落点距目标的位置如图：



甲炮射击结果



乙炮射击结果

↑ 乙较好

因为乙炮的弹着点较集中在中心附近。

# 方差 Variance

◆ **定义：** 设 $X$ 是一随机变量，若 $E[X-E(X)]^2$  存在，则称其为 $X$ 的方差，记成  $D(X)$ 或者  $\text{var}(X)$  ， 即

$$D(X) = \text{var}(X) = E[X - E(X)]^2$$

◆ **均方差（标准差）**

$$\sigma(X) = \sqrt{D(X)}$$

均方差与 $X$ 有相同的量纲



非负实数

方差刻画了随机变量的取值对于其数学期望的偏离程度，即体现随机变量取值离散程度或波动程度。

- 方差 $D(X) \geq 0$ ，且 $D(X)=0 \iff X$ 以概率1取常数。
- 方差越小，说明 $X$ 的取值越集中， $E(X)$ 的代表性就越好；
- 方差越大，说明 $X$ 的取值比较分散， $E(X)$ 的代表性就差。



由定义 $D(X) = E[X - E(X)]^2$ 知,

方差是随机变量 $X$ 的函数  $g(X) = [X - E(X)]^2$  的数学期望

$$D(X) = \begin{cases} \sum_k [x_k - E(X)]^2 p_k, \\ \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x) dx, \end{cases}$$

$X$ 为离散型,  
 $P(X=x_k)=p_k$

$X$ 为连续型,  
 $f(x)$ 为密度

# 方差的计算公式

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

公式变形：

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

**Proof.**

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

# 方差的计算公式

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

计算 $D(X)$ ：只需计算两类期望 $E(X)$ ,  $E(X^2)$

步骤：

Step 1: 计算期望  $E(X)$

Step 2: 计算期望  $E(X^2)$

Step 3: 套用公式：  $D(X) = E(X^2) - [E(X)]^2$

**例：** 设连续型随机变量 $X$  的密度函数为：

$$f(x) = \begin{cases} 2x, & x \in [0, 1] \\ 0, & x \notin [0, 1] \end{cases} \quad \text{求 } D(X)$$

**解：** Step 1: 计算期望  $E(X)$

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_0^1 x \cdot 2xdx = \frac{2}{3}$$

Step 2: 计算期望  $E(X^2)$

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)dx = \int_0^1 x^2 \cdot 2xdx = \frac{1}{2}$$

Step 3: 套用公式

$$D(X) = E(X^2) - [E(X)]^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}$$

# 常见分布的数学期望与方差

◆ 0-1分布:

X	0	1
P	$q$	$p$

其中 $p+q=1$

$$E(X)=p, \quad D(X)=pq$$

◆ 二项分布: 若 $X \sim B(n, p)$ , 则

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

$$E(X) = np, \quad D(X) = npq$$

◆ 泊松分布: 若 $X \sim \pi(\lambda)$ , 则

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

$$E(X) = D(X) = \lambda$$

# 常见分布的数学期望与方差

◆ 均匀分布：若  $X \sim U[a, b]$ ，则

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

$$E(X) = \frac{a+b}{2}, \quad D(X) = \frac{(b-a)^2}{12}$$

◆ 指数分布：若  $X \sim E(\lambda)$ ，则

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

$$E(X) = \frac{1}{\lambda}, \quad D(X) = \frac{1}{\lambda^2}$$

# 常见分布的数学期望与方差

◆ 正态分布：若  $X \sim N(\mu, \sigma^2)$ ，则

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

$$E(X) = \mu, \quad D(X) = \sigma^2$$

# 方差的性质

- ◆ 设 $C$ 是常数, 则  $D(C)=0$ ;
- ◆ 设 $C$ 是常数, 则  $D(CX)=C^2D(X)$ ;
- ◆ 若 $X$ 与 $Y$  独立, 则

独立是条件

$$D(X+Y)=D(X)+D(Y); \quad D(X-Y)=D(X)+D(Y);$$

$$D(aX+bY)=a^2D(X)+b^2D(Y); \quad D(aX+b)=a^2D(X)$$

期望性质对比:  
可推广为: 若 $X_1, X_2, \dots, X_n$ 相互独立, 则

$$E(C)=C, \quad E(CX)=CE(X), \quad E\left(\sum_{i=1}^n C_i X_i\right)=\sum_{i=1}^n C_i E(X_i);$$

$$E(X+Y)=E(X)+E(Y);$$

$$D\left(\sum_{i=1}^n C_i X_i\right)=\sum_{i=1}^n C_i^2 D(X_i);$$

若 $X$ 与 $Y$  独立, 则  $E(XY)=E(X) \cdot E(Y)$



**例：** 设随机变量 $X$ 的期望和方差分别为 $E(X)$ 和 $D(X)$ ,

且 $D(X) > 0$ , 求  $Y = \frac{X - E(X)}{\sqrt{D(X)}}$  的期望与方差.

$$E(Y) = 0, D(Y) = 1$$

称  $Y = \frac{X - E(X)}{\sqrt{D(X)}}$  为  $X$  的标准化随机变量.

**P99的例4, 例5**

## 小 结

本节介绍了随机变量方差的概念、性质及计算，给出了几种常用随机变量的方差。

# 第四节

## 协方差与相关系数

对于一维随机变量 $X$ 来说，

数学期望刻画了其取值的平均水平；

方差刻画了取值对于其数学期望的偏离程度。

对于二维随机变量 $(X, Y)$ 来说，

其分量 $X$ 和 $Y$ 有期望与方差之外，还应有一些数字特征，用以刻画 $X$ 与 $Y$ 之间的相关程度，其中最主要的就是下面要讨论的协方差和相关系数。

若 $X$ 与 $Y$ 独立, 则  $D(X+Y)=D(X)+D(Y)$

若 $X$ 与 $Y$ 不独立, 则  $D(X+Y)= ?$

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

从而,  $E\left[\left(X-E(X)\right)\left(Y-E(Y)\right)\right] \neq 0$  在一定程度上反映了二维随机变量 $(X,Y)$ 中的分量 $X$ 与 $Y$ 的某种相互关系。

# 协方差 Covariance

◆**定义**：若随机变量X的期望 $E(X)$ 和Y的期望 $E(Y)$ 都存在，则称 $\text{Cov}(X, Y) = E\{[X - E(X)][Y - E(Y)]\}$ 为X与Y的协方差。

◆**计算式**：
$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

当X和Y相互独立时， $\text{Cov}(X, Y) = 0$ ；  
所以计算协方差，归根结底还是计算数学期望  
计算期望是重点（二维随机变量如何计算期望）  
$$= E(XY) - E(X)E(Y)$$

# 协方差的性质

(1)  $\text{Cov}(X, X) = D(X)$

(2)  $\text{Cov}(X, Y) = \text{Cov}(Y, X);$

(3)  $\text{Cov}(aX, bY) = ab \text{Cov}(X, Y)$ ,  $a, b$ , 是常数

(4)  $\text{Cov}(X_1+X_2, Y) = \text{Cov}(X_1, Y) + \text{Cov}(X_2, Y)$

补充:  $\text{Cov}(X, a) = 0$

$$\text{Cov}(aX+b, cY+d) = ac \text{Cov}(X, Y);$$

$$\begin{aligned} \text{Cov}(aX_1+bY_1, cX_2+dY_2) = & ac\text{Cov}(X_1, X_2) + \\ & bc\text{Cov}(Y_1, X_2) + ad\text{Cov}(X_1, Y_2) + bd\text{Cov}(Y_1, Y_2) \end{aligned}$$

# 协方差的性质

$$(5) \quad D(X+Y)=D(X)+D(Y)+2\text{Cov}(X, Y)$$

$$D(X-Y)=D(X)+D(Y)-2\text{Cov}(X, Y)$$

推广：

$$D(X+Y+Z)$$

$$=D(X)+D(Y)+ D(Z)+ 2\text{Cov}(X, Y) +2\text{Cov}(X, Z)+2\text{Cov}(Y, Z)$$



# 期望、方差、协方差的性质对比

期望	方差	协方差
$E(c)=c$	$D(c)=0$	$Cov(c,X)=0$
$E(aX)=aE(X),$	$D(aX)=a^2D(X),$	$Cov(aX,bY)$ $=abCov(X,Y)$
$E(X+Y)$ $=E(X)+E(Y)$	$D(X+Y)=D(X)+$ $D(Y)+2Cov(X,Y)$	$Cov(X+Y,Z)$ $=Cov(X,Z)$ $+Cov(Y,Z)$
当X与Y独立时 $E(XY)=E(X)E(Y)$		

协方差的大小在一定程度上反映了 $X$ 和 $Y$ 相互间的关系，但它还受 $X$ 与 $Y$ 本身度量单位的影响。例如：

$$\text{Cov}(aX, bY) = ab \text{Cov}(X, Y)$$

为了克服这一缺点，对协方差进行标准化，这就引入了相关系数。

# 相关系数

◆定义： 设 $D(X) > 0, D(Y) > 0$ ，则称

$$\rho_{XY} = \frac{Cov(X, Y)}{\sqrt{D(X) D(Y)}} \text{ 为随机变量 } X \text{ 和 } Y \text{ 的相关系数。}$$

相关系数是一个无量纲的量，实际上是 $X$ 与 $Y$ 的标准化随机变量的协方差，即：

$$\rho_{XY} = \frac{E\{[X - E(X)][Y - E(Y)]\}}{\sqrt{D(X) D(Y)}} = Cov\left(\frac{X - E(X)}{\sqrt{D(X)}}, \frac{Y - E(Y)}{\sqrt{D(Y)}}\right)$$

记  $\rho_{XY}$  为  $\rho$

# 相关系数

◆ **X与Y的相关系数  $|\rho_{XY}|$  是刻画X与Y之间线性关系紧密程度的数字特征**

当  $|\rho_{XY}|$  较大时，X与Y之间线性关系的程度较好；

当  $|\rho_{XY}|$  较小时，X与Y之间线性关系的程度较差；

当  $|\rho_{XY}| = 0$ ，即  $\text{Cov}(X, Y) = 0$ ，称X与Y不相关，即X与Y不存在线性关系。

# 相关系数的计算

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

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

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

会计算各类期望： $E(X), E(X^2), E(Y), E(Y^2), E(XY)$

**例1:** 二维离散型随机变量(X,Y)的联合分布律为:

$X \backslash Y$	-1	0	1
0	0	$\frac{1}{3}$	0
1	$\frac{1}{3}$	0	$\frac{1}{3}$

(1) 求:  $\rho_{XY}$ ; (2) 问X与Y是否独立?

解:

$X \backslash Y$	-1	0	1	$p_X$
0	0	$\frac{1}{3}$	0	$\frac{1}{3}$
1	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{2}{3}$
$p_Y$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	

$$E(X) = \frac{2}{3},$$

$$E(X^2) = \frac{2}{3},$$

$$E(Y) = 0,$$

$$E(Y^2) = \frac{2}{3},$$

$$E(XY) = 0$$

$$D(X) = E(X^2) - [E(X)]^2 = \frac{2}{9}, \quad D(Y) = \frac{2}{3}$$

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0 \quad \rho_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{D(X) D(Y)}} = 0$$

(2) X与Y不独立

说明: X与Y虽然没有线性关系,但不独立,即可能存在其他关系.

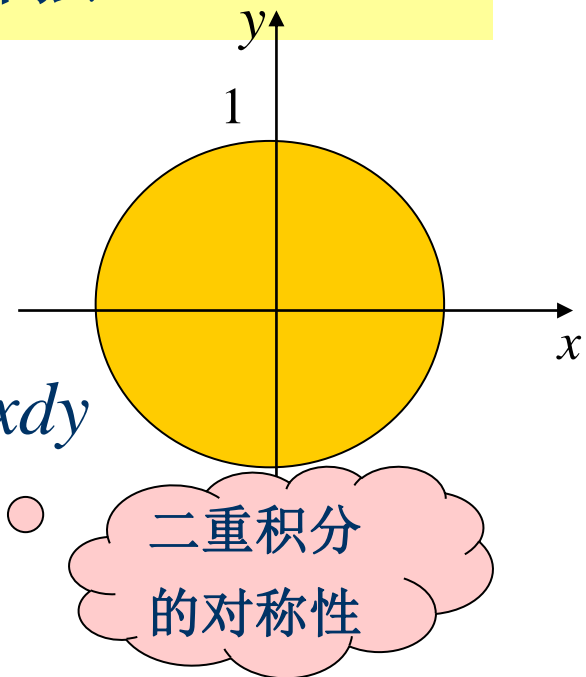
**例2:** 设 $(X,Y)$ 服从单位圆 $D=\{(x,y): x^2+y^2\leq 1\}$ 上的均匀分布, (1) 求  $\rho_{XY}$ , (2) 问 $X$ 与 $Y$ 是否独立?

**解:** 
$$f(x,y) = \begin{cases} \frac{1}{\pi}, & x^2 + y^2 \leq 1 \\ 0, & x^2 + y^2 > 1 \end{cases}$$

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xf(x,y) dxdy = \iint_{x^2+y^2 \leq 1} x/\pi dxdy \\ &= \pi^{-1} \int_{-1}^1 \left( \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} x dx \right) dy = 0 \end{aligned}$$

$$E(X^2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^2 f(x,y) dxdy = \iint_{x^2+y^2 \leq 1} x^2/\pi dxdy = \frac{1}{4}$$

由对称性可得,  $E(Y)=0$ ,  $E(Y^2)=1/4$





**例2:** 设 $(X,Y)$ 服从单位圆 $D=\{(x,y): x^2+y^2\leq 1\}$ 上的均匀分布, (1) 求  $\rho_{XY}$ ; (2) 问 $X$ 与 $Y$ 是否独立?

**解:** 
$$f(x,y) = \begin{cases} \frac{1}{\pi}, & x^2 + y^2 \leq 1 \\ 0, & x^2 + y^2 > 1 \end{cases}$$

**X与Y不独立**

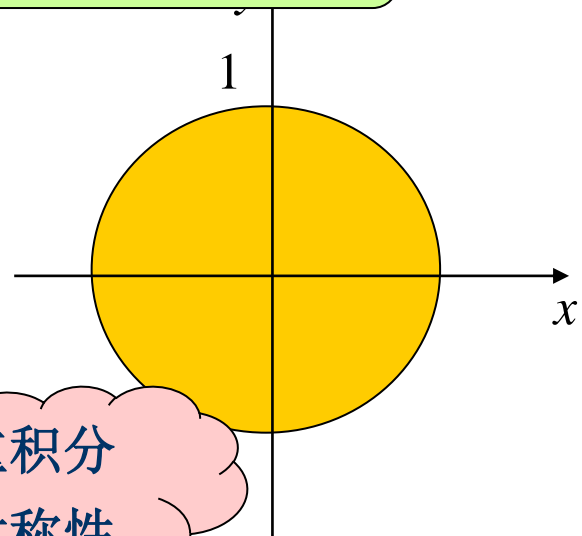
$$E(X) = E(Y) = 0, E(X^2) = E(Y^2) = \frac{1}{4}$$

$$E(XY) = \iint_{x^2+y^2 \leq 1} (xy/\pi) dx dy = 0$$

二重积分的对称性

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y) = 0$$

$$\rho_{XY} = \frac{\text{Cov}(X,Y)}{\sqrt{D(X) D(Y)}} = 0 \quad \text{即} X \text{与} Y \text{不线性相关。}$$



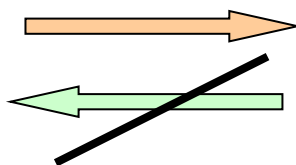
# 独立性与不相关性

■ 独立性与不相关性是讨论随机变量 $X$ 与 $Y$ 之间关系的两个概念；

➤  $X$ 和 $Y$ 独立：指 $X$ 与 $Y$ 互不影响，没有关系；

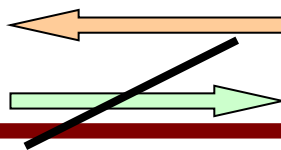
➤  $X$ 和 $Y$ 不相关：指 $X$ 与 $Y$ 之间没有线性关系。

$X$ 和 $Y$ 独立



$X$ 和 $Y$ 不(线性)相关

$X$ 和 $Y$ 不独立



$X$ 和 $Y$ (线性)相关

## 特 例

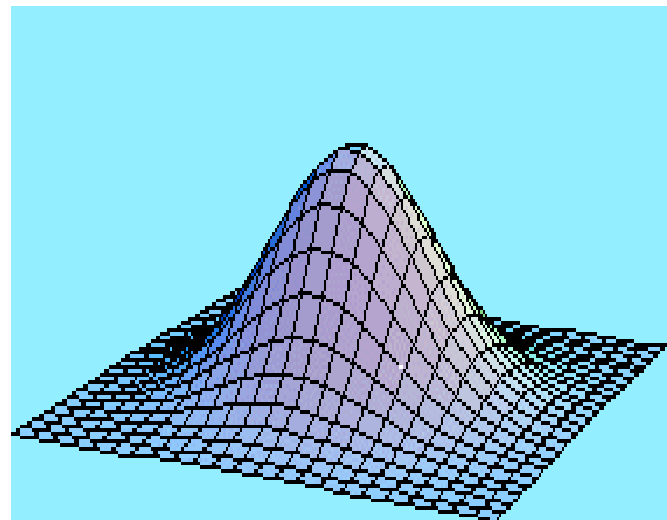
当 $(X,Y)$ 服从二维正态分布时，即

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

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

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

$$(-\infty < x < +\infty, -\infty < y < +\infty)$$



**X与Y独立**  $\longleftrightarrow$  **X与Y不相关**  $\longleftrightarrow$   $\rho = 0$

# 随机变量的矩

## ■ 定义:

- $X$  的  $k$  阶原点矩:  $E(X^k)$        $k=1, 2, \dots$
- $X$  的  $k$  阶中心矩:  $E\{[X-E(X)]^k\}$
- $X$  和  $Y$  的  $k+l$  阶混合原点矩:  $E(X^k Y^l)$        $k, l=1, 2, \dots$
- $X$  和  $Y$  的  $k+l$  阶混合中心矩:  
 $E\{[X-E(X)]^k [Y-E(Y)]^l\}$

有:  $X$  的期望  $E(X)$  是  $X$  的一阶原点矩;

方差  $D(X)$  是  $X$  的二阶中心矩;

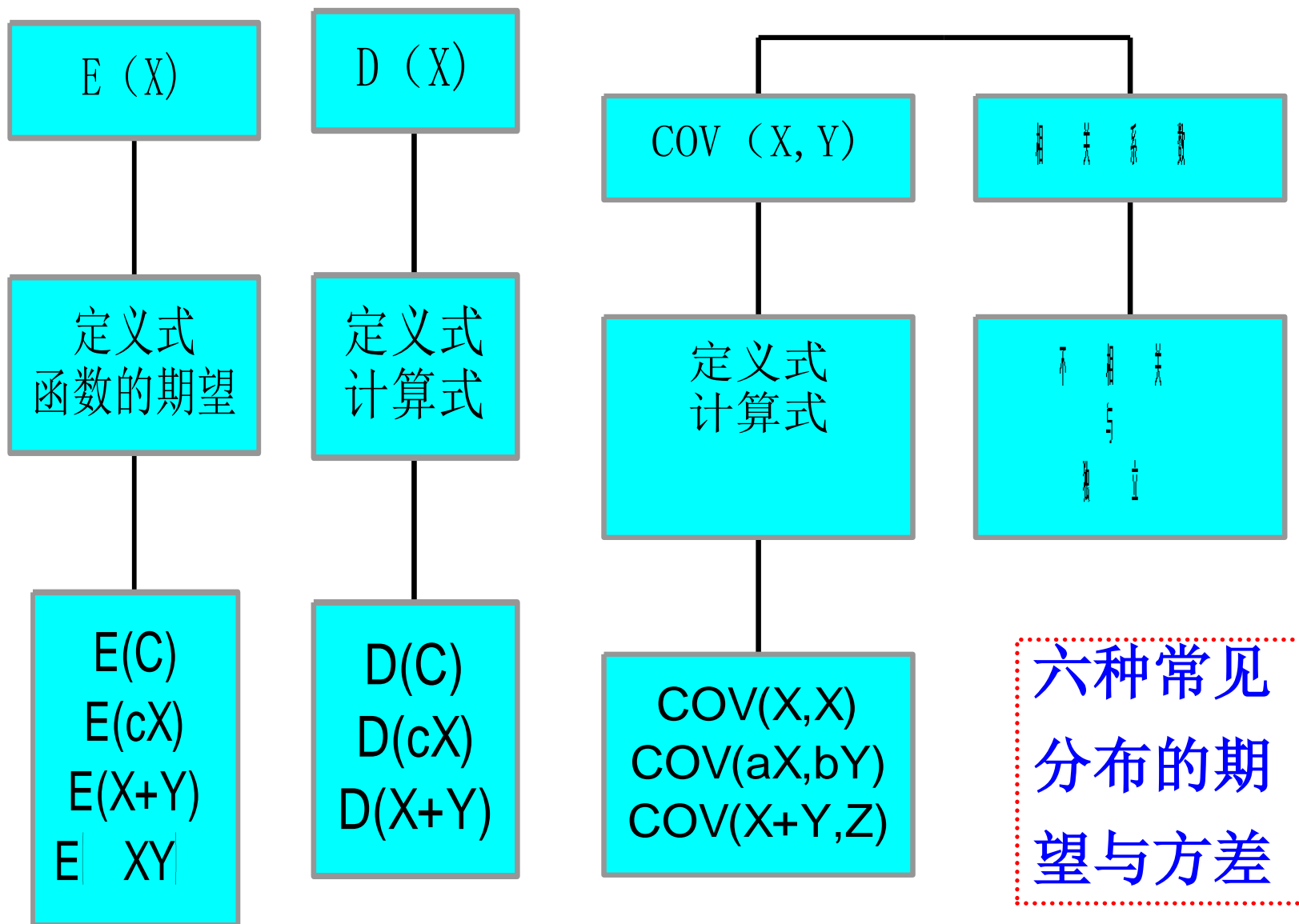
协方差  $\text{Cov}(X, Y)$  是  $X$  和  $Y$  的二阶混合中心矩.

矩的计算: 就是计算各类数学期望

# 小 结

本节首先介绍二维随机变量  $(X, Y)$  的分量  $X$  与  $Y$  的协方差及相关系数的概念、性质和计算；然后介绍随机变量的各种矩( $k$  阶原点矩、 $k$  阶中心矩、 $k+m$  阶混合原点矩、 $k+m$  阶混合中心矩),  $n$  维随机向量的协方差阵的概念、性质和计算.

# 数字特征小结



# 第五章

## 大数定理

### 中心极限定理

概率论与数理统计是研究**随机现象统计规律性**的一门学科。

随机现象的统计规律性只有在相同条件下进行**大量的重复试验**才能呈现出来。所以，要从随机现象中去寻求统计规律性，就应该对随机现象进行**大量的观测**。

研究随机现象的大量观测，常采用极限形式，由此导致了极限定理的研究。极限定理的内容很广泛，最重要的有两种：

**“大数定律”** 和 **“中心极限定理”**



# 大数定律

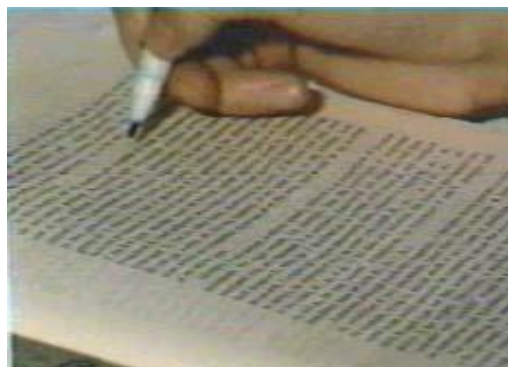
对随机现象进行大量重复的观测，各种结果的出现的频率具有稳定性。



大量地掷硬币  
正面出现频率



生产过程中  
废品率



字母使用频率

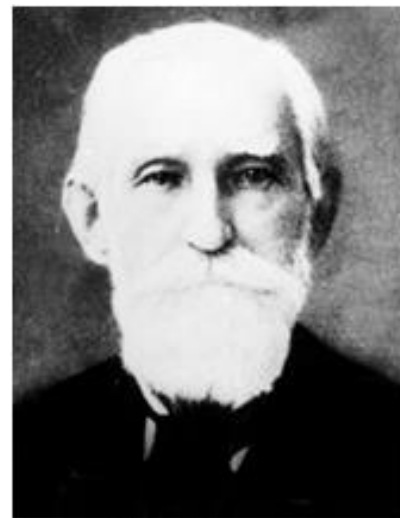
# 切比雪夫不等式

■ **定理：** 设随机变量 $X$  的数学期望  $E(X)=\mu$ ， 方差 $D(X)=\sigma^2$ ， 则对任给的  $\varepsilon > 0$ ， 有

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2},$$

或

$$P(|X - \mu| < \varepsilon) \geq 1 - \frac{\sigma^2}{\varepsilon^2}$$



切比雪夫, П. Л.

估计不等式

在随机变量 $X$ 的**分布未知**的情况下，只利用 $X$ 的期望和方差，即可对 $X$ 的概率分布进行估计。

**证明：** 只对 $X$  是连续型情况加以证明。

设 $X$  的概率密度函数为 $f(x)$ ，则有

$$\begin{aligned} P\{|X - \mu| \geq \varepsilon\} &= \int_{|x-\mu| \geq \varepsilon} f(x)dx \\ &\leq \int_{|x-\mu| \geq \varepsilon} \frac{(x-\mu)^2}{\varepsilon^2} f(x)dx \\ &\leq \int_{-\infty}^{+\infty} \frac{1}{\varepsilon^2} (x-\mu)^2 f(x)dx \\ &= \frac{1}{\varepsilon^2} \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx = \frac{\sigma^2}{\varepsilon^2} \end{aligned}$$

放大被积函数

放大积分区间

由切比雪夫不等式可以看出：若 $\sigma^2$  越小,则事件 $\{|X-E(X)| \leq \varepsilon\}$  的概率越大，即随机变量 $X$  集中在期望附近的可能性越大。

## 切比雪夫不等式的应用

**例：**已知正常男性成人血液中，每毫升白细胞数的平均值是7300，均方差是700，利用切比雪夫不等式估计每毫升血液含白细胞数在5200~9400之间的概率。

**解：**设X表示每毫升血液中含白细胞个数，则

$$E(X) = 7300, D(X) = 700^2$$

$$P(5200 \leq X \leq 9400) = P(|X - 7300| \leq 2100)$$

$$\geq 1 - \frac{D(X)}{2100^2} = 1 - \frac{1}{9} = \frac{8}{9}$$

# 练一练

设随机变量 $X$ 的方差为2.5，利用切比雪夫不等式估计概率

$$P(|X - E(X)| \geq 7.5)$$

答：

$$P(|X - E(X)| \geq 7.5) \leq \frac{2.5}{7.5^2} = \frac{1}{22.5}$$

# 大数定律

首先引入随机变量序列**相互独立**的概念。

■ **定义：** 设  $X_1, X_2, \dots$  是一随机变量序列。如果对任意的  $n > 1$ ,  $X_1, X_2, \dots, X_n$  相互独立, 则称  $X_1, X_2, \dots$  相互独立。

# 几个常见的大数定律

## ■ 定理1: 切比雪夫大数定律

设随机变量序列  $X_1, X_2, \dots$  相互独立, 具有相同的期望和方差:  $E(X_i) = \mu$ ,  $D(X_i) = \sigma^2$  ( $i=1, 2, \dots$ ), 则对任意的  $\varepsilon > 0$ , 有

$$\lim_{n \rightarrow +\infty} P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \mu\right| < \varepsilon\right) = 1$$

或者 
$$\lim_{n \rightarrow +\infty} P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \mu\right| \geq \varepsilon\right) = 0$$

证明：记  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$

$$E(\bar{X}) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \mu ,$$

$$D(\bar{X}) = D\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n D(X_i) = \frac{\sigma^2}{n}$$

$$P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \mu\right| < \varepsilon\right) = P\left(\left|\bar{X} - \mu\right| < \varepsilon\right) \geq 1 - \frac{\sigma^2}{n\varepsilon^2}$$

令  $n \rightarrow \infty$ ，并注意到概率小于等于1，结论成立。



## ■ 定理2: 伯努利大数定律

-----频率的稳定性

设 $X$ 表示在 $n$ 次独立重复试验中事件 $A$ 发生的次数,  $p$ 是每次试验中事件 $A$ 发生的概率, 即  $X \sim B(n, p)$



雅各布第一·伯努利

引入  $X_i = \begin{cases} 1, & \text{第 } i \text{ 次试验中事件 } A \text{ 发生,} \\ 0, & \text{第 } i \text{ 次试验中事件不 } A \text{ 发生,} \end{cases} \quad i = 1, 2, \dots, n.$

则 $X_1, X_2, \dots, X_n$  相互独立,都服从0-1分布,且  $X = \sum_{i=1}^n X_i$ ,

$$\frac{X}{n} = \frac{1}{n} \sum_{i=1}^n X_i \text{ 是 } n \text{ 次试验中 } A \text{ 发生的频率}$$

## ■ 定理：伯努利大数定律

-----频率的稳定性

设 $X$ 是  $n$  次独立重复试验中事件 $A$ 发生的次数， $p$ 是事件 $A$ 在每次试验中发生的概率，则对于任意正数 $\varepsilon$ ，恒有

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{X}{n} - p \right| < \varepsilon \right\} = 1$$

或者 
$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{X}{n} - p \right| \geq \varepsilon \right\} = 0$$

# 定理说明

◆ 伯努利大数定律表明：当重复试验次数 $n$ 充分大时，事件 $A$ 发生的频率 $X/n$ 与事件 $A$ 发生的概率 $p$ 有较大偏差的概率很小。即从理论上严格证明了频率具有的稳定性。

◆ 定理的应用：可通过多次重复一个试验，确定事件 $A$ 在每次试验中出现的概率

$$\frac{X}{n} \approx p = P(A)$$

# 中心极限定理 Central limit theorem

## ◆ 客观背景

在实际问题中，有许多随机现象可以看做是由很多因素独立影响的综合结果；而每一个因素对该现象的影响都很微小，但总起来却对总和有显著影响；那么描述这种随机现象的随机变量可以看成很多相互独立的起微小作用的因素的总和，它往往近似地服从正态分布。

# 中心极限定理



**例如：**炮弹射击的落点与目标的偏差，就受着许多随机因素(如瞄准，空气阻力，炮弹或炮身结构等)综合影响的. 每个随机因素的对弹着点（随机变量和）所起的作用都是很小的. 那么弹着点服从怎样分布呢？

# 中心极限定理

自从高斯指出测量误差服从正态分布之后，人们发现：正态分布在自然界中极为常见。



如果一个随机变量是由大量相互独立的随机因素的综合影响所造成，而每一个别因素对这种综合影响中所起的作用不大；则这种随机变量一般都服从或近似服从正态分布。

# 中心极限定理

中心极限定理是棣莫弗 (De Moivre) 在18世纪首先提出的，到现在内容已十分丰富；在这里，我们只介绍其中两个最基本的结论：

- ① 当  $n$  无限增大时，独立同分布随机变量之和的极限分布是正态分布；
- ② 当  $n$  很大时，二项分布可用正态分布近似。

# 中心极限定理

由于无穷个随机变量之和可能趋于 $\infty$ ，故我们不研究 $n$ 个随机变量之和本身，而考虑其标准化的随机变量：  
的极限分布。

$$Y_n = \frac{\sum_{k=1}^n X_k - E\left(\sum_{k=1}^n X_k\right)}{\sqrt{D\left(\sum_{k=1}^n X_k\right)}}$$

在概率论中，习惯于把随机变量之和的分布收敛于正态分布这一类定理都叫做中心极限定理。



## ■ 定理1: 独立同分布的中心极限定理

### 林德伯格-列维(Lindeberg-Levy)中心极限定理

设  $X_1, X_2, \dots$  是独立同分布的随机变量序列, 且  $E(X_i) = \mu, D(X_i) = \sigma^2$ , 则对任给  $x \in (-\infty, +\infty)$ , 均有

$$\lim_{n \rightarrow \infty} P \left( \frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} \leq x \right) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = \Phi(x)$$

其中  $\Phi(x)$  是标准正态分布  $N(0, 1)$  的分布函数。

$$\lim_{n \rightarrow \infty} P \left( \frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} \leq x \right) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = \Phi(x)$$

记  $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$  , 有

还有另一记法:

$$\lim_{n \rightarrow \infty} P \left( \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \leq x \right) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = \Phi(x)$$

# 定理说明

◆ 定理表明：对于独立的随机变量序列  $\{X_n\}$ ，不管  $X_i$  ( $i=1,2,\dots,n$ ) 服从什么分布，只要它们是同分布，且有有限的数学期望和方差，那么当  $n$  充分大时，这些随机变量之和近似地服从正态分布，即

$$\sum_{i=1}^n X_i \overset{\text{近似地}}{\sim} N(n\mu, n\sigma^2); \quad \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n\sigma}} \overset{\text{近似地}}{\sim} N(0,1)$$

◆ 定理的应用：在一般情况下，我们很难求出  $\sum_{i=1}^n X_i$  的分布的确切形式，但当  $n$  很大时，可以求出近似分布。

## ■ 定理2: 棣莫佛-拉普拉斯中心极限定理

设随机变量 $X$ 服从参数为 $(n, p)$ 的二项分布( $0 < p < 1$ ), 则对任意  $x \in (-\infty, +\infty)$ , 均有

$$\lim_{n \rightarrow \infty} P \left\{ \frac{X - np}{\sqrt{np(1-p)}} \leq x \right\} = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = \Phi(x)$$

■ 定理表明: 当 $n$ 很大,  $0 < p < 1$ 是一个定值时 (或者说,  $np(1-p)$ 也不太小时), 二项分布  $B(n, p)$  近似正态分布  $N(np, np(1-p))$ .

一般地，如果随机变量  $X \sim B(n, p)$ ，则有

$$\begin{aligned} P(a \leq X < b) &= P\left(\frac{a - np}{\sqrt{np(1-p)}} \leq \frac{X - np}{\sqrt{np(1-p)}} < \frac{b - np}{\sqrt{np(1-p)}}\right) \\ &\approx \Phi\left(\frac{b - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a - np}{\sqrt{np(1-p)}}\right) \end{aligned}$$

即  $B(n, p) \overset{\text{近似}}{\sim} N(np, np(1-p))$

**例1:** 设一批产品的强度服从期望为14, 方差为4的分布, 每箱中装有这种产品100件。求

(1) 每箱产品的平均强度超过14.5的概率;

(2) 每箱产品的平均强度超过期望14的概率。

**解:** 设 $X_i$ 是第 $i$ 件产品的强度( $i=1,2,\dots,100$ ), 且有 $E(X_i)=14$ ,  $D(X_i)=4$ ,  $X_i$ 相互独立。则每箱产品的平均强度为  $Y = \frac{1}{100} \sum_{i=1}^{100} X_i$

$$\text{且} \begin{cases} E(Y) = E\left(\frac{1}{100} \sum_{i=1}^{100} X_i\right) = \frac{1}{100} \sum_{i=1}^{100} E(X_i) = 14 \\ D(Y) = D\left(\frac{1}{100} \sum_{i=1}^{100} X_i\right) = \frac{1}{100^2} \sum_{i=1}^{100} D(X_i) = \frac{4}{100} \end{cases}$$

(1) 每箱产品的平均强度超过14.5的概率;

$$E(Y) = 14, \quad D(Y) = \frac{4}{100} \quad \overset{\text{近似}}{\therefore Y \sim N(14, 0.04)}$$

$$P(Y > 14.5) \approx 1 - \Phi\left(\frac{14.5 - 14}{\sqrt{0.04}}\right)$$

$$= 1 - \Phi(2.5) = 0.0062$$

(2) 每箱产品的平均强度超过期望14的概率

$$E(Y) = 14, \quad D(Y) = \frac{4}{100} \quad \therefore Y \overset{\text{近似}}{\sim} N(14, 0.04)$$

$$P(Y > 14)$$

$$\approx 1 - \Phi\left(\frac{14 - 14}{\sqrt{0.04}}\right) = 1 - \Phi(0) = 0.5$$



**例2:** 某公司有200名员工参加一种资格证书考试，按往年经验，考试通过率为0.8。试计算这200名员工至少有150人考试通过的概率。

**解法一:** 令  $X_i = \begin{cases} 1, & \text{第} i \text{个人考试通过,} \\ 0, & \text{第} i \text{个人考试未通过.} \end{cases} \quad i = 1, 2, \dots, 200$

$X_i$  相互独立.

$X_i$	0	1	$E(X_i) = 0.8, \quad D(X_i) = 0.16$ $i = 1, 2, \dots, 200$
$P$	0.2	0.8	

考试通过人数为  $\sum_{i=1}^{200} X_i$ ，且有

$$E\left(\sum_{i=1}^{200} X_i\right) = \sum_{i=1}^{200} E(X_i) = 160, \quad D\left(\sum_{i=1}^{200} X_i\right) = \sum_{i=1}^{200} D(X_i) = 32$$

$$\sum_{i=1}^{200} X_i \overset{\text{近似}}{\sim} N(160, 32)$$

于是

$$\begin{aligned} P\left(\sum_{i=1}^{200} X_i \geq 150\right) &= 1 - \Phi\left(\frac{150 - 160}{\sqrt{32}}\right) \\ &= 1 - \Phi(-1.77) \\ &= \Phi(1.77) = 0.96 \end{aligned}$$

**例2：**某公司有200名员工参加一种资格证书考试，按往年经验，考试通过率为0.8。试计算这200名员工至少有150人考试通过的概率。

**解法二：** 设X表示通过考试的人数，则

$$X \sim B(200, 0.8)$$

$$E(X) = 200 * 0.8 = 160, \quad D(X) = 200 * 0.8 * 0.2 = 32$$

所以：

$$P(X \geq 150) \begin{cases} = \sum_{i=150}^{200} C_{200}^i (0.8)^i (0.2)^{200-i} \\ \approx 1 - \Phi\left(\frac{150 - 160}{\sqrt{32}}\right) \\ = 1 - \Phi(-1.77) = \Phi(1.77) = 0.96 \end{cases}$$

**例3:** 已知一本300页的书中，每页印刷错误个数服从参数为0.2的泊松分布，求这本书印刷错误总数不多于70的概率。

**解:** 设 $X_i$ 表示第 $i$ 页上的印刷错误个数( $i=1,2,\dots,300$ ),

则  $X_i \sim \pi(0.2)$  且  $X_i$  相互独立, 且  $E(X_i)=D(X_i)=0.2$

$$E\left(\sum_{i=1}^{300} X_i\right) = \sum_{i=1}^{300} E(X_i) = 60, \quad D\left(\sum_{i=1}^{300} X_i\right) = \sum_{i=1}^{300} D(X_i) = 60$$

$$\sum_{i=1}^{300} X_i \stackrel{\text{近似}}{\sim} N(60, 60)$$

$$P\left(\sum_{i=1}^{300} X_i \leq 70\right) \approx \Phi\left(\frac{70-60}{\sqrt{60}}\right) = \Phi(1.29) = 0.95$$

**例4:** 某市保险公司开办一年人身保险业务,被保险人每年需交付保费160元。若一年内发生重大人身事故,其本人或家属获赔付金2万元。已知该市人员一年内发生重大人身事故的概率为0.005, 现有5000人参加此项保险。求保险公司一年内从此项业务所得到的总收益在20万元到40万元之间的概率。

**解:** 设X表示5000人中在一年内发生重大人身事故的人数, 则

$$X \sim B(5000, 0.005), \quad E(X)=25, \quad D(X)=24.875$$

$$P(20\text{万元} \leq \text{总收益} \leq 40\text{万元})$$

$$= P(20 \leq 0.016 \times 5000 - 2X \leq 40) = P(20 \leq X \leq 30)$$

$$\approx \Phi\left(\frac{30-25}{\sqrt{24.875}}\right) - \Phi\left(\frac{20-25}{\sqrt{24.875}}\right) = \Phi(1.0025) - \Phi(-1.0025)$$

$$= 2\Phi(1.0025) - 1 = 0.6839$$

**例5:** 设某单位设置一台电话总机,共有200个分机,设每个分机有5%的时间要使用外线通话,并且每个分机使用外线与否是相互独立的,该单位需要多少外线才能保证每个分机要使用外线时可供使用的概率达到0.9?

**解:** 设单位需要 $n$ 条外线, 设随机变量 $X$ 表示任一时间需要使用外线的分机数, 则

$$X \sim B(200, 0.05)$$

$$E(X) = 200 * 0.05 = 10, \quad D(X) = 200 * 0.05 * 0.95 = 9.5$$

保证每个分机要使用外线时可供使用的概率达到0.9



$$P(X \leq n) \geq 0.9$$

$$\begin{cases} P(X \leq n) = \Phi\left(\frac{n-10}{\sqrt{9.5}}\right) \\ 0.9 = \Phi(1.28) \end{cases} \quad \frac{n-10}{\sqrt{9.5}} \geq 1.28 \quad \Rightarrow \quad n \geq 14$$

练习：重复做10000次独立试验，A在每次试验中发生的概率为 $p$ ，分别利用切比雪夫不等式和中心极限定理估计：用A在10000次试验中发生的频率作为概率的 $p$ 的估计值时误差小于0.01的概率的近似值。

切比雪夫不等式用于：估计概率、证明不等式。

例：设 $X$ 服从参数为 $\lambda$ 的泊松分布，

证明：  $P(0 < X < 2\lambda) \geq \frac{\lambda - 1}{\lambda}$ .

证明：  $P(0 < X < 2\lambda) = P(|X - \lambda| < \lambda)$

$$\begin{aligned} &\geq 1 - \frac{DX}{\lambda^2} \\ &= 1 - \frac{\lambda}{\lambda^2} \\ &= \frac{\lambda^2 - \lambda}{\lambda^2} = \frac{\lambda - 1}{\lambda} \end{aligned}$$



- 利用中心极限定理：
  - 求区间内的概率
  - 已知概率求 **n** 的值。

## 小 结

本章介绍了大数定律和中心极限定理

中心极限定理表明：对于独立的随机变量序列  $\{X_n\}$ ，不管  $X_i$  ( $i=1,2,\dots,n$ ) 服从什么分布，只要它们是同分布，且有有限的数学期望和方差，那么当  $n$  充分大时，这些随机变量之和近似地服从正态分布，即

$$\sum_{i=1}^n X_i \overset{\text{近似地}}{\sim} N\left(E\left(\sum_{i=1}^n X_i\right), D\left(\sum_{i=1}^n X_i\right)\right)$$

## 第六章 数理统计基础

- ◆ 基本概念
- ◆ 三个常见分布
- ◆ 正态总体下统计量的分布

数理统计不同于一般的资料统计,它更侧重于应用随机现象本身的规律性进行资料的收集、整理和分析。

因此,数理统计中的方法和支持这些方法的相应理论是相当丰富的;概括起来可以归纳成两大类:

**参数估计:** 根据数据,对分布中的未知参数进行估计;

**假设检验:** 根据数据,对分布的未知参数的某种假设进行检验。

参数估计与假设检验构成了统计推断的两种基本形式,这两种推断渗透到了数理统计的每个分支。

# 数理统计中的几个概念

## ◆ 总体、个体

一个统计问题总有它明确的研究对象。

**总体：**研究对象的全体；

**个体：**总体中每个成员；

**总体的容量：**总体中所包含的个体的个数。



研究某批灯泡的质量

总体 { 有限总体  
无限总体

例如：研究某工厂生产的某种产品的废品率，则这种产品的全体就是总体，而每件产品都是一个个体。

实际上，我们真正关心的并不一定是总体或个体本身，而是总体或个体的某一项或某几项数量指标。

如：某电子产品的使用寿命，某天的最高气温，加工出来的某零件的长度等数量指标。

因此，有时也将总体理解为那些研究对象的某项数量指标的全体。

对一个总体，如果用 $X$  表示其数量指标。

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。所以， $X$  的值对不同的个体就取不同的值。因此，如果我们随机地抽取个体，则 $X$  的值也就随着抽取个体的不同而不同。

所以， $X$  是一个随机变量！

既然总体是随机变量 $X$ ，自然就有其概率分布。我们把 $X$  的分布称为总体分布。

总体的特性是由总体分布来刻画的。因此，常把总体和总体分布视为同义语。

**例：**若研究某地区  $N$  个农户的年收入。

在这里，总体既指这  $N$  个农户，又指我们所关心的  $N$  个农户的数量指标—他们的年收入( $N$  个数字)。

假定  $N$  户的年收入  $X$  只取以下各值: 0.5, 0.8, 1.0, 1.2和1.5; 取上述值的户数分别 $n_1, n_2, n_3, n_4$ 和 $n_5$   
( $n_1+n_2+n_3+n_4+n_5=N$ )。则 $X$ 为离散型分布，分布律为：

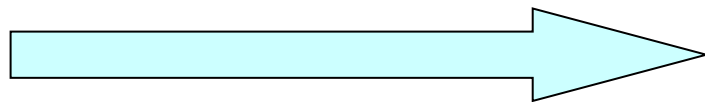
$X$	0.5	0.8	1	1.2	1.5
$P$	$\frac{n_1}{N}$	$\frac{n_2}{N}$	$\frac{n_3}{N}$	$\frac{n_4}{N}$	$\frac{n_5}{N}$



**例：**研究某批灯泡的寿命时，关心的数量指标就是寿命；那么，总体就可以用随机变量  $X$  表示，或用其分布函数  $F(x)$  表示。



寿命  $X$  可用一概率  
(指数)分布来刻画



寿命总体是指数分布总体

某批灯泡  
的寿命



鉴于此，常用随机变量的记号  
或用其分布函数表示总体。如  
说总体  $X$  或总体  $F(x)$ 。

类似地，在研究某地区中学生的营养状况时，若关心的数量指标是身高和体重，我们用 $X$ 和 $Y$ 分别表示身高和体重，那么此总体就可用二维随机变量 $(X,Y)$ 或其联合分布函数  $F(x,y)$ 来表示。



统计中，总体这个概念的要旨是：总体就是一个概率分布。

## ◆ 简单随机样本

总体分布一般是未知，或只知道是包含未知参数的分布

对总体进行研究的两种方法：

- 全面调查：如人口普查；
- 抽样调查：常用

按一定规则，从总体中抽取若干个体进行观察试验，以获得有关总体的信息，这一抽取过程称为**抽样**，所抽取的部分个体称为**样本**，样本中所包含的个体数目称为**样本容量**。

## ◆ 简单随机样本

数理统计是利用样本信息，对总体进行分析、估计、推断；故要求样本应具有**很好的代表性**。

**简单随机抽样：**采用独立、重复的随机抽样；

而抽得的这 $n$ 个个体称为一个**简单随机样本**，常表示为  $X_1, X_2, \dots, X_n$  或者  $(X_1, X_2, \dots, X_n)$

样本在抽样前，是一个 $n$ 维随机变量 $X_1, X_2, \dots, X_n$ ；  
在抽样后，是一组数据  $x_1, x_2, \dots, x_n$ ，称为样本的观测值， $n$ 为样本容量。

## ◆ 简单随机样本的性质

简单随机样本 $X_1, X_2, \dots, X_n$ 来自于总体 $X$ ，具有以下两条性质：

① 独立性：  $X_1, X_2, \dots, X_n$ 相互独立

② 同分布性：  $X_1, X_2, \dots, X_n$ 与总体 $X$ 具有相同的分布

**说明：**当我们说到总体及样本时，既指研究对象又指它们的某项数量指标。

## ◆ 统计量

由样本推断总体的某些情况时，需要对样本进行“加工”，构造出若干个样本的已知(确定)的函数，其作用是把样本中所含的某一方面的信息集中起来。

把不含任何未知参数的样本的函数称为统计量，它是完全由样本所决定的量。

即称样本 $X_1, \dots, X_n$ 的函数 $g(X_1, \dots, X_n)$ 是总体 $X$ 的一个统计量，如果 $g(X_1, \dots, X_n)$ 不含未知参数。

# 几个常见统计量

设总体为 $X$ ，样本为 $X_1, \dots, X_n$ ，样本观测值为 $x_1, \dots, x_n$

□ 样本均值

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

反映总体  
均值的信息

□ 样本方差

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

反映总体  
方差的信息

□ 样本标准差

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$$

它们的观测值用相应的小写字母 $x_1, \dots, x_n$ 表示

# 几个常见统计量

□ 样本  $k$  阶原点矩  $A_k = \frac{1}{n} \sum_{i=1}^n X_i^k$

反映总体  
 $k$  阶矩的信息

□ 样本  $k$  阶中心矩  $B_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k$

$$k=1, 2, \dots$$

反映总体  $k$  阶  
中心矩的信息

它们的观测值用相应的小写字母  $x_1, \dots, x_n$  表示



## 几点说明:

1. 总体均值 $E(X)$ 是常数, 而样本均值 $\bar{X}$ 是随机变量, 是两个不同的概念, 不能混淆; 当然两者之间有一定的关系。
2. 同样, 总体方差 $D(X)$ 与样本方差 $S^2$ 、总体矩与样本矩也是不同的概念。
3. 一些关系式:  $E(\bar{X}) = E(X)$ ,  $D(\bar{X}) = \frac{1}{n} D(X)$

# 数理统计中常用的三个分布

数理统计中常用的分布除正态分布外，还有三个非常有用的连续型分布，即

$\chi^2$ 分布	}	数理统计的三大分布 (都是连续型)
$t$ 分布		
$F$ 分布		
		它们都与正态分布有密切的联系



在本章中特别要求掌握对正态分布、 $\chi^2$ 分布、 $t$ 分布、 $F$ 分布的一些结论的熟练运用. 它们是后面各章的基础.

# 数理统计中常用的三个分布

## 一. $\chi^2$ 分布

1. 构造：设  $X_1, X_2, \dots, X_n$  相互独立，且都服从标准正态分布  $N(0, 1)$ ，则称随机变量

$$X = X_1^2 + X_2^2 + \dots + X_n^2$$

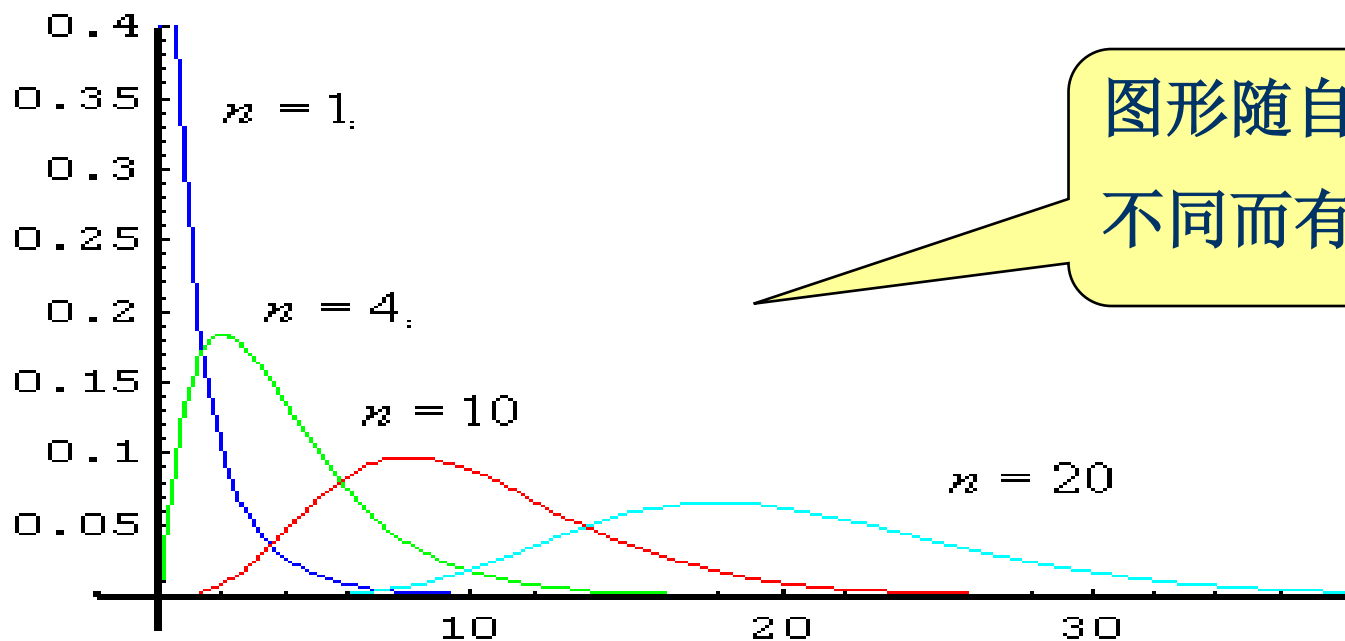
服从自由度为  $n$  的卡方分布，记成  $X \sim \chi^2(n)$ ；反之，若  $X \sim \chi^2(n)$ ，则  $X$  可以分解成  $n$  个相互独立的标准正态随机变量的平方和。

自由度是指独立随机变量的个数： $df=n$

# $\chi^2$ 分布

## 2. 图形: $\chi^2$ 分布的概率密度函数 $f(x)$ 曲线

$$f(x) = \begin{cases} \frac{1}{2^{n/2} \Gamma(n/2)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$



图形随自由度的不同而有所改变

# $\chi^2$ 分布

## 3. 性质:

① 期望与方差: 若  $X \sim \chi^2(n)$ , 则

$$E(X)=n, D(X)=2n$$

② 可加性: 若  $X \sim \chi^2(n)$ ,  $Y \sim \chi^2(m)$ , 且相互独立, 则

$$X + Y \sim \chi^2(n+m)$$

③ 补充: 设  $X_1, X_2, \dots, X_n$  为取自正态总体  $X \sim N(\mu, \sigma^2)$  的样本, 则

$$\frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} \sim \chi^2(n)$$

注: 由中心极限定理可以推出,  $n$  充分大时,  $\frac{X - n}{\sqrt{2n}}$  近似于标准正态分布  $N(0,1)$ 。

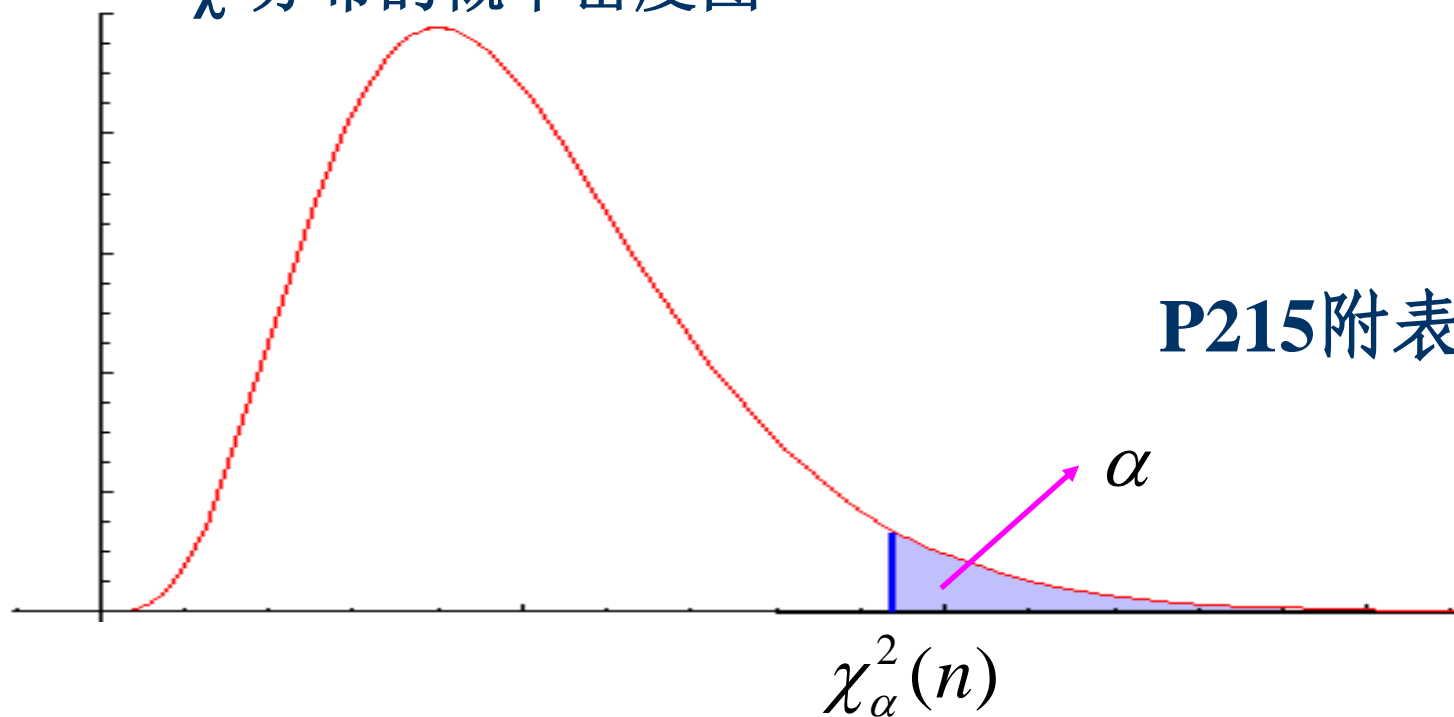
# $\chi^2$ 分布

4. 分位点：设 $X \sim \chi^2(n)$ ，若对于 $\alpha$ ：  $0 < \alpha < 1$ ，满足

$$P(X > \chi_{\alpha}^2(n)) = \int_{\chi_{\alpha}^2(n)}^{+\infty} f(x) dx = \alpha;$$

的点  $\chi_{\alpha}^2(n)$  为  $\chi^2$  分布的上侧分位点。

$\chi^2$ 分布的概率密度图



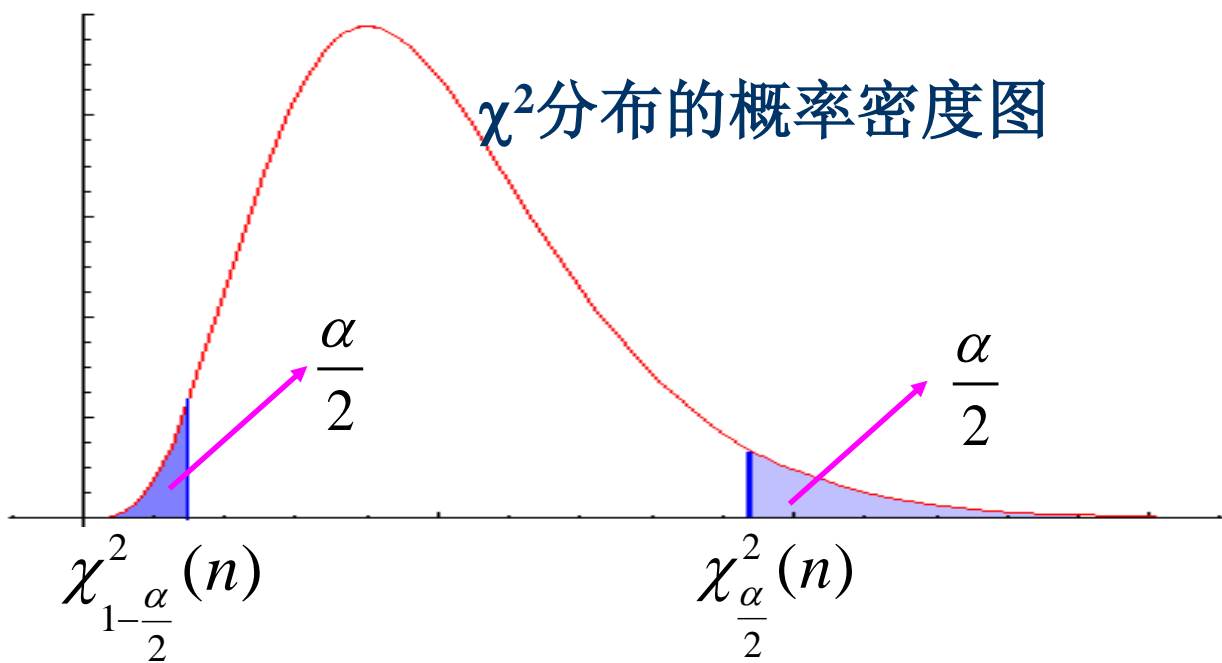
P215附表4

# $\chi^2$ 分布

◆  $\chi^2$ 分布的**单侧**分位点 $\chi_{\alpha}^2(n)$ :  $P(X > \chi_{\alpha}^2(n)) = \alpha$

◆  $\chi^2$ 分布的**双侧**分位点:  $P(a < X < b) = 1 - \alpha$

$$a = \chi_{1-\frac{\alpha}{2}}^2(n), \quad b = \chi_{\frac{\alpha}{2}}^2(n)$$



# 数理统计中常用的三个分布

## 二. $t$ 分布

1. 构造：若  $X \sim N(0, 1)$ ,  $Y \sim \chi^2(n)$ ,  $X$  与  $Y$  独立，则

称随机变量

$$T = \frac{X}{\sqrt{Y/n}} \sim t(n)$$

服从自由度为  $n$  的  $t$  分布，记成  $T \sim t(n)$ ;

反之，若  $T \sim t(n)$ ，则有相互独立的  $X \sim N(0, 1)$ ,

$Y \sim \chi^2(n)$ ，使  $T = \frac{X}{\sqrt{Y/n}}$

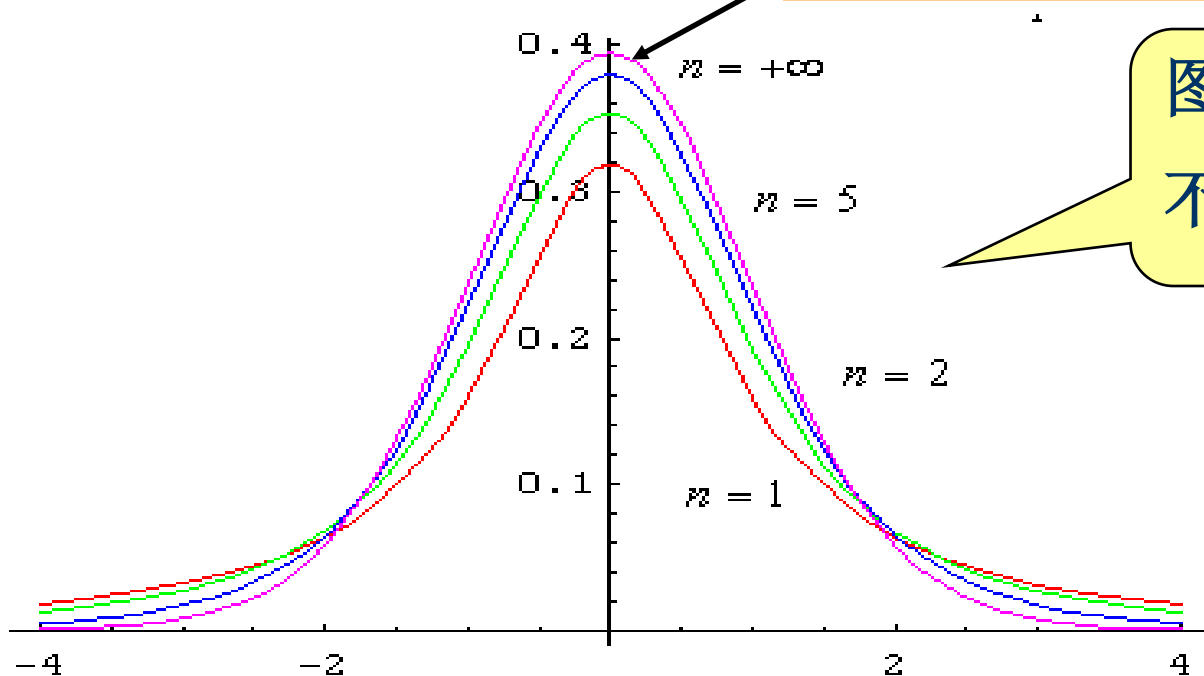


# $t$ 分布

## 2. 图形: $t$ 分布的概率密度函数 $f(x)$ 曲线

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

标准正态分布



图形随自由度的不同而有所改变

# $t$ 分布

## 3. 性质:

### ① 极限分布:

$$\lim_{n \rightarrow \infty} f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = \varphi(t), -\infty < x < \infty$$

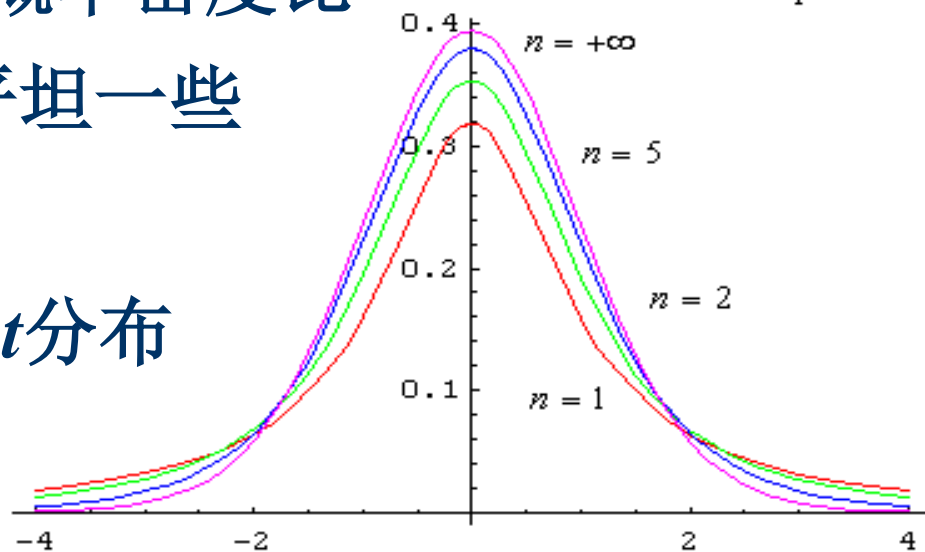
即 $t$ 分布的极限分布是标准正态分布

### ② 期望和方差: 若 $X \sim t(n)$ , 则

$E(X)=0$ , 因为 $f(x)$ 关于 $y$ 轴对称;

$D(X)>1$ , 因为 $t$ 分布的概率密度比标准正态分布的图形要平坦一些

当 $n>45$ 时, 可用 $N(0,1)$ 代替 $t$ 分布



# $t$ 分布

4. 单侧分位点：设  $X \sim t(n)$ ，若对于  $\alpha: 0 < \alpha < 1$ ，满足

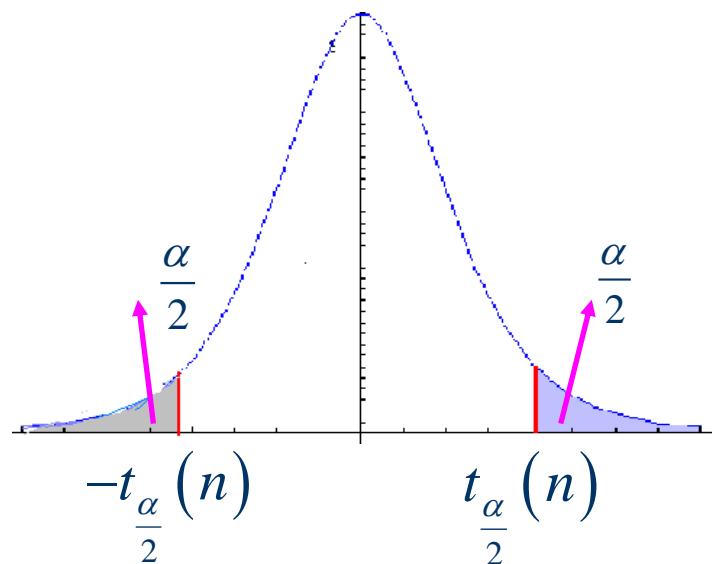
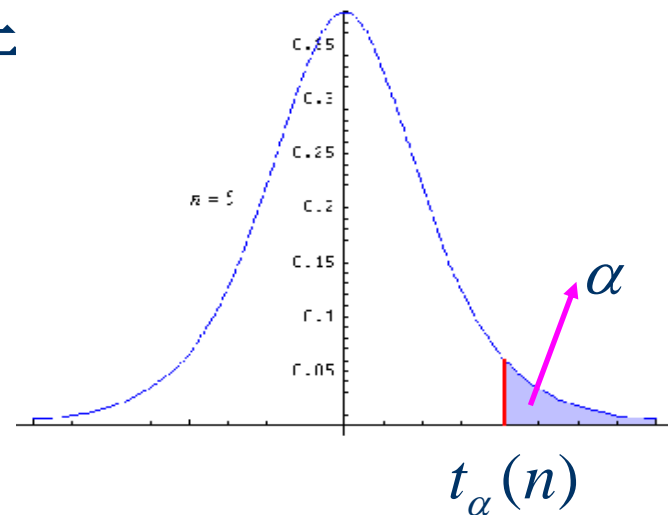
$$P(X > t_{\alpha}(n)) = \int_{t_{\alpha}(n)}^{+\infty} f(x) dx = \alpha$$

的点  $t_{\alpha}(n)$  为  $t$  分布的上侧分位点

双侧分位点：设  $X \sim t(n)$ ，若满足

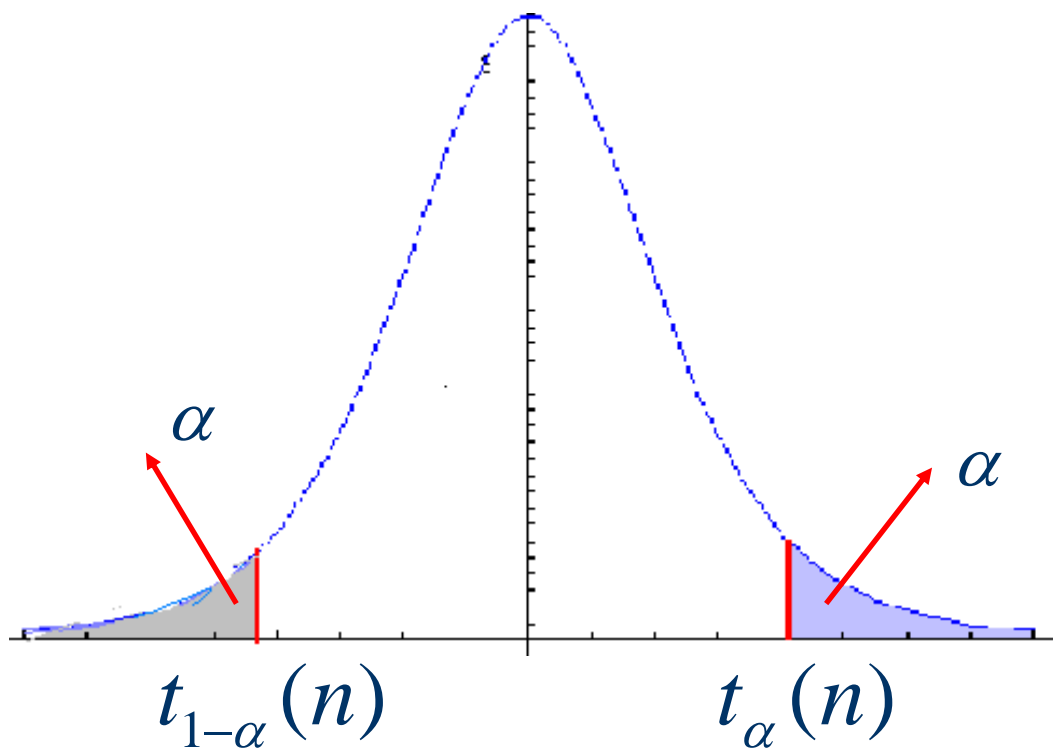
$$P(a < X < b) = 1 - \alpha$$

$$a = -t_{\frac{\alpha}{2}}(n), \quad b = t_{\frac{\alpha}{2}}(n)$$



# $t$ 分布

注:  $t_{1-\alpha}(n) = -t_{\alpha}(n)$



# 数理统计中常用的三个分布

## 三. $F$ 分布

1. 构造: 若  $X \sim \chi^2(n_1)$ ,  $Y \sim \chi^2(n_2)$ ,  $X$  与  $Y$  独立, 则称随机变量

$$F = \frac{X / n_1}{Y / n_2} \sim F(n_1, n_2)$$

服从第一自由度为  $n_1$ , 第二自由度为  $n_2$  的  $F$  分布, 记成  $F \sim F(n_1, n_2)$ ;

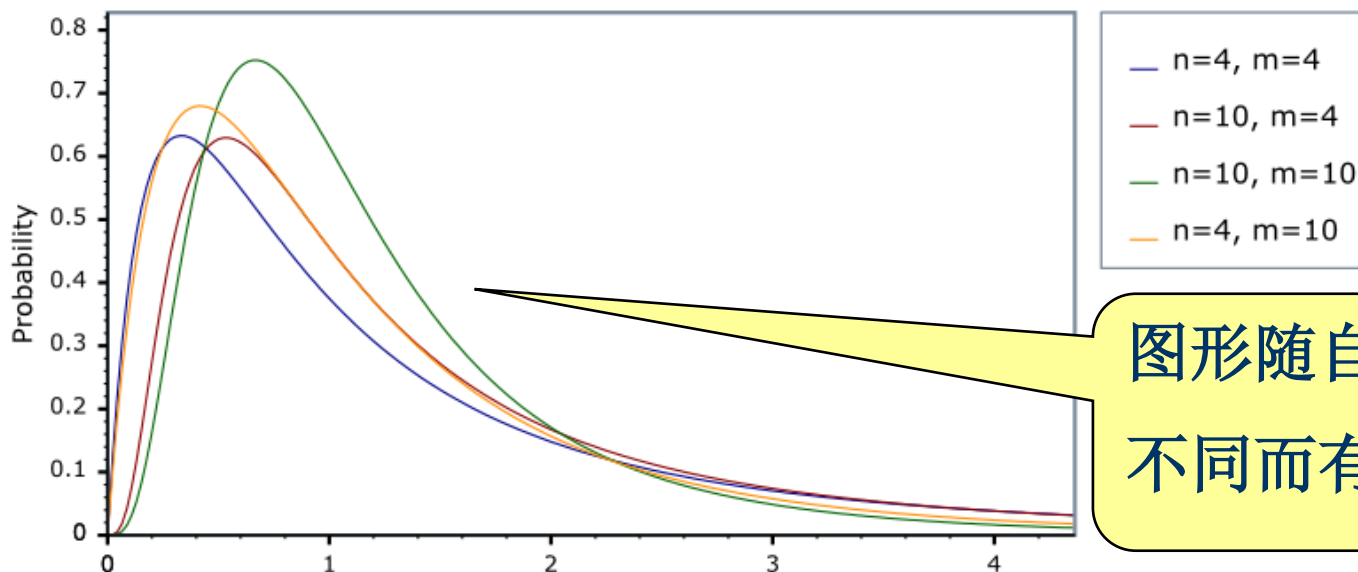
反之, 若  $F \sim F(n_1, n_2)$ , 则有相互独立的  $X \sim \chi^2(n_1)$ ,  $Y \sim \chi^2(n_2)$ , 使

$$F = \frac{X / n_1}{Y / n_2}$$

# $F$ 分布

## 2. 图形: $t$ 分布的概率密度函数 $f(x)$ 曲线

$$f(x) = \begin{cases} \frac{\Gamma\left(\frac{n_1 + n_2}{2}\right) \left(\frac{n_1}{n_2}\right)^{n_1/2} x^{\frac{n_1}{2}-1}}{\Gamma\left(\frac{n_1}{2}\right) \Gamma\left(\frac{n_2}{2}\right) \left(1 + \frac{n_1}{n_2} x\right)^{(n_1+n_2)/2}}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$



$F(n,m)$ 的密度曲线

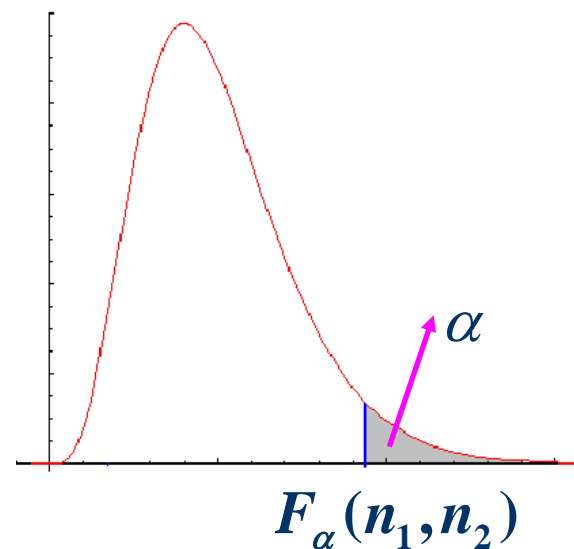
图形随自由度的不同而有所改变

# $F$ 分布

3. 单侧分位点：设  $X \sim F(n_1, n_2)$ ，若对于  $\alpha$ ：  $0 < \alpha < 1$ ，满足

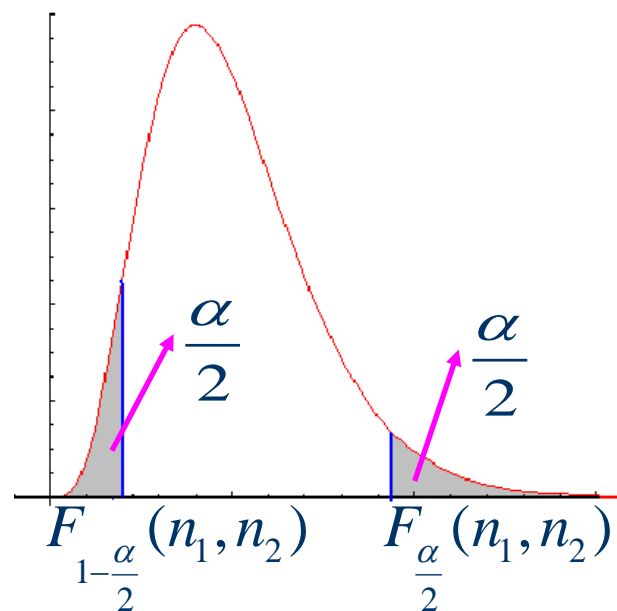
$$P(X > F_{\alpha}(n_1, n_2)) = \int_{F_{\alpha}(n_1, n_2)}^{+\infty} f(x) dx = \alpha$$

的点  $F_{\alpha}(n_1, n_2)$  为  $F$  分布的上侧分位点



双侧分位点：设  $X \sim F(n_1, n_2)$ ，若满足  $P(a < X < b) = 1 - \alpha$

$$a = F_{1-\frac{\alpha}{2}}(n_1, n_2), \quad b = F_{\frac{\alpha}{2}}(n_1, n_2)$$



# $F$ 分布

## 4. 性质:

① 若  $F \sim F(n_1, n_2)$  , 则  $\frac{1}{F} \sim F(n_2, n_1)$

由F分布  
的定义

②  $F_{1-\alpha}(n_1, n_2) = \frac{1}{F_{\alpha}(n_2, n_1)}$

证明: 由  $F \sim F(n_1, n_2)$ , 可得

$$P(F > F_{1-\alpha}(n_1, n_2)) = 1 - \alpha \implies P\left(\frac{1}{F} < \frac{1}{F_{1-\alpha}(n_1, n_2)}\right) = 1 - \alpha$$

$$\implies P\left(\frac{1}{F} > \frac{1}{F_{1-\alpha}(n_1, n_2)}\right) = \alpha$$

$$\text{由 } \frac{1}{F} \sim F(n_2, n_1) \text{ 可得 } P\left(\frac{1}{F} > F_{\alpha}(n_2, n_1)\right) = \alpha$$

得证!



## ***F*** 分布

◆在通常*F*分布表中，只对 $\alpha$ 比较小的值，如 $\alpha = 0.01, 0.05, 0.025$ 及 $0.1$ 等列出了分位点。但有时我们也需要知道 $\alpha$ 比较大的分位点，它们在*F*分布表中查不到。这时我们就可利用分位点的关系式把它们计算出来。

$$F_{1-\alpha}(n_1, n_2) = \frac{1}{F_{\alpha}(n_2, n_1)}$$

例：对 $n_1=12, n_2=9, \alpha=0.95$ ，我们在*F*分布表中查不到 $F_{0.95}(12, 9)$ ，但由上式，知

$$F_{0.95}(12, 9) = \frac{1}{F_{0.05}(9, 12)} = \frac{1}{2.80} = 0.375.$$

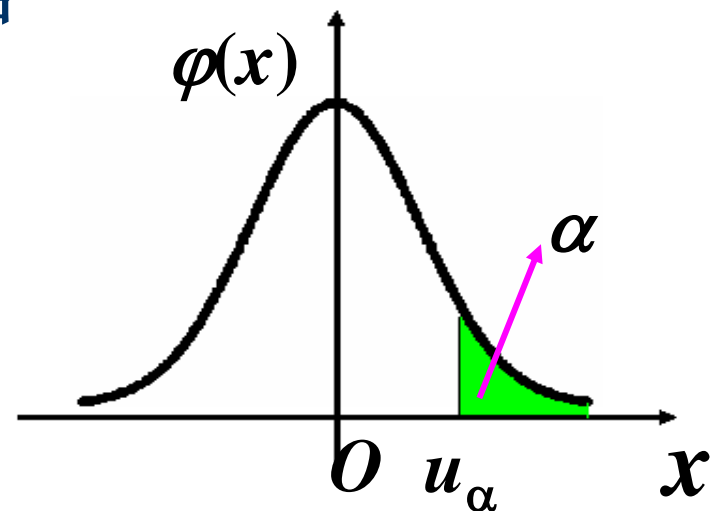
可从*F*分布  
表中查到

# 正态分布的分位点

◆单侧分位点：设 $X \sim N(0, 1)$ ，若  
对于 $\alpha$ ： $0 < \alpha < 1$ ，满足

$$P(X > u_\alpha) = \int_{u_\alpha}^{+\infty} \varphi(x) dx = \alpha$$

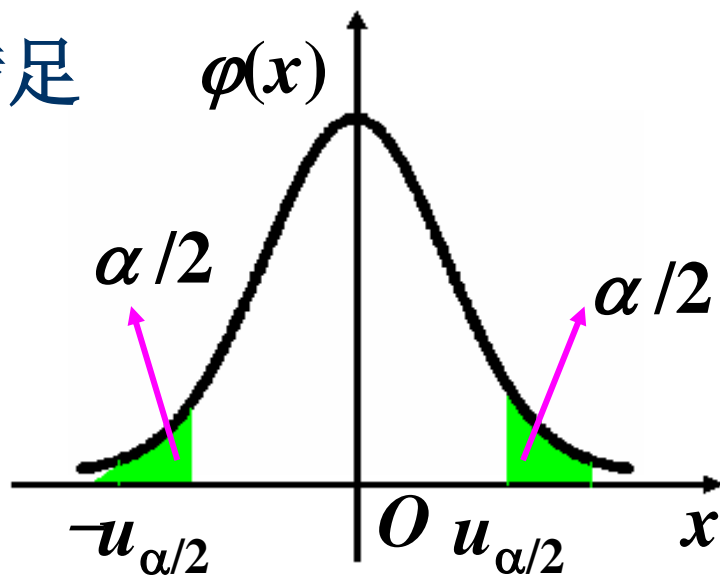
的点  $u_\alpha$  为  $N(0, 1)$  的上侧分位点



双侧分位点：设 $X \sim N(0, 1)$ ，满足

$$P(a < X < b) = 1 - \alpha$$

$$a = -u_{\frac{\alpha}{2}}, \quad b = u_{\frac{\alpha}{2}}$$



**例1:** 设总体 $X \sim N(0,1)$ ,  $X_1, X_2, \dots, X_n$ 为简单随机样本, 试问下列统计量各服从什么分布?

$$(1) \frac{X_1 - X_2}{\sqrt{X_3^2 + X_4^2}}; \quad (2) \frac{\sqrt{n-1}X_1}{\sqrt{\sum_{i=2}^n X_i^2}}; \quad (3) \frac{\left(\frac{n}{3}-1\right)\sum_{i=1}^3 X_i^2}{\sum_{i=4}^n X_i^2}.$$

**解:** (1) 因为 $X_i \sim N(0,1)$ ,  $i=1, 2, \dots, n$ , 所以

$$X_1 - X_2 \sim N(0, 2) \quad \longrightarrow \quad \frac{X_1 - X_2}{\sqrt{2}} \sim N(0,1)$$

$$X_3^2 + X_4^2 \sim \chi^2(2)$$

$$\frac{X_1 - X_2}{\sqrt{X_3^2 + X_4^2}} = \frac{(X_1 - X_2)/\sqrt{2}}{\sqrt{\frac{X_3^2 + X_4^2}{2}}} \sim t(2)$$

**例1:** 设总体 $X \sim N(0,1)$ ,  $X_1, X_2, \dots, X_n$ 为简单随机样本, 试问下列统计量各服从什么分布?

$$(1) \frac{X_1 - X_2}{\sqrt{X_3^2 + X_4^2}}; \quad (2) \frac{\sqrt{n-1}X_1}{\sqrt{\sum_{i=2}^n X_i^2}}; \quad (3) \frac{\left(\frac{n}{3}-1\right)\sum_{i=1}^3 X_i^2}{\sum_{i=4}^n X_i^2}.$$

**解:** (2) 因为 $X_i \sim N(0,1)$ , 所以

$$\sum_{i=2}^n X_i^2 \sim \chi^2(n-1)$$

$$\frac{\sqrt{n-1}X_1}{\sqrt{\sum_{i=2}^n X_i^2}} = \frac{X_1}{\sqrt{\sum_{i=2}^n X_i^2 / (n-1)}} \sim t(n-1)$$

**例1:** 设总体 $X \sim N(0,1)$ ,  $X_1, X_2, \dots, X_n$ 为简单随机样本, 试问下列统计量各服从什么分布?

$$(1) \frac{X_1 - X_2}{\sqrt{X_3^2 + X_4^2}}; \quad (2) \frac{\sqrt{n-1}X_1}{\sqrt{\sum_{i=2}^n X_i^2}}; \quad (3) \frac{\left(\frac{n}{3}-1\right)\sum_{i=1}^3 X_i^2}{\sum_{i=4}^n X_i^2}.$$

**解:** (3) 因为

$$\sum_{i=1}^3 X_i^2 \sim \chi^2(3), \quad \sum_{i=4}^n X_i^2 \sim \chi^2(n-3)$$

所以

$$\frac{\left(\frac{n}{3}-1\right)\sum_{i=1}^3 X_i^2}{\sum_{i=4}^n X_i^2} = \frac{\sum_{i=1}^3 X_i^2 / 3}{\sum_{i=4}^n X_i^2 / (n-3)} \sim F(3, n-3)$$

**例2:** 若 $T \sim t(n)$ , 问 $T^2$ 服从什么分布?

**解:** 因为 $T \sim t(n)$ , 可以认为

$$T = \frac{U}{\sqrt{V/n}},$$

其中 $V \sim \chi^2(n)$ ,  $U \sim N(0,1)$ ,  $U^2 \sim \chi^2(1)$

$$T^2 = \frac{U^2/1}{V/n} \sim F(1, n)$$

# 一个正态总体下的三个统计量的分布

设总体为 $X \sim N(\mu, \sigma^2)$ ，样本为 $X_1, \dots, X_n$ ，

样本均值和样本方差为：

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

# 正态总体下的统计量的分布

设总体为 $X \sim N(\mu, \sigma^2)$ , 样本为 $X_1, \dots, X_n$

◆ 定理6.1  $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

$\mu$ 和 $\sigma^2$ 已

知

证明:  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$

$$E(\bar{X}) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \mu$$

是 $n$ 个独立的正态随机变量的线性组合, 故服从正态分布

$$D(\bar{X}) = \frac{1}{n^2} \sum_{i=1}^n D(X_i) = \frac{\sigma^2}{n}$$

$$\therefore \bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$

标准化得  $\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0, 1)$



# 正态总体下的统计量的分布

设总体为 $X \sim N(\mu, \sigma^2)$ ，样本为 $X_1, \dots, X_n$

◆ 定理6.2  $\bar{X}$ 与 $S^2$ 相互独立;

$\mu$  未知  
 $\sigma^2$  已知

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} \sim \chi^2(n-1);$$

上式的自由度为什么是 $n-1$ ?

从表面上看， $\sum_{i=1}^n (X_i - \bar{X})^2$ 是 $n$ 个正态随机变量的平方和，但实际上它们不是独立的，它们之间有一种线性约束关系：

$$\sum_{i=1}^n (X_i - \bar{X}) = \left( \sum_{i=1}^n X_i \right) - n\bar{X} = 0$$

这表明：当这 $n$ 个正态随机变量中有 $n-1$ 个取值给定时，剩下的一个的取值就跟着唯一确定了，故在这 $n$ 项平方和中只有 $n-1$ 项是独立的，所以上式的自由度是 $n-1$ .

# 正态总体下的统计量的分布

设总体为 $X \sim N(\mu, \sigma^2)$ , 样本为 $X_1, \dots, X_n$

◆ 定理6.2  $\bar{X}$ 与 $S^2$ 相互独立;

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} \sim \chi^2(n-1);$$

$\mu$ 未知  
 $\sigma^2$ 已知

对比:

$$\sum_{i=1}^n \left( \frac{X_i - \mu}{\sigma} \right)^2 = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} \sim \chi^2(n)$$

$\mu$ 和 $\sigma^2$ 已知

# 正态总体下的统计量的分布

设总体为 $X \sim N(\mu, \sigma^2)$ , 样本为 $X_1, \dots, X_n$

◆ 定理6.3

$$\frac{\bar{X} - \mu}{S / \sqrt{n}} \sim t(n-1) \quad \cdot \quad \cdot \quad \cdot$$

$\sigma^2$ 未知  
 $\mu$ 已知

证明：由定理6.1和定理6.2,

$$U = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0, 1), \quad V = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

且U与V独立,根据t分布的构造

$$\frac{\frac{\bar{X} - \mu}{\sigma / \sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2} / (n-1)}} = \frac{\bar{X} - \mu}{S / \sqrt{n}} \sim t(n-1)$$

得证!

# 正态总体下的统计量的分布

设总体为 $X \sim N(\mu, \sigma^2)$ , 样本为 $X_1, \dots, X_n$

◆ 定理6.3

$$\frac{\bar{X} - \mu}{S / \sqrt{n}} \sim t(n-1).$$

$\sigma^2$ 未知  
 $\mu$ 已知

对比:

$$\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0, 1).$$

$\mu$ 和 $\sigma^2$ 已  
知

# 一个正态总体下的统计量的分布

设总体为 $X \sim N(\mu, \sigma^2)$ , 样本为 $X_1, \dots, X_n$

◆ 结论: 
$$\frac{\bar{X} - \mu}{S / \sqrt{n}} \sim t(n-1) \leftrightarrow \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0, 1)$$

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} \sim \chi^2(n-1)$$

$$\sum_{i=1}^n \left( \frac{X_i - \mu}{\sigma} \right)^2 = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} \sim \chi^2(n)$$



**例1:** 设总体 $X \sim N(\mu, 4^2)$ ,  $X_1, X_2, \dots, X_{10}$ 是 $n=10$ 简单随机样本,  $S^2$ 为样本方差, 已知 $P(S^2 > \alpha) = 0.1$ , 求 $\alpha$ .

$$\text{求 } P(S^2 > \alpha) = 0.1$$

?

**解:** 因为 $n=10$ ,  $n-1=9$ ,  $\sigma^2=4^2$ , 所以  $\frac{9S^2}{4^2} \sim \chi^2(9)$

$$P(S^2 > \alpha) = P\left(\frac{9S^2}{4^2} > \frac{9\alpha}{4^2}\right) = 0.1$$

$$\text{所以 } \frac{9\alpha}{4^2} = \chi_{0.1}^2(9) \approx 14.684$$

$$\alpha \approx 14.684 \times \frac{16}{9} = 26.105$$

查表

**例2:** 在设计导弹发射装置时，重要内容之一是研究弹着点偏离目标中心的距离的方差。对于某类导弹发射装置，弹着点偏离目标中心的距离服从  $N(\mu, \sigma^2)$ ，这里  $\sigma^2 = 100\text{米}^2$ 。现在进行了25次发射试验，用  $S^2$  记这25次试验中弹着点偏离目标中心的距离的样本方差。求： $S^2$  超过50米<sup>2</sup>的概率。

**解：**根据定理6.2，知

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2 \quad \text{其中 } n = 25, \sigma^2 = 100$$

$$\begin{aligned} \text{所以 } P(S^2 > 50) &= P\left\{\frac{(n-1)S^2}{\sigma^2} > \frac{(n-1) \times 50}{\sigma^2}\right\} \\ &= P\left\{\chi_{25-1}^2 > \frac{24 \times 50}{100}\right\} = P(\chi_{24}^2 > 12) \end{aligned}$$

查P<sub>215</sub>附表4, 得到:

$$P(S^2 > 50) = P(\chi_{24}^2 > 12) > 0.975$$



## 两个正态总体下的三个统计量的分布

➤ 设总体为  $X \sim N(\mu_1, \sigma_1^2)$ ，有样本为  $X_1, X_2, \dots, X_{n_1}$ ，

样本均值和样本方差为：

$$\bar{X} = \frac{1}{n_1} \sum_{i=1}^{n_1} X_i \quad S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_i - \bar{X})^2$$

➤ 设总体为  $Y \sim N(\mu_2, \sigma_2^2)$ ，有样本为  $Y_1, Y_2, \dots, Y_{n_2}$ ，

样本均值和样本方差为：

$$\bar{Y} = \frac{1}{n_2} \sum_{i=1}^{n_2} Y_i \quad S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (Y_i - \bar{Y})^2$$

两个样本相互独立

# 两个正态总体下的统计量的分布

## ◆ 定理6.4

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

两个正态  
总体结论

$$\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$$

一个正态  
总体结论

# 两个正态总体下的统计量的分布

◆ 定理6.4 
$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

证明：由定理6.1，知

$$\bar{X} \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right), \quad \bar{Y} \sim N\left(\mu_2, \frac{\sigma_2^2}{n_2}\right)$$

由两样本相互独立， $\bar{X}$ 与 $\bar{Y}$ 也相互独立

$$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

# 两个正态总体下的统计量的分布

◆ 定理6.5: 当 $\sigma_1^2$ ,  $\sigma_2^2$ 未知, 但两者相等时,

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_{\omega} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$

$$\text{其中 } S_{\omega}^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

称为混合样本方差

# 两个正态总体下的统计量的分布

◆ 定理6.5的证明：由定理6.2可得

$$\frac{(n_1 - 1)S_1^2}{\sigma^2} \sim \chi_{n_1 - 1}^2, \quad \frac{(n_2 - 1)S_2^2}{\sigma^2} \sim \chi_{n_2 - 1}^2$$

且二者相互独立，由 $\chi^2$ 分布的可加性：

$$\frac{(n_1 - 1)S_1^2}{\sigma^2} + \frac{(n_2 - 1)S_2^2}{\sigma^2} \sim \chi_{n_1 + n_2 - 2}^2$$

由定理6.4及t分布的构造，可得最后结论.

# 两个正态总体下的统计量的分布

◆ 定理6.6: 
$$\frac{\frac{(n_1-1)S_1^2}{\sigma_1^2} / (n_1-1)}{\frac{(n_2-1)S_2^2}{\sigma_2^2} / (n_2-1)} = \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F(n_1-1, n_2-1)$$

证明: 由定理6.2

$$\frac{(n_1-1)S_1^2}{\sigma_1^2} \sim \chi^2(n_1-1) \quad \frac{(n_2-1)S_2^2}{\sigma_2^2} \sim \chi^2(n_2-1)$$

由F分布的构造

$$\frac{\frac{(n_1-1)S_1^2}{\sigma_1^2} / (n_1-1)}{\frac{(n_2-1)S_2^2}{\sigma_2^2} / (n_2-1)} = \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F(n_1-1, n_2-1) \quad \text{得证!}$$

**例：**从总体 $X \sim N(\mu, 3)$ ,  $Y \sim N(\mu, 5)$ 中分别抽取 $n_1=10$ ,  $n_2=15$ 的两独立样本, 求两个样本方差之比大于1.272的概率.

**解：**因为  $\frac{S_1^2}{S_2^2} \cdot \frac{5}{3} \sim F(9, 14)$

$$\begin{aligned} \text{所以 } P\left(\frac{S_1^2}{S_2^2} > 1.272\right) &= P\left(\frac{S_1^2}{S_2^2} \cdot \frac{5}{3} > 1.272 \times \frac{5}{3}\right) \\ &= P\left(\frac{S_1^2}{S_2^2} \cdot \frac{5}{3} > 2.12\right) \end{aligned}$$

由F分布表,  $F_{0.1}(9, 14) = 2.12$ ,

$$\text{即 } P\left(\frac{S_1^2}{S_2^2} > 1.272\right) = 0.1.$$

## 小 结

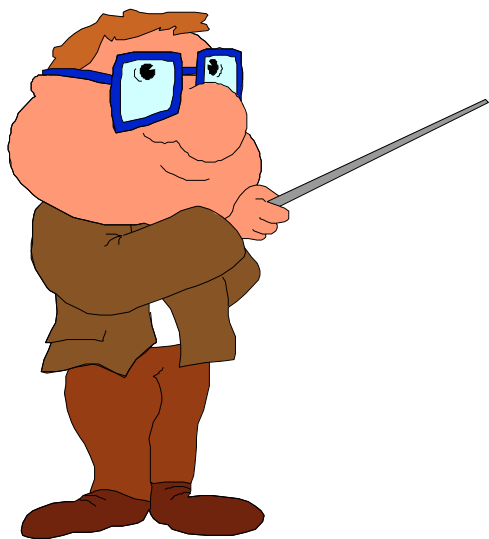
本讲首先介绍数理统计中三个常用的重要统计量的分布： $\chi^2$ 分布、 $t$ 分布和  $F$  分布；然后以定理的形式 给出了正态总体样本均值与样本方差的分布及其相关结论。



# 第七章 参数估计

- ◆ 参数的点估计
- ◆ 估计量的评选标准
- ◆ 参数的区间估计

## 数理统计的任务：



- 总体分布类型的判断；
- 总体分布中未知参数的推断：

参数估计 与 假设检验

# 参数估计

参数估计问题是利用从总体抽样得到的信息来估计总体的某些参数或者参数的某些函数.

参数估计问题的一般提法:

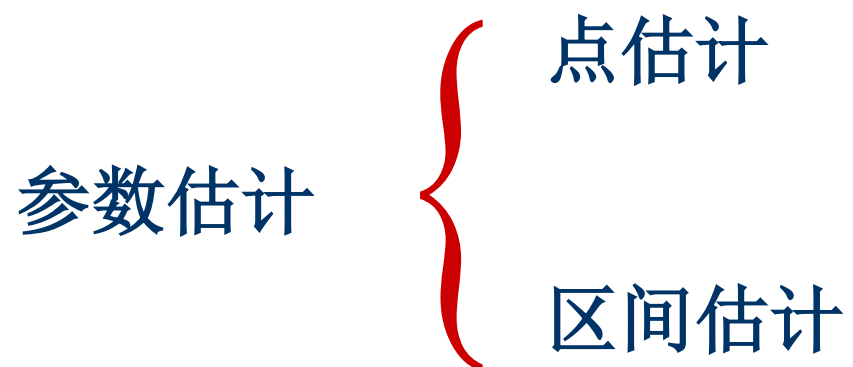
设总体  $X$  的分布函数为  $F(x, \theta)$ , 其中  $\theta$  为未知参数或参数向量, 现从该总体中抽样, 得到样本

$$X_1, X_2, \dots, X_n$$

依据样本对参数  $\theta$  做出估计, 或估计参数  $\theta$  的某个已知函数  $g(\theta)$ 。

这类问题称为参数估计

# 参数估计



# 参数估计

◆ 例：我们要估计某班男生的平均身高。

假定身高服从正态分布  $N(\mu, 0.1^2)$

现从该总体选取容量为5的样本，我们的任务是要根据选出的样本(5个数)求出总体均值  $\mu$  的估计，而全部信息就由这5个数组成。设这5个数是：

1.65, 1.67, 1.75, 1.78, 1.69

估计  $\mu$  为1.73，这是点估计。

估计  $\mu$  在区间 [1.65, 1.74] 内，这是区间估计。

# 参数的点估计

**例：** 已知某地区新生婴儿的体重 $X \sim N(\mu, \sigma^2)$   
( $\mu, \sigma$  未知)



...



随机抽查100个婴儿，得100个体重数据

10, 7, 6, 6.5, 5, 5.2, ...

而全部信息就由这100个数组成。

据此,我们应如何估计 $\mu$  和  $\sigma$  呢?

# 参数的点估计

为估计  $\mu$ ：

我们需要构造出适当的统计量  $T(X_1, X_2, \dots, X_n)$ 。

一旦当有了样本，就将样本值代入到该统计量中，算出一个值，用来作为  $\mu$  的估计值。

$T(X_1, X_2, \dots, X_n)$  称为参数  $\mu$  的点估计量，把样本值代入  $T(X_1, X_2, \dots, X_n)$  中，得到  $\mu$  的一个点估计值。

# 参数的点估计

✚ 寻求估计量的方法

1. 矩估计法

2. 最大似然法

3. 最小二乘法

4. 贝叶斯方法 .....

这里我们主要介绍前面两种方法



# 矩估计

矩估计是基于“**替换**”思想建立起来的一种参数估计方法。最早由英国统计学家 **K.皮尔逊** 提出。



**原理：**用同阶、同类的**样本矩**来估计**总体矩**

**基本思想：**由于样本来源于总体，样本矩在一定程度上反映了总体矩，而且由大数定理

$$\lim_{n \rightarrow \infty} P \left( \left| \frac{1}{n} \sum_{i=1}^n X_i^k - E(X^k) \right| > \varepsilon \right) = 0 \quad \forall \varepsilon > 0, k = 1, 2, \dots$$

可知：样本矩依概率收敛于总体矩。因此只要总体X的k阶原点矩存在，就可以用样本矩作为总体矩的估计量，而且这种估计可以到达任意精确的程度。

# 矩估计

□ 矩估计就是用相应的样本矩去估计总体矩

➤ 总体  $k$  阶原点矩:  $a_k = E(X^k)$

➤ 样本  $k$  阶原点矩:  $A_k = \frac{1}{n} \sum_{i=1}^n X_i^k$

# 矩估计

## ◆ 矩估计的一般步骤:

- ① 设总体  $X$  的分布函数中含  $k$  个未知参数  $\theta_1, \theta_2, \dots, \theta_k$ , 且总体  $X$  的  $k$  阶原点矩  $E(X^k)$  都存在, 则可建立如下方程组:

$$\begin{cases} \frac{1}{n} \sum_{i=1}^n X_i = E(X) \\ \frac{1}{n} \sum_{i=1}^n X_i^2 = E(X^2) \\ \dots\dots\dots \\ \frac{1}{n} \sum_{i=1}^n X_i^k = E(X^k) \end{cases}$$

等式左边  $X_1, X_2, \dots, X_n$  是样本, 等式右边的  $E(X^k)$  是  $\theta_1, \theta_2, \dots, \theta_k$  的函数

得到了关于  $\theta_1, \theta_2, \dots, \theta_k$  的  $k$  个方程

# 矩估计

◆ 矩估计的一般步骤:

②解方程组

$$\begin{cases} \hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \dots, X_n) \\ \hat{\theta}_2 = \hat{\theta}_2(X_1, X_2, \dots, X_n) \\ \dots\dots\dots \\ \hat{\theta}_k = \hat{\theta}_k(X_1, X_2, \dots, X_n) \end{cases}$$

它们就是 $\theta_1, \theta_2, \dots, \theta_k$ 的矩估计量，是随机变量；

将样本用观测值 $x_1, x_2, \dots, x_n$ 代人,得到 $\theta_1, \theta_2, \dots, \theta_k$ 的矩估计值

# 矩估计的计算

**例1:** 设总体  $X$  的概率密度为

$$f(x) = \begin{cases} (\alpha + 1)x^\alpha, & 0 < x < 1 \\ 0, & \text{其他} \end{cases}$$

其中  $\alpha > -1$  是未知参数, 求  $\alpha$  的矩估计.

**解:** 先求总体的期望  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = E(X)$

$$\begin{aligned} E(X) &= \int_0^1 x \cdot (\alpha + 1)x^\alpha dx \\ &= (\alpha + 1) \int_0^1 x^{\alpha+1} dx \\ &= \frac{\alpha + 1}{\alpha + 2} \end{aligned}$$

# 矩估计的计算

由矩估计法，令

$$\bar{X} = \frac{\alpha + 1}{\alpha + 2}$$

样本矩

总体矩

解得：

$$\hat{\alpha} = \frac{2\bar{X} - 1}{1 - \bar{X}}$$

为 $\alpha$  的矩估计量。

**注意：**要在参数上边加上“ $\wedge$ ”，表示参数的估计。它是统计量。

# 矩估计的计算

**例2:** 设  $X_1, X_2, \dots, X_n$  是取自总体  $X$  的简单样本,  $X$  有概率密度函数

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-(x-\mu)/\theta}, & x \geq \mu \\ 0, & \text{其他} \end{cases} \quad \begin{cases} \frac{1}{n} \sum_{i=1}^n X_i = E(X) \\ \frac{1}{n} \sum_{i=1}^n X_i^2 = E(X^2) \end{cases}$$

其中  $\mu, \theta$  为未知参数, 且  $\theta > 0$ , 求  $\mu$  和  $\theta$  的矩估计

**解:** 
$$\begin{cases} E(X) = \int_{\mu}^{+\infty} x \cdot \frac{1}{\theta} e^{-(x-\mu)/\theta} dx = \theta + \mu \\ E(X^2) = \int_{\mu}^{+\infty} x^2 \cdot \frac{1}{\theta} e^{-(x-\mu)/\theta} dx = \theta^2 + (\theta + \mu)^2 \end{cases}$$

令 
$$\begin{cases} \bar{X} = \mu + \theta \\ \frac{1}{n} \sum_{i=1}^n X_i^2 = \theta^2 + (\mu + \theta)^2 \end{cases} \quad \longrightarrow \quad \begin{cases} \hat{\theta} = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2} \\ \hat{\mu} = \bar{X} - \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2} \end{cases}$$

## 矩估计的计算

**例3:** 设总体 $X$ 的均值为 $\mu$ , 方差为 $\sigma^2$ , 求 $\mu$  和  $\sigma^2$  的矩估计

**解:** 由 
$$\begin{cases} E(X) = \mu \\ D(X) = E(X^2) - [E(X)]^2 = \sigma^2 \end{cases}$$

列出方程组: 
$$\begin{cases} \bar{X} = \mu \\ \frac{1}{n} \sum_{i=1}^n X_i^2 = \mu^2 + \sigma^2 \end{cases}$$

故均值 $\mu$ , 方差 $\sigma^2$ 的矩估计为

$$\begin{cases} \hat{\mu} = \bar{X}, \\ \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \stackrel{\text{即}}{=} \frac{n-1}{n} S^2. \end{cases}$$



# 矩估计的计算

◆ 例3的特殊情况:

□ 正态总体  $N(\mu, \sigma^2)$  中  $\mu$  和  $\sigma^2$  的矩估计为

$$\begin{cases} \hat{\mu} = \bar{X} \\ \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \end{cases}$$

□ 若总体  $X \sim U(a, b)$ , 则  $a, b$  的矩估计为

$$\begin{cases} \bar{X} = \frac{a+b}{2} \\ \sigma^2 = \frac{(b-a)^2}{12} \end{cases} \Rightarrow \begin{cases} \hat{a} = \bar{X} - \sqrt{3}\sigma \\ \hat{b} = \bar{X} + \sqrt{3}\sigma \end{cases}$$

# 矩估计

- ◆ 矩估计法的优点：直观简便，不必知道总体的分布类型，只需总体的 $k$ 阶原点矩存在即可。
- ◆ 矩估计法的缺点：总体的 $k$ 阶原点矩不一定存在；当总体类型已知时，没有充分利用总体分布提供的信息。因此它的估计精度可能要比别的估计法低。

# 极大似然估计

极大似然估计法是在总体的分布类型已知前提下，使用的一种参数估计法。

它首先是由德国数学家**高斯**在1821年提出的。

然而，这个方法常归功于英国统计学家**费歇**。



Gauss



Fisher

**费歇**在1922年重新发现了这一方法，研究了这种方法的一些性质，并给出了求参数极大似然估计一般方法——**极大似然估计原理**。

# 极大似然估计

## ◆ 极大似然估计法的基本思想

设总体的分布函数形式 $F(x, \theta)$ 已知，但含有未知参数 $\theta$ ； $\theta$ 可以取很多值。在 $\theta$ 的一切可能取值中选一个使样本观测值出现的概率最大的 $\theta$ 值，作为 $\theta$ 的估计值。

# 极大似然估计思想举例

## ➤ 黑箱问题:

设一个黑箱子里方有黑球和白球，比例为3:1，用实验的方法判断哪个颜色的球多。

➤ 分析：已知任取一个，抽到黑球和白球的概率分别为  $\frac{3}{4}$  和  $\frac{1}{4}$ 。

➤ 现假设有放回的抽取  $n=3$  次，则得到黑球的数量  $X \sim B(n, p)$ 。

用这个实验来估计  $p = \frac{3}{4}$  还是  $p = \frac{1}{4}$

	分布形式		0	1	2	3
总体一	$B(3, 1/4)$	$P(X=x_i)$	27/64	27/64	9/64	1/64
总体二	$B(3, 3/4)$	$P(X=x_i)$	1/64	9/64	27/64	27/64

	分布形式	黑球数量	0	1	2	3
总体一	$B(3, 1/4)$	$P(X=x_i)$	$27/64$	$27/64$	$9/64$	$1/64$
总体二	$B(3, 3/4)$	$P(X=x_i)$	$1/64$	$9/64$	$27/64$	$27/64$

可以看出，试验（样本）结果如果三次取到0个黑球，此时使得“ $X=0$ ”这个结果发生的概率在 $p=1/4$ 时候达到最大，所以认为 $p=1/4$ 。

对于每个 $x$ 的值，就是要选一个参数 $p$ 的值( $1/4$ 或者 $3/4$ ),使得 $P(X=x_i)$ 的概率达到最大。

# 极大似然估计原理

➤ 假设总体 $X$ 为连续型随机变量

设总体  $X$  的分布为  $f(x, \theta)$ ,  $X_1, X_2, \dots, X_n$  是来自总体  $X$  的简单样本。则样本的联合概率函数为

$$L(\theta) \overset{\Delta}{=} L(\underbrace{x_1, x_2, \dots, x_n}_{\text{视为变量}}, \underbrace{\theta}_{\text{被看作固定, 但未知的参数}}) = \prod_{i=1}^n f(x_i, \theta)$$

视为变量

被看作固定,  
但未知的参数

称  $L(\theta)$  为  $\theta$  的似然函数, 这里  $x_1, x_2, \dots, x_n$  是样本的观测值

# 极大似然估计原理

似然函数:

$$L(\theta) = \prod_{i=1}^n f(x_i, \theta)$$

$L(\theta)$  看作参数  $\theta$  的函数, 它可作为  $\theta$  将以多大可能产生样本值  $x_1, x_2, \dots, x_n$  的一种度量.

最大似然估计法就是用使  $L(\theta)$  达到最大值的  $\hat{\theta}$  去估计  $\theta$ . 即

$$L(\hat{\theta}) = \max_{\theta \in \Theta} L(\theta)$$

$\theta$  可能变化空间,  
称为参数空间。

称  $\theta(x_1, \dots, x_n)$  为  $\theta$  的最大似然估计值(MLE).

相应的统计量  $\theta(X_1, \dots, X_n)$  称为  $\theta$  的最大似然估计量.



# 极大似然估计

## ◆ 似然函数

① 当总体 $\mathbf{X}$ 为离散型随机变量时，若有

$$P(X = x) = p(x, \theta)$$

则似然函数为

$$\begin{aligned} L(\theta) &= P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) \\ &= \prod_{i=1}^n P(X_i = x_i) = \prod_{i=1}^n p(x_i, \theta) \end{aligned}$$

② 当总体 $\mathbf{X}$ 为连续型随机变量时，若概率密度为 $f(x, \theta)$ ,

则似然函数为

$$L(\theta) = \prod_{i=1}^n f(x_i, \theta)$$

# 极大似然估计

◆ 求极大似然估计的一般步骤为：

1. 求似然函数  $L(\theta) = L(x_1, x_2, \dots, x_n; \theta)$

2. 一般地，求出  $\ln L(\theta)$  及似然方程

$$\frac{d \ln L(\theta)}{d \theta} = 0$$

3. 解似然方程得到最大似然估计值

$$\hat{\theta} = \hat{\theta}(x_1, x_2, \dots, x_n)$$

4. 最后得到最大似然估计量

$$\hat{\theta} = \hat{\theta}(X_1, X_2, \dots, X_n)$$

# 极大似然估计

◆ 三点说明:

➤ 求似然函数  $L(\theta)$  的最大值点, 可应用微积分中的技巧。由于  $\ln(x)$  是  $x$  的增函数, 所以  $\ln L(\theta)$  与  $L(\theta)$  在  $\theta$  的同一点处达到各自的最大值。

➤ 若  $\theta = (\theta_1, \theta_2, \dots, \theta_k)$  是向量, 上述似然方程需用似然方程组

$$\begin{cases} \frac{\partial \ln L(\theta_1, \theta_2, \dots, \theta_k)}{\partial \theta_1} = 0 \\ \dots\dots\dots \\ \frac{\partial \ln L(\theta_1, \theta_2, \dots, \theta_k)}{\partial \theta_k} = 0 \end{cases}$$

➤ 用上述方法求参数的极大似然估计有时行不通, 这时要用极大似然原理来求。

# 极大似然估计的计算

**例4:** 设 $X_1, X_2, \dots, X_n$ 是取自总体  $X \sim B(1, p)$  的一个样本, 求参数  $p$  的极大似然估计。

**解:** 似然函数为:

$$P(X = x) = p^x (1 - p)^{1-x}$$

$$\begin{aligned} L(p) &= \prod_{i=1}^n P(X_i = x_i) = \prod_{i=1}^n p^{x_i} (1 - p)^{1-x_i} \\ &= p^{\sum_{i=1}^n x_i} (1 - p)^{n - \sum_{i=1}^n x_i} \end{aligned}$$

对数似然函数为

$$\ln L(p) = \sum_{i=1}^n x_i \ln(p) + \left( n - \sum_{i=1}^n x_i \right) \ln(1 - p)$$

# 极大似然估计的计算

$$\ln L(p) = \sum_{i=1}^n x_i \ln(p) + \left( n - \sum_{i=1}^n x_i \right) \ln(1-p)$$

对  $p$  求导，并令其等于零，得

$$\frac{d \ln L(p)}{dp} = \frac{1}{p} \sum_{i=1}^n x_i - \frac{1}{1-p} (n - \sum_{i=1}^n x_i) = 0.$$

上式等价于

$$\frac{\bar{x}}{p} = \frac{1 - \bar{x}}{1 - p}$$

$$n \bar{x}$$

解得

注意  $\hat{p}$

$$\hat{p} = \bar{x}$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

极大似然估计值为  $\hat{p} = \bar{x}$

极大似然估计量为  $\hat{p} = \bar{X}$

# 极大似然估计的计算

**例4(续):** 设 $X_1, X_2, \dots, X_n$ 是取自总体  $X \sim B(1, p)$  的一个样本, 求参数  $p$  的矩估计。

**解:** 
$$\frac{1}{n} \sum_{i=1}^n X_i = E(X) = p$$

$\therefore$  矩估计量为 
$$\hat{p} = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}$$

矩估计值为 
$$\hat{p} = \bar{x}$$

# 极大似然估计的计算

**例5:** 设总体  $X$  服从泊松分布  $P(\lambda)$ , 求参数  $\lambda$  的极大似然估计

**解:** 由  $X$  的概率分布函数为

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

得到  $\lambda$  的似然函数为

$$L(\lambda) = \prod_{i=1}^n P(X_i = x_i) = \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} = e^{-n\lambda} \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!},$$

对数似然函数为

$$\ln L(\lambda) = -n\lambda + (\ln \lambda) \sum_{i=1}^n x_i - \sum_{i=1}^n \ln(x_i!)$$

# 极大似然估计的计算


$$\ln L(\lambda) = -n\lambda + (\ln \lambda) \sum_{i=1}^n x_i - \sum_{i=1}^n \ln(x_i !)$$

似然方程为

$$\frac{d}{d\lambda} \ln L(\lambda) = -n + \frac{1}{\lambda} \sum_{i=1}^n x_i = 0$$

其解为

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$$



极大似然  
估计值



# 极大似然估计的计算

**例6：**求正态总体  $N(\mu, \sigma^2)$  参数  $\mu$  和  $\sigma^2$  的极大似然估计(注：我们把  $\sigma^2$  看作一个参数)

**解：**似然函数为

$$\begin{aligned} L(\mu, \sigma^2) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}} \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2} \end{aligned}$$

对数似然函数为

$$\ln L(\mu, \sigma^2) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

# 极大似然估计的计算

$$\ln L(\mu, \sigma^2) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2,$$

似然方程组为

$$\begin{cases} \frac{\partial}{\partial \mu} \ln L(\mu, \sigma^2) = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) = 0, \\ \frac{\partial}{\partial \sigma^2} \ln L(\mu, \sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - \mu)^2 = 0. \end{cases}$$

解得

$$\begin{cases} \hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x} \\ \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \end{cases}$$

极大似然  
估计值

# 极大似然估计的计算

**例 7:** 设  $X \sim U[a, b]$ , 求  $a, b$  的极大似然估计

**解:** 因

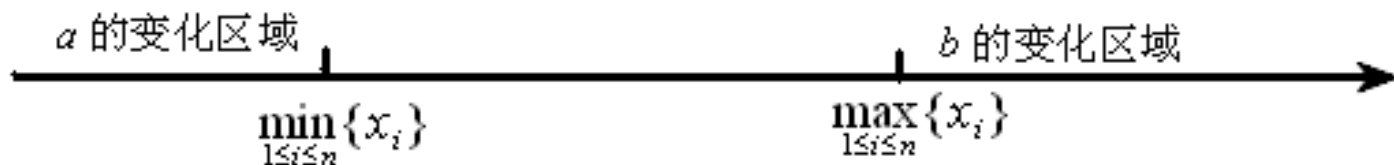
$$f(x, a, b) = \begin{cases} \frac{1}{b-a}, & x \in [a, b] \\ 0, & x \notin [a, b] \end{cases}$$

所以 
$$L(a, b) = \prod_{i=1}^n f(x_i, a, b)$$

$$= \begin{cases} \frac{1}{(b-a)^n}, & x_i \in [a, b], i = 1, 2, \dots, n \\ 0, & \text{其他} \end{cases}$$

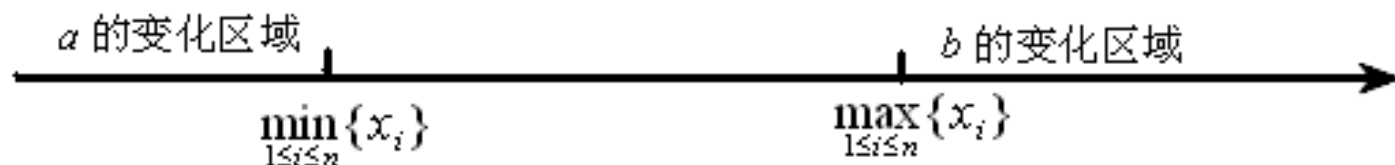
$$L(a,b) = \begin{cases} \frac{1}{(b-a)^n}, & x_i \in [a,b], \quad i = 1, 2, \dots, n, \\ 0, & \text{其他} \end{cases}$$

$$= \begin{cases} \frac{1}{(b-a)^n}, & a \leq \min_{1 \leq i \leq n} \{x_i\} \leq \max_{1 \leq i \leq n} \{x_i\} \leq b, \\ 0, & \text{其他.} \end{cases}$$



由上式看到： $L(a,b)$ 作为 $a$ 和 $b$ 的二元函数是不连续的，所以我们不能用似然方程组来求极大似然估计，而必须从极大似然估计的定义出发，求 $L(a,b)$ 的最大值。

$$L(a,b) = \begin{cases} \frac{1}{(b-a)^n}, & a \leq \min_{1 \leq i \leq n} \{x_i\} \leq \max_{1 \leq i \leq n} \{x_i\} \leq b, \\ 0, & \text{其他} . \end{cases}$$



为使 $L(a,b)$  达到最大,  $b-a$  应该尽量地小; 但  $b$  不能小于  $\max\{x_1, x_2, \dots, x_n\}$ 。否则,  $L(a,b) = 0$ 。

类似地,  $a$  不能大于  $\min\{x_1, x_2, \dots, x_n\}$ 。

因此,  $a$  和  $b$  的极大似然估计为

$$\hat{a} = \min_{1 \leq i \leq n} \{x_i\}, \quad \hat{b} = \max_{1 \leq i \leq n} \{x_i\}.$$

# 小 结

本讲首先介绍参数矩估计的基本思想以及求矩估计的步骤，给出多个求参数矩估计的例子；然后介绍参数极大似然估计的基本原理，求极大似然估计的基本方法，给出多个求参数极大似然矩估计的例子。

点估计

```
graph TD; A([点估计]) --> B([矩估计法]); A --> C([极大似然估计法]); B --> D([基本步骤]); C --> E([基本步骤]);
```

A flowchart illustrating the classification of Point Estimation (点估计) methods. The root node is '点估计' (Point Estimation). It branches into two main methods: '矩估计法' (Method of Moments) and '极大似然估计法' (Maximum Likelihood Estimation). Both methods lead to their respective '基本步骤' (Basic Steps).

矩估计法

极大似然估计法

基本步骤

基本步骤

## 第二节 估计量的评选标准

● 无偏性

● 有效性

● 一致性



# 估计量的评选标准

## ◆ 问题的提出

从前面可以看到，对于同一个参数，用不同的估计方法求出的估计量可能不相同；而且很明显，原则上任何统计量都可以作为未知参数的估计量。

这就需要讨论以下问题：

- (1) 对于同一个参数究竟采用哪一个估计量好？
- (2) 评价估计量的标准是什么？

# 估计量的评选标准

常用的几条标准是：

- ① 无偏性
- ② 有效性
- ③ 一致性(相合性)

# 无偏性

## ◆ 定义

设总体 $X$ 的分布参数为 $\theta$ , 设 $(X_1, X_2, \dots, X_n)$ 是来自总体的一个样本,

$$\hat{\theta} = \hat{\theta}(X_1, X_2, \dots, X_n)$$

是 $\theta$ 的一个估计量(注意:它是一个统计量, 是随机变量), 若

$$E(\hat{\theta}) = \theta$$

对一切可能的 $\theta$ 成立, 则称 $\hat{\theta}$ 为 $\theta$ 的一个无偏估计量。一个估计量如果不是无偏的, 就称它是有偏估计量。

◆ 无偏估计的**实际意义**: 保证了无系统误差, 估计时不会始终偏大或者偏小。

## 无偏性的计算

**例：** 设 $X_1, X_2, \dots, X_n$ 为来自均值为 $\mu$ 的总体 $X$ 的随机样本，考虑 $\mu$ 的如下几个估计量是不是无偏估计量？

$$\hat{\mu}_1 = X_1, \quad \hat{\mu}_2 = \frac{X_1 + X_2}{2}, \quad \hat{\mu}_3 = \frac{X_1 + X_2 + X_{n-1} + X_n}{4},$$

$$\hat{\mu}_4 = 2X_1, \quad \hat{\mu}_5 = \frac{X_1 + X_2}{3}$$

**解：** 由 $X_1, X_2, \dots, X_n$ 相互独立，且与总体 $X$ 同分布，故有

$$E(X_i) = \mu, i=1, 2, \dots, n$$

$$E(\hat{\mu}_1) = E(X_1) = \mu \quad \longrightarrow \quad \hat{\mu}_1 \text{ 是 } \mu \text{ 的无偏估计}$$

$$\hat{\mu}_1 = X_1, \quad \hat{\mu}_2 = \frac{X_1 + X_2}{2}, \quad \hat{\mu}_3 = \frac{X_1 + X_2 + X_{n-1} + X_n}{4},$$

$$\hat{\mu}_4 = 2X_1, \quad \hat{\mu}_5 = \frac{X_1 + X_2}{3}$$

$$E(\hat{\mu}_2) = E\left(\frac{X_1 + X_2}{2}\right) = \mu$$

$$E(\hat{\mu}_3) = E\left(\frac{X_1 + X_2 + X_{n-1} + X_n}{4}\right) = \mu$$

$$E(\hat{\mu}_4) = E(2X_1) = 2\mu$$

$$E(\hat{\mu}_5) = E\left(\frac{X_1 + X_2}{3}\right) = \frac{2}{3}\mu$$

所以

$\hat{\mu}_2$ 和 $\hat{\mu}_3$ 是 $\mu$ 的无偏估计

$\hat{\mu}_4$ 和 $\hat{\mu}_5$ 是 $\mu$ 的有偏估计

## 无偏性的计算

**注：** 设总体  $X$  的均值为  $\mu$ ，方差为  $\sigma^2$ ， $X_1, X_2, \dots, X_n$

为来自总体  $X$  的随机样本，记

$\bar{X}$  与  $S^2$  分别为样本均值与样本方差，即

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

则  $E(\bar{X}) = \mu$ ， $E(S^2) = \sigma^2$

P<sub>119</sub>

**结论：** 样本均值和样本方差分别是总体均值  $\mu$  和总体方差  $\sigma^2$  的无偏估计.

# 有效性

## ◆ 问题的提出

一个参数往往有不只一个无偏估计, 怎样去比较多个无偏估计的优劣呢?

其中一个标准是看它们哪一个的取值更集中于待估参数的真值附近, 即看哪一个估计量的方差更小, 这就是**有效性**的概念。

# 有效性

## ◆ 定义

设  $\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \dots, X_n)$  与  $\hat{\theta}_2 = \hat{\theta}_2(X_1, X_2, \dots, X_n)$  都是参数  $\theta$  的无偏估计，若  $D(\hat{\theta}_1) \leq D(\hat{\theta}_2)$ ，则称  $\hat{\theta}_1$  比  $\hat{\theta}_2$  更有效。

◆ 有效性的意义在于：当用估计量去估计参数时，除了无系统误差外，还要求估计的精确度更高。



**例：** 设 $X_1, X_2, \dots, X_n$ 为来自均值为 $\mu$ 的总体 $X$ 的随机样本，考虑 $\mu$ 的如下几个**无偏**估计量哪一个更有效？

$$\hat{\mu}_1 = X_1, \quad \hat{\mu}_2 = \frac{1}{n-1} \sum_{i=2}^n X_i, \quad \hat{\mu}_3 = \frac{1}{n} \sum_{i=1}^n X_i$$

**解：** 由 $X_1, X_2, \dots, X_n$ 相互独立，且与总体 $X$ 同分布，故有

$$D(X_i) = D(X) = \sigma^2, \quad i=1, 2, \dots, n$$

$$D(\hat{\mu}_1) = D(X_1) = \sigma^2$$

$$D(\hat{\mu}_2) = D\left(\frac{1}{n-1} \sum_{i=2}^n X_i\right) = \frac{1}{(n-1)^2} \sum_{i=2}^n D(X_i) = \frac{\sigma^2}{n-1}$$

$$D(\hat{\mu}_3) = D\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n D(X_i) = \frac{\sigma^2}{n}$$

$$D(\hat{\mu}_3) \leq D(\hat{\mu}_2) \leq D(\hat{\mu}_1) \quad \longrightarrow \quad \hat{\mu}_3 \text{ 比 } \hat{\mu}_1 \text{ 和 } \hat{\mu}_2 \text{ 更有效}$$

这表明：当用样本均值去估计总体均值时，使用**全样本**总比不使用全样本要好。

# 一致性

## ◆ 问题的提出

- 估计量的无偏性和有效性都是在样本容量 $n$ 固定的情况下讨论的，但是一般估计量的好坏也依赖于样本容量。
- 当样本容量 $n$ 增大时，关于总体的信息也随之增加，此时估计量的值会更精确。
- 特别当 $n \rightarrow +\infty$ 时，估计值和参数的真值几乎完全一致，这就是估计量的一致性(或称为相合性)。

# 一致性

## ◆ 定义

设  $\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \dots, X_n)$  为未知参数  $\theta$  的无偏估计量，  
若对于任意的  $\varepsilon > 0$ ，都有

$$\lim_{n \rightarrow +\infty} P\left\{|\hat{\theta} - \theta| > \varepsilon\right\} = 0$$

或者 
$$\lim_{n \rightarrow +\infty} P\left\{|\hat{\theta} - \theta| \leq \varepsilon\right\} = 1$$

则称  $\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \dots, X_n)$  为  $\theta$  的一致 (或相合) 估计量.

◆ 一致估计量的意义：只要样本容量足够大，估计量就可以用任意接近于1的概率把参数真实值估计到任意的精度。

● 样本的  $k$  阶原点矩是总体  $X$  的  $k$  阶原点矩的一致估计量

# 第七章 参数估计

## ➤ 参数的区间估计

## 参数的区间估计

◆ 点估计就是利用样本计算出的值 (即实轴上点) 来估计未知参数。

➤ 优点：可直接地告诉人们 “未知参数大致是多少”  
；

➤ 缺点：并未反映出估计的误差范围 (精度)。

故点估计在使用上还有不尽如人意之处；而区间估计正好弥补了点估计的这一不足之处。

◆ 区间估计：以区间的形式给出包含参数 $\theta$ 真实值的可靠程度。

# 参数的区间估计

◆ 定义：设总体 $X$ 的分布中含有一个未知的参数 $\theta$ ,  $(X_1, X_2, \dots, X_n)$ 是来自总体的样本，对给定的 $\alpha (0 < \alpha < 1)$ , 如果由样本确定两个统计量

$$\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \dots, X_n), \quad \hat{\theta}_2 = \hat{\theta}_2(X_1, X_2, \dots, X_n)$$

使得

$$P\left\{\hat{\theta}_1(X_1, X_2, \dots, X_n) < \theta < \hat{\theta}_2(X_1, X_2, \dots, X_n)\right\} = 1 - \alpha$$

则称随机区间  $(\hat{\theta}_1, \hat{\theta}_2)$  为参数 $\theta$ 的置信度(或置信水平)为  $1-\alpha$ 的置信区间,  $\hat{\theta}_1$ 为置信上限,  $\hat{\theta}_2$ 为置信下限。

# 参数的区间估计

几点说明:

① 置信区间 $(\hat{\theta}_1, \hat{\theta}_2)$ 是一个随机区间, 与样本有关。

$$\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \dots, X_n), \quad \hat{\theta}_2 = \hat{\theta}_2(X_1, X_2, \dots, X_n)$$

② 对于一个给定的样本, 置信区间可能包含未知参数 $\theta$ 的真实值, 也可能不包含; 但是包含的概率为 $1-\alpha$ 。

$$P\{\hat{\theta}_1(X_1, X_2, \dots, X_n) < \theta < \hat{\theta}_2(X_1, X_2, \dots, X_n)\} = 1 - \alpha$$

③ 置信度的大小常根据实际需要来确定

通常 $\alpha$ 取:  $\alpha=0.1, 0.05, 0.01$ ;

相应的置信度为 $1-\alpha=0.9, 0.95, 0.99$ 。

# 参数的区间估计

评价区间估计的好坏有两个要素：

## ① 精度

可以用区间长度来刻画，长度越长，精度越低。

## ② 置信度 $1-\alpha$

在样本容量 $n$ 固定时，当置信度 $1-\alpha$ 增大时， $\alpha$ 变小，此时置信区间的长度变长。

即置信区间的置信度越高，则精度越低；反之，精度越高，置信度越低。

可靠度与精度是一对矛盾；

一般是在保证置信度的条件下尽可能提高精度



# 区间估计的思想

**例：** 设某厂生产的灯泡使用寿命 $X \sim N(\mu, 100^2)$ ，现随机抽取5只，测量其寿命如下：

1455, 1502, 1370, 1610, 1430

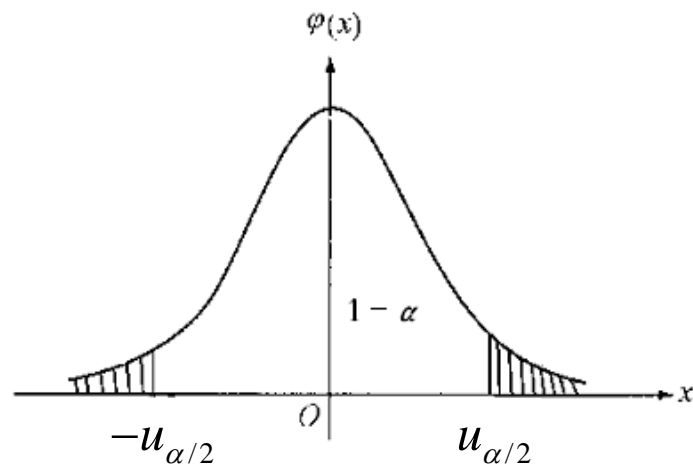
则该厂灯泡的平均使用寿命 $\mu$ 的**点估计**值为

$$\bar{x} = \frac{1}{5}(1455 + 1502 + 1370 + 1610 + 1430) = 1473.4$$

**区间估计**呢？即平均寿命的范围有多大，又有多大的可能性是在该区间内呢？

平均寿命的范围有多大, 有多大的概率在该区间内呢?  
如果要求有95%的把握判断 $\mu$ 在1473.4左右, 则由U统计量可知

$$U = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$



由

$$P\left\{\left|\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right| < \varepsilon\right\} = 2\Phi(\varepsilon) - 1 = 0.95 \quad \longrightarrow \quad \Phi(\varepsilon) = 0.975$$

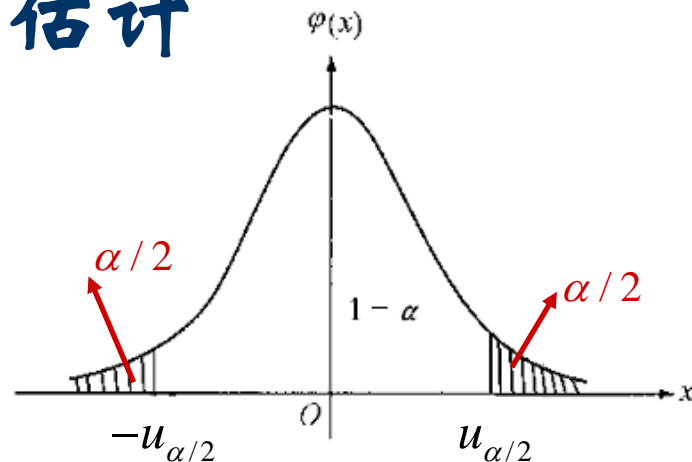
查表得  $\varepsilon = 1.96$  利用分位点查表

区间估计为  $\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}}$

# 总体均值 $\mu$ 的区间估计

$$\text{令 } U = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0, 1)$$

构造一个含 $\mu$ 的函数,要求其分布为已知



$$P(-u_{\alpha/2} \leq U \leq u_{\alpha/2}) = 1 - \alpha$$

利用已知分布,求出分位点

$$-u_{\alpha/2} \leq U \leq u_{\alpha/2} \text{ 即 } -u_{\alpha/2} \leq \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \leq u_{\alpha/2}$$

代入U

$$\text{有 } \bar{X} - \frac{\sigma}{\sqrt{n}} u_{\alpha/2} \leq \mu \leq \bar{X} + \frac{\sigma}{\sqrt{n}} u_{\alpha/2}$$

解出 $\mu$ 的置信度为  
 $1 - \alpha$ 的置信区间为

$$\hat{\mu}_1$$

$$\hat{\mu}_2$$

$$(\hat{\mu}_1, \hat{\mu}_2)$$

# 参数的区间估计

◆ 正态总体均值 $\mu$ 的区间估计：设 $X \sim N(\mu, \sigma^2)$

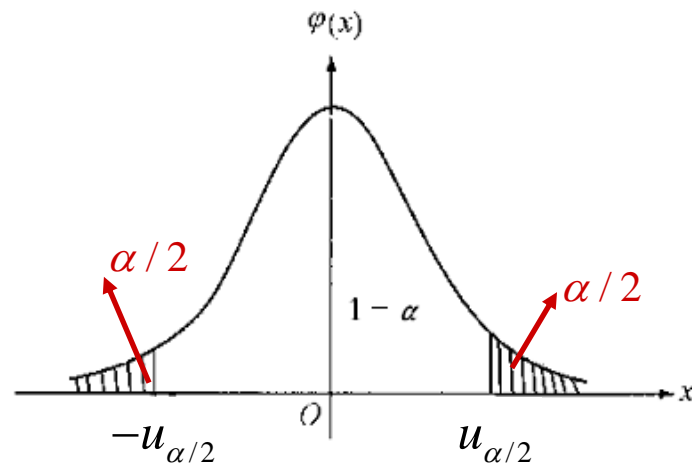
① 总体方差 $\sigma^2$ 已知时

构造 
$$U = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0, 1)$$

$\mu$ 的置信度为 $1-\alpha$ 的置信区间为

$$\left( \bar{X} - \frac{\sigma}{\sqrt{n}} u_{\alpha/2}, \bar{X} + \frac{\sigma}{\sqrt{n}} u_{\alpha/2} \right)$$

其中 $u_{\alpha/2}$ 是标准正态分布的 $\alpha/2$ 上侧分位点.



# 参数的区间估计

◆ 正态总体均值 $\mu$ 的区间估计：设 $X \sim N(\mu, \sigma^2)$

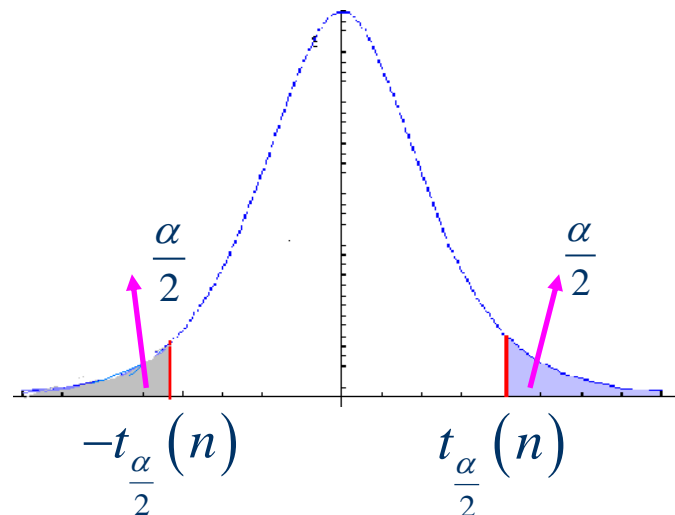
② 总体方差 $\sigma^2$ 未知时

构造  $T = \frac{\bar{X} - \mu}{S / \sqrt{n}} \sim t(n-1)$

$\mu$  的置信度为 $1-\alpha$ 的置信区间为

$$\left( \bar{X} - \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1), \bar{X} + \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1) \right)$$

其中 $t_{\alpha/2}(n-1)$ 是分布 $t(n-1)$ 的 $\alpha/2$ 上侧分位点.



## 区间估计的计算

**例：**某厂生产的一种塑料口杯的重量 $X$ 被认为服从正态分布，今随机抽取9个，测得其重量为(单位：克)：

21.1, 21.3, 21.4, 21.5, 21.3, 21.7, 21.4, 21.3, 21.6

试用95%的置信度估计全部口杯的平均重量.

$$\frac{\bar{X} - \mu}{S / \sqrt{n}} \sim t(n-1) \quad \left( \bar{X} - \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1), \bar{X} + \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1) \right)$$

**解：**由题可知： $X \sim N(\mu, \sigma^2)$ ，且由 $1-\alpha=0.95$ ，得

$$\alpha=0.05 \quad \bar{x} = 21.4, S = 0.18, n = 9, t_{0.025}(8) = 2.306$$

$$\frac{S}{\sqrt{n}} t_{\alpha/2}(8) = \frac{0.18}{\sqrt{9}} \cdot 2.306 = 0.13836$$

全部口杯的平均重量的95%置信区间为(21.26, 21.54)

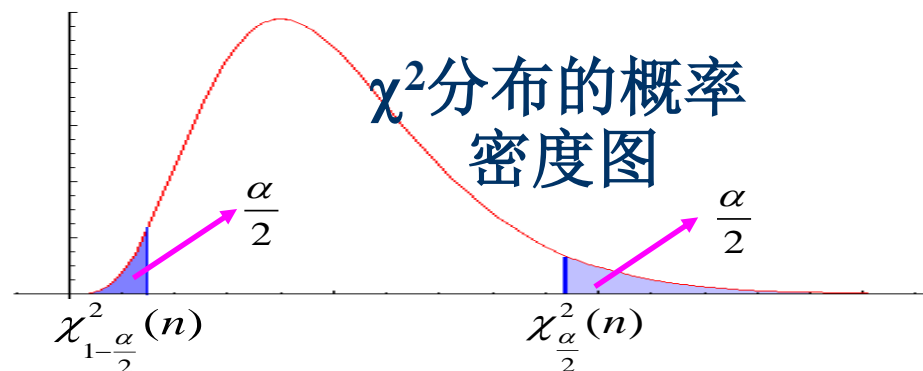
# 参数的区间估计

◆ 正态总体方差 $\sigma^2$ 的区间估计：设 $X \sim N(\mu, \sigma^2)$

①  $\mu$ 已知时总体方差 $\sigma^2$ 的区间估计

构造

$$\sum_{i=1}^n \left( \frac{X_i - \mu}{\sigma} \right)^2 = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} \sim \chi^2(n)$$



$\sigma^2$ 的置信度为 $1-\alpha$ 的置信区间为

$$\left( \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi^2_{\alpha/2}(n)}, \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi^2_{1-\alpha/2}(n)} \right)$$

# 参数的区间估计

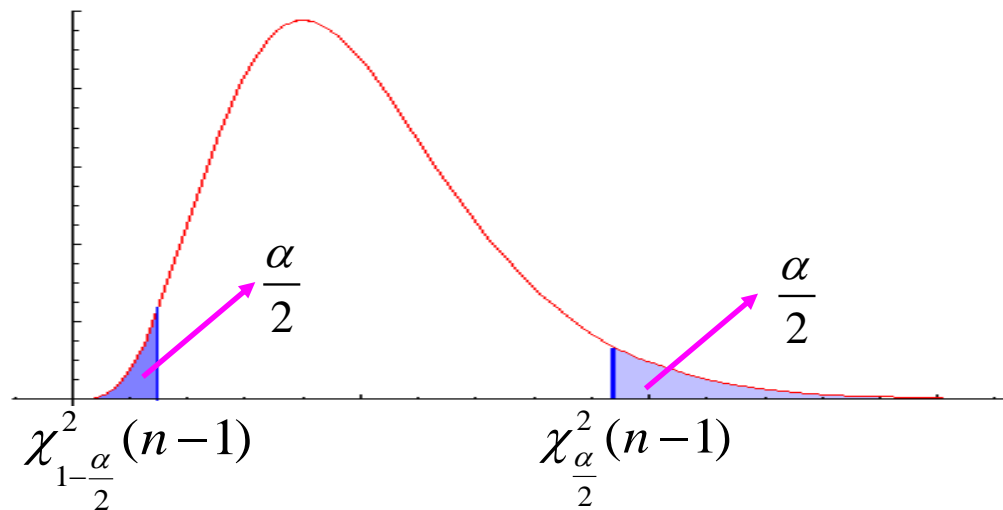
◆ 正态总体方差 $\sigma^2$ 的区间估计：设 $X \sim N(\mu, \sigma^2)$

②  $\mu$ 未知时总体方差 $\sigma^2$ 的区间估计

构造 
$$\frac{(n-1)S^2}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 \sim \chi^2(n-1)$$

$\sigma^2$ 的置信度为 $1-\alpha$ 的置信区间为

$$\begin{aligned} & \left( \frac{(n-1)S^2}{\chi_{\alpha/2}^2(n-1)}, \frac{(n-1)S^2}{\chi_{1-\alpha/2}^2(n-1)} \right) \\ &= \left( \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\chi_{\alpha/2}^2(n-1)}, \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\chi_{1-\alpha/2}^2(n-1)} \right) \end{aligned}$$





## 区间估计的计算

**例：**为估计一物体的重量 $\mu$ ，将其称量10次，得到重量的测量值 (单位：千克) 如下：

10.1, 10.0, 9.8, 10.5, 9.7, 10.1, 9.9, 10.2, 10.3, 9.9

设它们服从正态分布 $X \sim N(\mu, \sigma^2)$ 。求 $\sigma^2$ 的置信度为0.95的置信区间。 $\left( \frac{(n-1)S^2}{\chi_{\alpha/2}^2(n-1)}, \frac{(n-1)S^2}{\chi_{1-\alpha/2}^2(n-1)} \right)$

**解：** $n=10$ ， $\alpha=0.05$ ， $S^2=0.0583$ ，查附表得：

$$\chi_{0.975}^2(9) = 2.70, \chi_{0.025}^2(9) = 19.023$$

$$\begin{aligned} \left( \frac{(n-1)S^2}{\chi_{\alpha/2}^2(n-1)}, \frac{(n-1)S^2}{\chi_{1-\alpha/2}^2(n-1)} \right) &= \left( \frac{9 \times 0.0583}{19.023}, \frac{9 \times 0.0583}{2.7} \right) \\ &= (0.028, 0.194) \end{aligned}$$

# 两个正态总体的区间估计

## ◆ 背景

在实际应用中，经常会遇到两个正态总体的区间估计问题。

**例如：**考察一项新技术对提高产品的某项质量指标的作用  
将实施新技术前的产品质量指标看成正态总体 $N(\mu_1, \sigma_1^2)$ ，  
实施新技术后产品质量指标看成正态总体 $N(\mu_2, \sigma_2^2)$ 。

于是，评价新技术的效果问题，就归结为研究两个正态总体均值之差和总体方差之比的区间估计。

**又如：**比较甲乙两厂生产某种药物的治疗效果时，可以把两个厂的药效分别看成服从正态分布的两个总体.那么药效的差异，即是两个正态总体均值和方差的差异。

## 两个正态总体的区间估计

➤ 设  $X_1, X_2, \dots, X_{n_1}$  是抽自正态总体  $X$  的简单随机样本,  $X \sim N(\mu_1, \sigma_1^2)$ , 样本均值与样本方差为

$$\bar{X} = \frac{1}{n_1} \sum_{i=1}^{n_1} X_i, \quad S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_i - \bar{X})^2$$

➤ 设  $Y_1, Y_2, \dots, Y_{n_2}$  是抽自正态总体  $Y$  的简单随机样本,  $Y \sim N(\mu_2, \sigma_2^2)$ , 样本均值与样本方差为

$$\bar{Y} = \frac{1}{n_2} \sum_{i=1}^{n_2} Y_i, \quad S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (Y_i - \bar{Y})^2$$

➤  $(X_1, X_2, \dots, X_{n_1})$  和  $(Y_1, Y_2, \dots, Y_{n_2})$  是独立样本

# 两个正态总体的区间估计

◆ 两个正态总体均值差 $\mu_1 - \mu_2$ 的区间估计:

① 设 $\sigma_1^2$ 和 $\sigma_2^2$ 都已知

利用

$$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right),$$

$\mu_1 - \mu_2$ 的置信度为 $1-\alpha$ 的置信区间为

$$\left( (\bar{X} - \bar{Y}) - u_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, (\bar{X} - \bar{Y}) + u_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right)$$

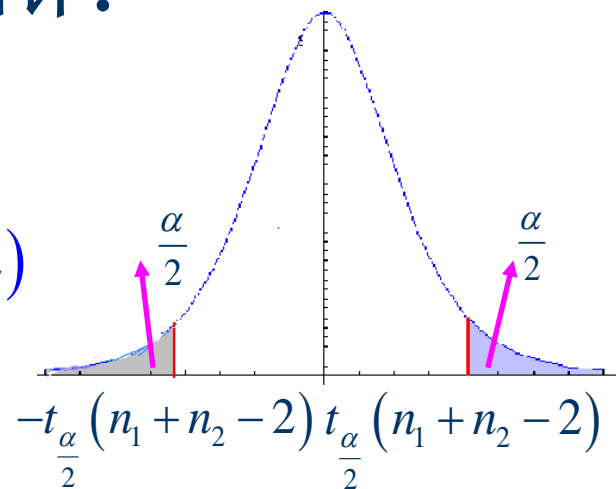
# 两个正态总体的区间估计

◆ 两个正态总体均值差 $\mu_1 - \mu_2$ 的区间估计:

② 设 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ , 但未知时

利用 
$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$

$\mu_1 - \mu_2$ 的置信度为 $1 - \alpha$ 的置信区间为



$$\left( (\bar{X} - \bar{Y}) \mp t_{\alpha/2}(n_1 + n_2 - 2) S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right)$$

其中 
$$S_w^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

# 两个正态总体的区间估计

◆ 两个正态总体方差比  $\frac{\sigma_1^2}{\sigma_2^2}$  的区间估计:

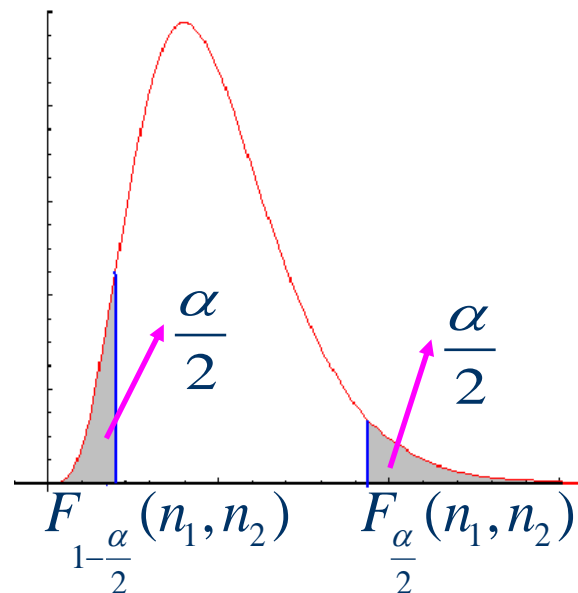
➤ 设  $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$  都未知

利用

$$F = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim F(n_1 - 1, n_2 - 1)$$

$\frac{\sigma_1^2}{\sigma_2^2}$  的置信度为  $1-\alpha$  的置信区间为

$$\left( \frac{S_1^2 / S_2^2}{F_{\alpha/2}(n_1 - 1, n_2 - 1)}, \frac{S_1^2 / S_2^2}{F_{1-\alpha/2}(n_1 - 1, n_2 - 1)} \right)$$



# 区间估计的计算

**例：**比较棉花品种的优劣

假设用甲、乙两种棉花纺出的棉纱强度分别为 $X \sim N(\mu_1, 2.18^2)$ 和 $Y \sim N(\mu_2, 1.76^2)$ 。试验者从这两种棉纱中分别抽取样本  $X_1, X_2, \dots, X_{200}$  和  $Y_1, Y_2, \dots, Y_{100}$ ，样本均值分别为：

$$\bar{X} = 5.32, \bar{Y} = 5.76$$

求  $\mu_1 - \mu_2$  的置信系数为 0.95 的区间估计。

**解：**  $\alpha = 0.05, u_{\alpha/2} = u_{0.025} = 1.96, \sigma_1 = 2.18, \sigma_2 = 1.76, n_1 = 200, n_2 = 100,$

可得  $\mu_1 - \mu_2$  的置信度为  $1 - \alpha$  的置信区间为

$$\left( (\bar{X} - \bar{Y}) - u_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, (\bar{X} - \bar{Y}) + u_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right)$$

$$= (-0.899, 0.019)$$



## 区间估计的计算

**例：**某公司利用两条自动化流水线灌装矿泉水。设这两条流水线所装矿泉水的体积 (单位:毫升) $X \sim N(\mu_1, \sigma^2)$  和  $Y \sim N(\mu_2, \sigma^2)$ . 现从生产线上分别抽取  $X_1, X_2, \dots, X_{12}$  和  $Y_1, Y_2, \dots, Y_{17}$ , 样本均值与样本方差分别为:

$$\bar{X} = 501.1, S_1^2 = 2.4; \quad \bar{Y} = 499.7, S_2^2 = 4.7$$

求  $\mu_1 - \mu_2$  的置信系数为0.95的区间估计。

**解：** $n_1=12, n_2=17, \alpha=0.05, t_{n_1+n_2-2}(0.025)=t_{27}(0.025)=2.05$ , 再由其他已知条件可算出

$$S_w = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{(12 - 1) \times 2.4 + (17 - 1) \times 4.7}{12 + 17 - 2}} = 1.94$$

故 $\mu_1 - \mu_2$  的置信系数为  $1-\alpha$  的置信区间

$$\left( (\bar{X} - \bar{Y}) \mp t_{n_1+n_2-2}(\alpha/2) S_w \sqrt{n_1^{-1} + n_2^{-1}} \right) = (-0.101, 2.901)$$



## 两个正态总体的区间估计

➤ 两个正态总体均值差 $\mu_1 - \mu_2$ 的区间估计的说明

在这两个例子中， $\mu_1 - \mu_2$  的置信区间都包含了零，也就是说： $\mu_1$ 可能大于 $\mu_2$ ，也可能小于 $\mu_2$ 。  
这时我们认为二者没有显著差异。

# 区间估计的计算

**例：**某公司利用两条自动化流水线灌装矿泉水。设这两条流水线所装矿泉水的体积 (单位:毫升)  $X \sim N(\mu_1, \sigma^2)$  和  $Y \sim N(\mu_2, \sigma^2)$ . 现从生产线上分别抽取  $X_1, X_2, \dots, X_{12}$  和  $Y_1, Y_2, \dots, Y_{17}$ , 样本均值与样本方差分别为:

$$\bar{X} = 501.1, \quad S_1^2 = 2.4; \quad \bar{Y} = 499.7, \quad S_2^2 = 4.7$$

求 的置信系数为**0.95**的区间估计。

**解：**  $n_1=12, n_2=17, \alpha=0.05$ , 再由其他已知条件可算出

$$S_w = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{(12 - 1) \times 2.4 + (17 - 1) \times 4.7}{12 + 17 - 2}} = 1.94$$

查表  $t_{27}(0.025)=2.05$ , 故  $\mu_1 - \mu_2$  的置信系数为  $1 - \alpha$  的置信区间

$$\left( (\bar{X} - \bar{Y}) \mp t_{n_1+n_2-2}(\alpha/2) S_w \sqrt{n_1^{-1} + n_2^{-1}} \right) = (-0.101, 2.901)$$

## 区间估计的计算

**例：**研究由机器  $A$  和机器  $B$  生产的钢管的内径，随机地抽取机器  $A$  生产的钢管 18 只，测得样本方差  $s_2^2 = 0.29(mm^2)$ 。随机地取机器  $B$  生产的钢管 13 只，测得样本方差  $s_1^2 = 0.34(mm^2)$ ；设两样本相互独立，且设由机器  $A$  和机器  $B$  生产的钢管的内径分别服从正态分布  $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2)$ ，这里  $\mu_i, \sigma_i^2$  ( $i = 1, 2$ ) 均未知。试求方差比  $\sigma_1^2 / \sigma_2^2$  的置信水平为 0.90 的置信区间。

# 区间估计的计算

**解：**  $\alpha=0.10$ ,  $\alpha/2=0.05$ ,  $1-\alpha/2=0.95$

$$n_1 = 18, s_1^2 = 0.34, n_2 = 13, s_2^2 = 0.29$$

$$F_{0.05}(17,12) = 2.59, F_{0.95}(17,12) = \frac{1}{F_{0.05}(12,17)} = \frac{1}{2.38}$$

故两总体方差比 $\sigma_1^2 / \sigma_2^2$ 的置信水平为0.90的置信区间为

$$\left( \frac{S_1^2 / S_2^2}{F_{\alpha/2}(n_1-1, n_2-1)}, \frac{S_1^2 / S_2^2}{F_{1-\alpha/2}(n_1-1, n_2-1)} \right) = (0.45, 2.79)$$

**【说明】** 方差比的置信区间包含了1，也就是说： $\sigma_1$ 可能大于 $\sigma_2$ ，也可能小于 $\sigma_2$ 。这时我们认为二者没有显著差异。

## 小 结

本章首先介绍了参数点估计的两种方法：  
矩估计和极大似然估计；接着介绍了点估计的  
评优准则，包括：无偏性，有效性和一致性；  
然后介绍了区间估计的基本概念，详细讨论了  
正态总体下均值和方差的区间估计。

# 第八章 假设检验

- ◆ 假设检验的基本步骤
- ◆ 正态总体参数的假设检验

# 假设检验

在总体的分布函数完全未知或只知其形式、但不知其参数的情况下，为了推断总体的某些性质，提出某些关于总体的假设.

- 提出总体服从泊松分布的假设;
- 对正态分布总体提出数学期望等于 $\mu_0$ 的假设等。

假设检验就是根据得到样本对所提出的假设作出判断：是接受，还是拒绝.

通常借助于直观分析和理论分析相结合的做法,其基本原理就是人们在实际问题中经常采用的所谓实际推断原理:“一个小概率事件在一次试验中几乎是不可能发生的”.

# 假设检验的基本思想

**例：**某工厂生产长度为 10 cm 的零件，根据以往生产的实际情况，可以认为：长度值  $X$  服从正态分布  $N(\mu, 0.1^2)$ 。现在随机抽取 10 个零件，测得它们的值为：

9.9, 10.1, 10.2, 9.7, 9.9, 9.9, 10.0, 10.5, 10.1, 10.2.

问：从样本看，能否认为该厂生产的零件的长度的平均值  $\mu = 10$  cm?



**例：**某工厂生产长度为 10 cm 的零件，根据以往生产的实际情况，可以认为：长度值  $X$  服从正态分布  $N(\mu, 0.1^2)$ 。现在随机抽取 10 个零件，测得它们的值为：9.9, 10.1, 10.2, 9.7, 9.9, 9.9, 10.0, 10.5, 10.1, 10.2. 问：从样本看，能否认为该厂生产的零件的长度的平均值  $\mu = 10$  cm？

## I. 如何建立检验模型

- 确定总体：记  $X$  为该厂生产的零件的长度，  
则  $X \sim N(\mu, 0.1^2)$ ;
- 明确任务：通过样本推断 “ $X$  的均值  $\mu$  是否等于 10cm”;
- 假设：下面的任务是要通过样本检验 “ $X$  的均值  $\mu = 10$ ” 这一假设是否成立。

# 假设检验的基本思想

## ✓ 一些概念

在数理统计中，把“ $X$  的均值  $\mu=10$ ”这样一个待检验的假设记为“原假设”或“零假设”，记“ $H_0: \mu=10$ ”。

原假设的对立面是“ $X$  的均值  $\mu \neq 10$ ”，称为“对立假设”或“备择假设”，记成“ $H_1: \mu \neq 10$ ”。

把原假设和对立假设合写在一起，就是：

$$H_0: \mu=10; H_1: \mu \neq 10$$

**例：**某工厂生产长度为 10 cm 的零件，根据以往生产的实际情况，可以认为：长度值  $X$  服从正态分布  $N(\mu, 0.1^2)$ 。现在随机抽取 10 个零件，测得它们的值为：9.9, 10.1, 10.2, 9.7, 9.9, 9.9, 10.0, 10.5, 10.1, 10.2. 问：从样本看，能否认为该厂生产的零件的长度的平均值  $\mu = 10$  cm？

## II. 解决问题的思路

因样本均值是  $\mu$  的一个很好的估计。所以当  $\mu = 10$ ，即原假设  $H_0$  成立时， $|\bar{X} - 10|$  应比较小；如果该值过大，想必  $H_0$  不成立。

于是，我们就用  $|\bar{X} - 10|$  的大小检验  $H_0$  是否成立。

合理的做法应该是：找出一个界限  $c$ ，

当  $|\bar{X} - 10| < c$  时，接受原假设  $H_0$ ；

当  $|\bar{X} - 10| \geq c$  时，拒绝原假设  $H_0$ 。

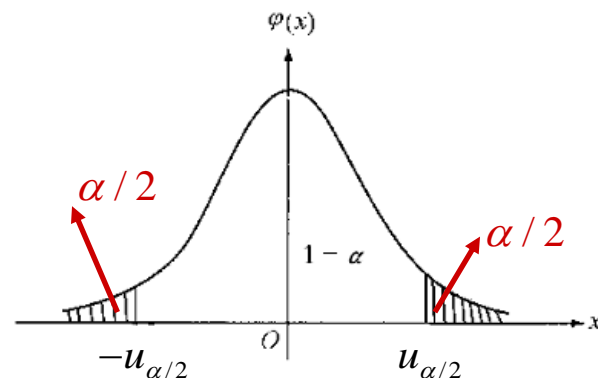
# 假设检验的基本思想

◆ 这里的问题是：如何确定常数 $c$ 呢？

$$\bar{X} \sim N\left(\mu, \frac{0.1^2}{10}\right) \text{ 或 } \frac{\bar{X} - \mu}{0.1/\sqrt{10}} \sim N(0, 1)$$

于是，当原假设 $H_0: \mu=10$ 成立时，有

$$\frac{\bar{X} - 10}{0.1/\sqrt{10}} \sim N(0, 1)$$



为确定常数 $c$ ，我们考虑一个很小的正数 $\alpha$ ，如 $\alpha=0.05$ 。当原假设 $H_0: \mu=10$ 成立时，有

$$P\left\{\frac{|\bar{X} - 10|}{0.1/\sqrt{10}} \geq u_{\alpha/2}\right\} = \alpha$$

$$\text{即 } P\left\{|\bar{X} - 10| \geq \left(0.1/\sqrt{10}\right)u_{\alpha/2}\right\} = \alpha \Rightarrow c = \left(0.1/\sqrt{10}\right)u_{\alpha/2}$$

# 假设检验的基本思想

于是，我们就得到如下**检验准则**：

当 $|\bar{X} - 10| < c$ 时，接受原假设  $H_0$ ；

当 $|\bar{X} - 10| \geq c$ 时，拒绝原假设  $H_0$ . 其中  $c = (0.1 / \sqrt{10}) u_{\alpha/2}$

称  $U = \frac{\bar{X} - 10}{0.1 / \sqrt{10}}$  为检验统计量； $\alpha$  称为显著性水平。

称  $|U| = \frac{|\bar{X} - 10|}{0.1 / \sqrt{10}} \geq u_{\alpha/2}$  为原假设  $H_0$  的拒绝域。

1. 建立检验统计量
2. 根据检验统计量的分布确定拒绝域

假设检验最  
关键的两步

# 假设检验的基本思想

**例：**某工厂生产长度为 10 cm 的零件，根据以往生产的实际情况，可以认为：长度值  $X$  服从正态分布  $N(\mu, 0.1^2)$ 。现在随机抽取 10 个零件，测得它们的值为：

9.9, 10.1, 10.2, 9.7, 9.9, 9.9, 10.0, 10.5, 10.1, 10.2.

问：从样本看，能否认为该厂生产的零件的长度的平均值  $\mu = 10$  cm?

用以上检验准则处理我们的问题

$$\alpha = 0.05, u_{\alpha/2} = u_{0.025} = 1.96, \bar{X} = 10.05$$

$$c = \left(0.1 / \sqrt{10}\right) u_{\alpha/2} = \left(0.1 / \sqrt{10}\right) \times 1.96 \approx 0.062$$

$$|\bar{X} - 10| = 0.05 < c = 0.062 \quad \text{接受原假设 } H_0: \mu = 10$$

# 假设检验的基本思想

## III. 方法原理

因为，当原假设是 $H_0: \mu=10$ 成立时，

$$P\left\{|\bar{X}-10|\geq\left(0.1/\sqrt{10}\right)u_{\alpha/2}\right\}=\alpha$$

所以，当 $\alpha$ 很小时，若 $H_0$ 为真(正确)，则检验统计量落入拒绝域是一小概率事件(概率很小，为 $\alpha$ )。前面我们曾提到：“**通常认为小概率事件在一次试验中基本上不会发生**”。那么，如果小概率事件发生了，即：

$$\left\{|\bar{X}-10|\geq\left(0.1/\sqrt{10}\right)u_{\alpha/2}\right\}$$

就拒绝接受 $H_0$ 成立，即认为 $H_0$ 不成立。

# 假设检验的基本思想

## IV. 两类错误与显著性水平

假设检验的依据是：小概率事件在一次试验中很难发生，但很难发生不等于不发生。因而当我们检验一个假设 $H_0$ 时，有可能犯以下两类错误之一：

- ①  $H_0$ 是正确的，但被我们拒绝了，这就犯了“弃真”的错误，即抛弃了正确假设，我们称之为犯第一类错误；
- ②  $H_0$ 是不正确的，但被我们接受了，这就犯了“取伪”的错误，即采用了伪假设，我们称之为犯第二类错误；

在假设检验问题中，作出判断依据的只是一个样本，检验统计量总是随机的，所以我们总是以一定的概率犯以上两类错误。



# 假设检验的基本思想

通常用 $\alpha$ 和 $\beta$ 记犯第一、第二类错误的概率，即

$$\alpha = P\{\text{拒绝}H_0 \mid H_0\text{为真}\}, \quad \beta = P\{\text{接受}H_0 \mid H_0\text{为假}\}$$

在检验问题中，犯“弃真”和“取伪”两类错误都总是不可避免的。

当样本容量  $n$  一定时，若减少犯第一类错误的概率，则犯第二类错误的概率往往增大；反之亦然。若要使犯两类错误的概率都减小，除非增加样本容量所以，犯两类错误的概率不能同时得到控制。

在统计学中，通常把控制犯第一类错误的概率，而不考虑犯第二类错误的概率的检验，称为**显著性检验**。一般事先选定一个数 $\alpha$  ( $0 < \alpha < 1$ ，一般 $\alpha = 0.1, 0.05, 0.01$ )，要求犯第一类错误的概率不超过 $\alpha$ ，称 $\alpha$ 为假设检验的**显著性水平**，简称**水平**。

# 假设检验的基本思想

## ➤ 假设检验问题通常叙述为:

在显著性水平 $\alpha$ 下, 检验假设  $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$

## ➤ 双边假设检验与单边假设检验

形如 $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$ 中, 备择假设  $H_1$ 表示  $\mu$ 可能大于 $\mu_0$ , 也可能小于  $\mu_0$ , 称为双边备择假设. 这类假设检验的拒绝域位于两侧, 我们称之为双侧(双边)假设检验。

形如 $H_0: \mu \leq \mu_0, H_1: \mu > \mu_0$ 中, 这类假设检验的拒绝域位于一侧, 我们称之为单侧(单边)假设检验。

# 假设检验的一般步骤

1. 根据实际问题，提出原假设 $H_0$ 和备择假设 $H_1$ ；
2. 根据实际问题的要求，选取适当的检验统计量并确定拒绝域的形式；
3. 根据给定的显著性水平 $\alpha$ ， $P(\text{拒绝}H_0 \mid H_0\text{为真}) = \alpha$ ，求出拒绝域。
4. 根据样本观测值计算出检验统计量的值，从而确定是接受 $H_0$ ，还是拒绝 $H_0$ 。

## 第二节 一个正态总体的假设检验

- 一个正态总体均值的假设检验
- 一个正态总体方差的假设检验


# 一个正态总体均值的假设检验

设总体 $X \sim N(\mu, \sigma^2)$ ,  $(X_1, X_2, \dots, X_n)$ 是来自 $X$ 的样本, 样本均值与样本方差如下:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

关于总体均值 $\mu$ 的假设检验的分类:

□ 已知方差 $\sigma^2$ :

①  $H_0: \mu = \mu_0$  

②  $H_0: \mu \leq \mu_0$  

③  $H_0: \mu \geq \mu_0$  

□ 未知方差 $\sigma^2$ :

①  $H_0: \mu = \mu_0$

②  $H_0: \mu \leq \mu_0$

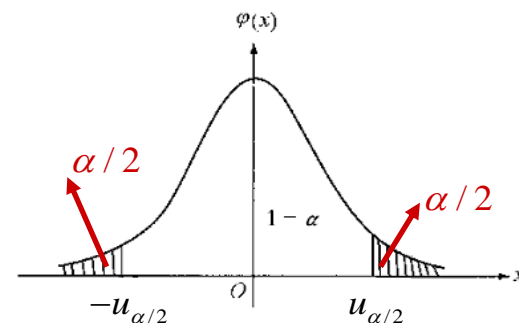
③  $H_0: \mu \geq \mu_0$

# 一个正态总体均值的假设检验

◆  $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$

① 已知方差  $\sigma^2$

构造检验统计量:  $U = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1)$

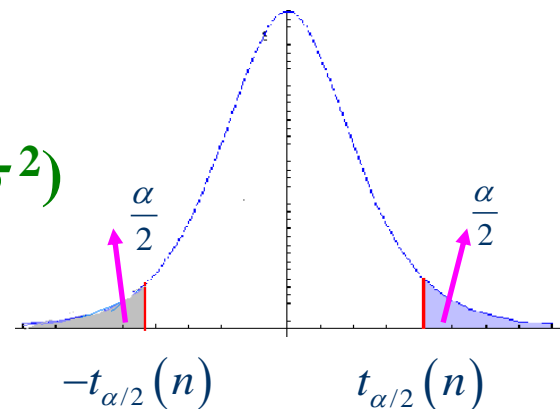


对于给定的显著性水平  $\alpha$ , 拒绝域为:  $W = \{|U| \geq u_{\alpha/2}\}$

这种检验法称作 **U检验法**

② 未知方差  $\sigma^2$  (用样本方差  $S^2$  代替未知的  $\sigma^2$ )

构造检验统计量:  $T = \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \sim t(n-1)$



对于给定的显著性水平  $\alpha$ , 拒绝域为:  $W = \{|T| \geq t_{\alpha/2}(n-1)\}$

这种检验法称作 **t 检验法**

检验统计量落入  
拒绝域的概率为  $\alpha$

# 假设检验的计算

**例：**某工厂生产长度为 10 cm 的零件，根据以往生产的实际情况，可以认为：长度值  $X$  服从正态分布  $N(\mu, \sigma^2)$ 。现在随机抽取 10 个零件，测得它们的值为：9.9, 10.1, 10.2, 9.7, 9.9, 9.9, 10.0, 10.5, 10.1, 10.2。问：从样本看，能否认为该厂生产的零件的长度的平均值  $\mu = 10$  cm？

**解：**建立假设： $H_0: \mu=10$ ； $H_1: \mu \neq 10$

$$n=10, \alpha=0.05, t_{\alpha/2}(n-1)=t_{0.025}(9)=2.2622$$

$$\text{计算得 } \bar{X}=10.05, S^2=0.05, S=0.224$$

$$|T| = \left| \frac{\bar{X} - 10}{S / \sqrt{n}} \right| = \left| \frac{10.05 - 10}{0.224 / \sqrt{10}} \right| = 0.7059 < 2.2622$$

所以，接受原假设  $H_0: \mu=10$ ，认为该厂生产的零件长度的平均值  $\mu=10$  cm。

# 一个正态总体均值的假设检验

总体  $X \sim N(\mu, \sigma^2)$

◆  $H_0: \mu \leq \mu_0$  (单边假设检验)

① 已知方差  $\sigma^2$

构造检验统计量:  $U = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$  不一定服从  $N(0,1)$

当  $H_0$  成立时, 由题设

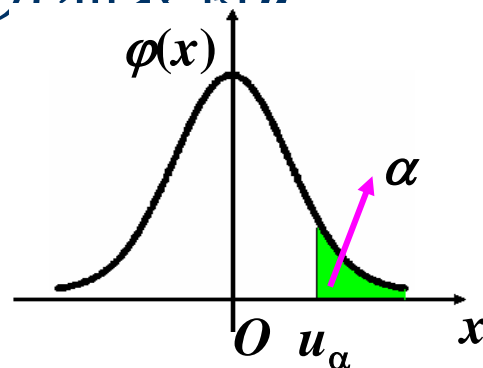
$$U' = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

且有  $U \leq U'$ , 对于给定的显著性水平  $\alpha$ , 查正态分布表得

即有 
$$P\left\{\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > u_\alpha\right\} \leq P\left\{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > u_\alpha\right\} = \alpha$$

取拒绝域 
$$V = \left\{\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > u_\alpha\right\}$$

最后计算检验统计量的值, 若  $U > u_\alpha$ , 则拒绝  $H_0$



检验统计量落入  
拒绝域的概率  $\leq \alpha$



# 一个正态总体均值的假设检验

◆  $H_0: \mu \leq \mu_0$

① 已知方差  $\sigma^2$

构造检验统计量:  $U = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$  不一定服从  $N(0,1)$

对于给定的显著性水平  $\alpha$ , 拒绝域为:  $V = \{U > u_\alpha\}$

这种检验法称作 **U检验法**

② 未知方差  $\sigma^2$  (常用样本方差  $S^2$  代替未知的  $\sigma^2$ )

构造检验统计量:  $T = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}$  不一定服从  $t(n-1)$

对于给定的显著性水平  $\alpha$ , 拒绝域为:  $V = \{T > t_\alpha(n-1)\}$

这种检验法称作 **t 检验法**

检验统计量落入  
拒绝域的概率  $\leq \alpha$

# 假设检验的计算

**例：**某种电子元件的寿命 $X$ (以小时计)服从正态分布 $N(\mu, \sigma^2)$ , 现测得16只元件的寿命如下：

159 280 101 212 224 379 179 264

222 362 168 250 149 260 485 170

问是否有理由认为元件的平均寿命大于225(小时)？

**解：**建立假设： $H_0 : \mu \leq \mu_0 = 225, H_1 : \mu > 225$

取显著性水平 $\alpha=0.05$ ,  $n=16$ ,  $t_{\alpha}(n-1)=t_{0.05}(15)=1.7531$

计算得  $\bar{x} = 241.5, s = 98.7259$

$$T = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} = 0.6685 < 1.7531$$

故接受 $H_0$ , 认为元件的平均寿命不大于225小时.

# 一个正态总体均值的假设检验

◆  $H_0: \mu \geq \mu_0$

① 已知方差  $\sigma^2$

构造检验统计量:  $U = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$  不一定服从  $N(0,1)$

对于给定的显著性水平  $\alpha$ , 拒绝域为:  $V = \{U < -u_\alpha\}$

这种检验法称作 **U检验法**

② 未知方差  $\sigma^2$  (常用样本方差  $S^2$  代替未知的  $\sigma^2$ )

构造检验统计量:  $T = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}$  不一定服从  $t(n-1)$

对于给定的显著性水平  $\alpha$ , 拒绝域为:  $V = \{T < -t_\alpha(n-1)\}$

这种检验法称作 **t 检验法**

检验统计量落入  
拒绝域的概率  $\leq \alpha$

# 一个正态总体方差的假设检验

设总体 $X \sim N(\mu, \sigma^2)$ ,  $(X_1, X_2, \dots, X_n)$ 是来自 $X$ 的样本, 样本均值与样本方差如下:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

关于总体方差 $\sigma^2$ 的假设检验的分类

①  $H_0 : \sigma^2 = \sigma_0^2$

②  $H_0 : \sigma^2 \leq \sigma_0^2$

③  $H_0 : \sigma^2 \geq \sigma_0^2$

# 一个正态总体方差的假设检验 (以总体均值未知为前提)

①  $H_0: \sigma^2 = \sigma_0^2$  (双边假设检验)

思路分析:

利用样本方差  $S^2$  是  $\sigma^2$  的一个无偏估计, 且  $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$ .  
当原假设  $H_0: \sigma^2 = \sigma_0^2$  成立时,  $S^2$  和  $\sigma_0^2$  应该比较接近, 即比值  $S^2/\sigma_0^2$  应接近于 1。所以这个比值过大或过小时, 应拒绝原假设。

合理的做法是: 找两个合适的界限  $c_1$  和  $c_2$ , 使得

- 当  $c_1 < (n-1)S^2/\sigma_0^2 < c_2$  时, 接受  $H_0$ ;
- 当  $(n-1)S^2/\sigma_0^2 \leq c_1$  或  $(n-1)S^2/\sigma_0^2 \geq c_2$  时, 拒绝  $H_0$ .

# 一个正态总体方差的假设检验

问题是 $c_1$ 与 $c_2$ 如何确定？

当原假设  $H_0: \sigma^2 = \sigma_0^2$  成立时，由定理6.2有

$$(n-1)S^2 / \sigma_0^2 \sim \chi^2(n-1)$$

$$P\left\{\chi_{1-\alpha/2}^2(n-1) < \frac{(n-1)S^2}{\sigma_0^2} < \chi_{\alpha/2}^2(n-1)\right\} = 1 - \alpha$$

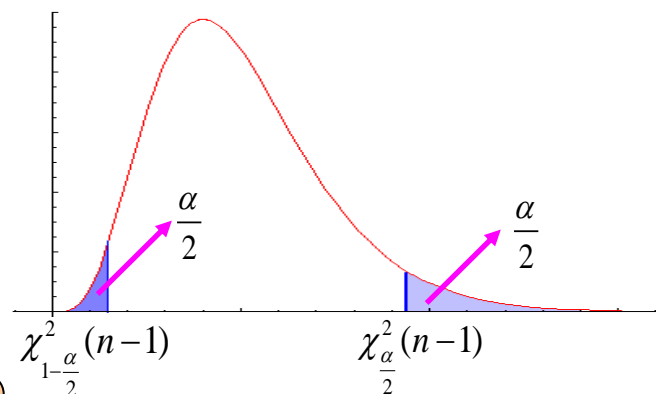
检验统计量

故  $H_0$  的拒绝域为

$$W = \left\{ \frac{(n-1)S^2}{\sigma_0^2} < \chi_{1-\alpha/2}^2(n-1) \text{ 或 } \frac{(n-1)S^2}{\sigma_0^2} > \chi_{\alpha/2}^2(n-1) \right\}$$

上述检验法称为  $\chi^2$  检验法

检验统计量落入拒绝域的概率为 $\alpha$



# 一个正态总体方差的假设检验

②  $H_0: \sigma^2 \leq \sigma_0^2; H_1: \sigma^2 > \sigma_0^2$

由定理6.2有

$$(n-1)S^2 / \sigma^2 \sim \chi^2(n-1)$$

当 $H_0$ 成立时，构造检验统计量： $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$

$$P\left\{\frac{(n-1)S^2}{\sigma_0^2} > \chi_\alpha^2(n-1)\right\} \leq P\left\{\frac{(n-1)S^2}{\sigma^2} > \chi_\alpha^2(n-1)\right\} = \alpha$$

对于给定的显著性水平 $\alpha$ ，拒绝域为：

$$V = \left\{ \chi^2 > \chi_\alpha^2(n-1) \right\}$$

检验统计量落入拒绝域的概率 $\leq \alpha$

# 一个正态总体方差的假设检验

构造检验统计量:  $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$

①  $H_0: \sigma^2 = \sigma_0^2$

对于给定的显著性水平 $\alpha$ ，拒绝域为:

$$W = \{ \chi^2 < \chi_{1-\alpha/2}^2(n-1) \text{ 或 } \chi^2 > \chi_{\alpha/2}^2(n-1) \}$$

②  $H_0: \sigma^2 \leq \sigma_0^2$

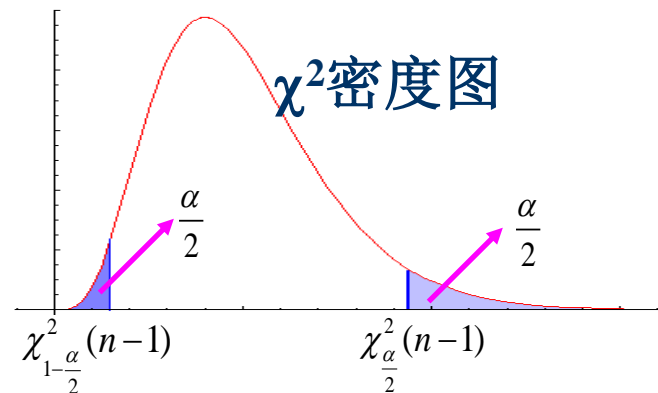
对于给定的显著性水平 $\alpha$ ，拒绝域为:

$$V = \{ \chi^2 > \chi_{\alpha}^2(n-1) \}$$

③  $H_0: \sigma^2 \geq \sigma_0^2$

对于给定的显著性水平 $\alpha$ ，拒绝域为:

$$V = \{ \chi^2 < \chi_{1-\alpha}^2(n-1) \}$$





# 假设检验的计算

**例：**某厂生产的某种型号的电池，其寿命长期以来服从方差  $\sigma^2=5000$ (小时<sup>2</sup>) 的正态分布，现有一批这种电池，从它生产情况来看，寿命的波动性有所变化. 现随机的取26只电池，测出其寿命的样本方差  $S^2=9200$ (小时<sup>2</sup>). 问根据这一数据能否推断这批电池的寿命的波动性较以往的有显著的变化？  $\alpha=0.02$

拒绝域为  $W = \{ \chi^2 < \chi_{1-\alpha/2}^2(n-1) \text{ 或 } \chi^2 > \chi_{\alpha/2}^2(n-1) \}$

**解：**首先建立假设： $H_0: \sigma^2=5000$ ,  $H_1: \sigma^2 \neq 5000$

$$n=26, \quad \alpha=0.02, \quad \sigma_0^2=5000$$

$$\chi_{\alpha/2}^2(n-1) = \chi_{0.01}^2(25) = 44.314, \quad \chi_{1-\alpha/2}^2(n-1) = \chi_{0.99}^2(25) = 11.524$$

$$\text{检验统计量 } \chi^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{25 \times 9200}{5000} = 46 > 44.314$$

所以拒绝 $H_0$ ,认为这批电池寿命的波动性较以往有显著的变化.

# 假设检验的计算

**例：**某自动车床生产的产品尺寸服从正态分布,按规定产品尺寸的方差 $\sigma^2$ 不得超过0.1.为检验该自动车床的工作精度,随机的取25件产品,测得样本方差 $S^2=0.1975$ ,  $\bar{x} = 3.86$ .问该车床生产的产品是否达到所要求的精度?  $\alpha=0.05$

拒绝域为 $W = \{\chi^2 > \chi_{\alpha}^2(n-1)\}$

**解：**首先建立假设:  $H_0: \sigma^2 \leq 0.1$ ,  $H_1: \sigma^2 > 0.1$

$n=25$ ,  $\alpha=0.05$ ,  $\sigma_0^2=0.1$ ,  $\chi_{\alpha}^2(n-1) = \chi_{0.05}^2(24) = 36.415$

检验统计量  $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{24 \times 0.1975}{0.1} = 47.4 > 36.415$

所以拒绝 $H_0$ , 认为该车床生产的产品没有达到所要求的精度.

## 第三节 两个正态总体的假设检验

- 两个正态总体均值的假设检验
- 两个正态总体方差的假设检验

# 两个正态总体的假设检验

➤ 设 $X$ 和 $Y$ 是两个相互独立的正态总体,  $X \sim N(\mu_1, \sigma_1^2)$ ,  $Y \sim N(\mu_2, \sigma_2^2)$ ,  $(X_1, X_2, \dots, X_{n_1})$ 和 $(Y_1, Y_2, \dots, Y_{n_2})$ 为分别来自总体 $X$ 和 $Y$ 的样本, 样本均值与样本方差为:

$$\bar{X} = \frac{1}{n_1} \sum_{i=1}^{n_1} X_i, \quad S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_i - \bar{X})^2$$

$$\bar{Y} = \frac{1}{n_2} \sum_{i=1}^{n_2} Y_i, \quad S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (Y_i - \bar{Y})^2$$

□ 关于总体期望的假设检验:

①  $H_0: \mu_1 = \mu_2$

②  $H_0: \mu_1 \leq \mu_2$

③  $H_0: \mu_1 \geq \mu_2$

□ 关于总体方差的假设检验:

①  $H_0: \sigma_1^2 = \sigma_2^2$

②  $H_0: \sigma_1^2 \leq \sigma_2^2$

③  $H_0: \sigma_1^2 \geq \sigma_2^2$

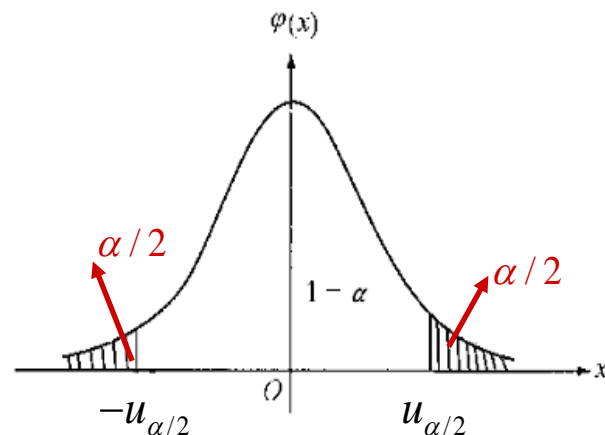
# 两个正态总体均值的假设检验

◆ 当 $\sigma_1^2$ 和 $\sigma_2^2$ 已知时

①  $H_0: \mu_1 = \mu_2, H_1: \mu_1 \neq \mu_2$

由定理6.4可知

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$



当 $H_0$ 成立时，构造检验统计量：
$$U = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

对于给定的显著性水平 $\alpha$ ，拒绝域为：
$$W = \{|U| \geq u_{\alpha/2}\}$$

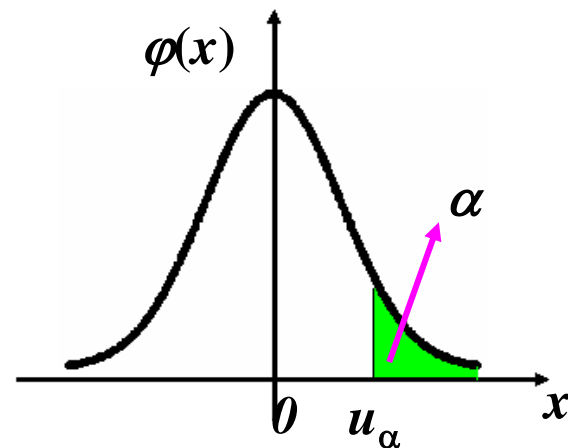
检验统计量落入拒绝域的概率为 $\alpha$

# 两个正态总体均值的假设检验

◆ 当 $\sigma_1^2$ 和 $\sigma_2^2$ 已知时

②  $H_0: \mu_1 \leq \mu_2, H_1: \mu_1 > \mu_2$

由定理6.4可知 
$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$



当 $H_0$ 成立时( $\mu_1 - \mu_2 \leq 0$ ), 构造检验统计量:

有

$$P\left\{\frac{(\bar{X} - \bar{Y})}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > u_\alpha\right\} \leq P\left\{\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > u_\alpha\right\} = \alpha$$

$$U = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

对于给定的显著性水平 $\alpha$ , 拒绝域为:  $W = \{U > u_\alpha\}$

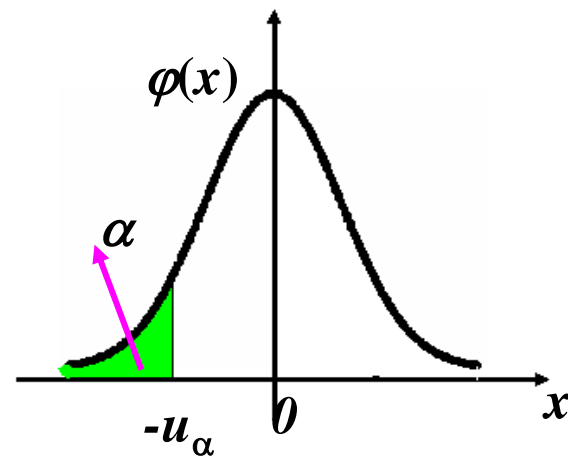
检验统计量落入拒绝域的概率 $\leq \alpha$

# 两个正态总体均值的假设检验

◆ 当 $\sigma_1^2$ 和 $\sigma_2^2$ 已知时

③  $H_0: \mu_1 \geq \mu_2, H_1: \mu_1 < \mu_2$

由定理6.4可知 
$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$



当 $H_0$ 成立时( $\mu_1 - \mu_2 \geq 0$ ), 构造检验统计量:

有 
$$P\left\{\frac{(\bar{X} - \bar{Y})}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < -u_\alpha\right\} \leq P\left\{\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < -u_\alpha\right\} = \alpha$$

$$U = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

对于给定的显著性水平 $\alpha$ , 拒绝域为:  $W = \{U < -u_\alpha\}$

检验统计量落入拒绝域的概率 $\leq \alpha$

# 两个正态总体均值的假设检验

◆ 当 $\sigma_1^2$ 和 $\sigma_2^2$ 未知，但 $\sigma_1^2 = \sigma_2^2$ 时

①  $H_0: \mu_1 = \mu_2, H_1: \mu_1 \neq \mu_2$

由定理6.5可知

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_{\omega} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2) \quad \text{其中 } S_{\omega}^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

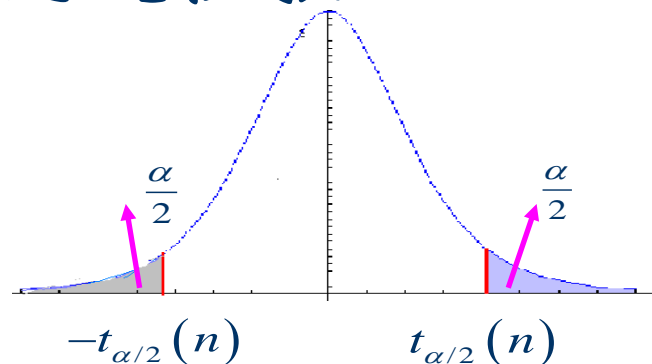
当 $H_0$ 成立时，构造检验统计量：

$$T = \frac{\bar{X} - \bar{Y}}{S_{\omega} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$

对于给定的显著性水平 $\alpha$ ，拒绝域为：

$$W = \{|T| \geq t_{\alpha/2}(n_1 + n_2 - 2)\}$$

检验统计量落入拒绝域的概率为 $\alpha$





# 说明

上面，我们假定  $\sigma_1^2 = \sigma_2^2$ 。当然，这是个不得已而强加上去的条件，因为如果不加此条件，就无法使用简单易行的  $t$  检验。

在实用中，只要有理由认为  $\sigma_1^2$  和  $\sigma_2^2$  相差不是太大，往往就可使用上述方法。通常是：如果方差比检验未被拒绝(见下节)，就认为  $\sigma_1^2$  和  $\sigma_2^2$  相差不是太大。

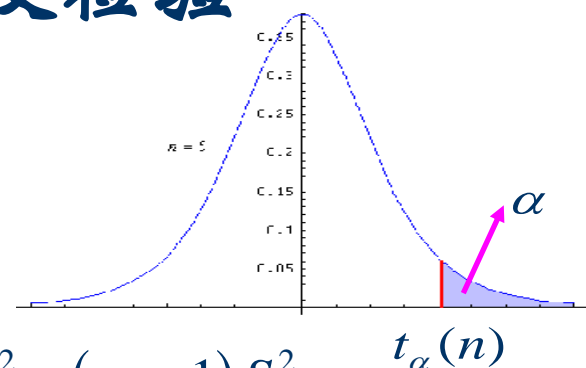
# 两个正态总体均值的假设检验

◆ 当 $\sigma_1^2$ 和 $\sigma_2^2$ 未知，但 $\sigma_1^2 = \sigma_2^2$ 时

②  $H_0: \mu_1 \leq \mu_2$ ,  $H_1: \mu_1 > \mu_2$

由定理6.5可知

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_\omega \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2) \quad \text{其中 } S_\omega^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$



当 $H_0$ 成立时( $\mu_1 - \mu_2 \leq 0$ ), 构造检验统计量:

有

$$P\left\{\frac{(\bar{X} - \bar{Y})}{S_\omega \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} > t_\alpha\right\} \leq P\left\{\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_\omega \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} > t_\alpha\right\} = \alpha$$

$$T = \frac{\bar{X} - \bar{Y}}{S_\omega \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

对于给定的显著性水平 $\alpha$ , 拒绝域为:

检验统计量落入拒绝域的概率 $\leq \alpha$

$$W = \{T > t_\alpha(n_1 + n_2 - 2)\}$$

# 两个正态总体均值的假设检验

◆ 当 $\sigma_1^2$ 和 $\sigma_2^2$ 未知, 但 $\sigma_1^2=\sigma_2^2$ 时

③  $H_0: \mu_1 \geq \mu_2, H_1: \mu_1 < \mu_2$

由定理6.5可知

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_\omega \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2) \quad \text{其中 } S_\omega^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

当 $H_0$ 成立时( $\mu_1 - \mu_2 \geq 0$ ), 构造检验统计量:  $T = \frac{\bar{X} - \bar{Y}}{S_\omega \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

有  $P\left\{\frac{(\bar{X} - \bar{Y})}{S_\omega \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} < -t_\alpha\right\} \leq P\left\{\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_\omega \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} < -t_\alpha\right\} = \alpha$

对于给定的显著性水平 $\alpha$ , 拒绝域为:

检验统计量落入拒绝域的概率 $\leq \alpha$

$$W = \{T < -t_\alpha(n_1 + n_2 - 2)\}$$

	原假设 $H_0$	检验统计量	备择假设 $H_1$	拒绝域
1	$\mu \leq \mu_0$ $\mu \geq \mu_0$ $\mu = \mu_0$ $(\sigma^2 \text{已知})$	$U = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \neq \mu_0$	$U > u_\alpha$ $U < -u_\alpha$ $ U  \geq u_{\alpha/2}$
2	$\mu \leq \mu_0$ $\mu \geq \mu_0$ $\mu = \mu_0$ $(\sigma^2 \text{未知})$	$T = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \neq \mu_0$	$T > t_\alpha(n-1)$ $T < -t_\alpha(n-1)$ $ T  \geq t_{\alpha/2}(n-1)$
3	$\mu_1 \leq \mu_2$ $\mu_1 \geq \mu_2$ $\mu_1 = \mu_2$ $(\sigma_1^2, \sigma_2^2 \text{已知})$	$U = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \neq \mu_0$	$U > u_\alpha$ $U < -u_\alpha$ $ U  \geq u_{\alpha/2}$
4	$\mu_1 \leq \mu_2$ $\mu_1 \geq \mu_2$ $\mu_1 = \mu_2$ $(\sigma_1^2 = \sigma_2^2 = \sigma^2 \text{未知})$	$T = \frac{\bar{X} - \bar{Y}}{S_\omega \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ $S_\omega^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 2)S_2^2}{n_1 + n_2 - 2}$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \neq \mu_0$	$T > t_\alpha(n_1 + n_2 - 2)$ $T < -t_\alpha(n_1 + n_2 - 2)$ $ T  \geq t_{\alpha/2}(n_1 + n_2 - 2)$

# 假设检验的计算

**例：**假设有A和B两种药，欲比较它们在服用2小时后在血液中的含量是否一样。对药品A，随机抽取8个病人服药，服药2小时后，测得8个病人血液中药物浓度(用适当的单位)分别为：

1.23, 1.42, 1.41, 1.62, 1.55, 1.51, 1.60, 1.76

对药品B，随机抽取6个病人服药，服药2小时后，测得血液中药的浓度分别为：

1.76, 1.41, 1.87, 1.49, 1.67, 1.81

假定这两组观测值抽自具有共同方差的两个正态总体，在显著性水平 $\alpha=0.10$ 下，检验病人血液中这两种药的浓度是否有显著不同？

# 假设检验的计算

**解：**问题就是从两个独立总体  $N(\mu_1, \sigma^2)$  和  $N(\mu_2, \sigma^2)$  中分别抽取样本  $(X_1, X_2, \dots, X_8)$  和  $(Y_1, Y_2, \dots, Y_6)$ ，样本均值和样本方差分别为：

$$\bar{X} = 1.51, S_1^2 = 0.03, \bar{Y} = 1.66, S_2^2 = 0.21, \alpha = 0.1$$

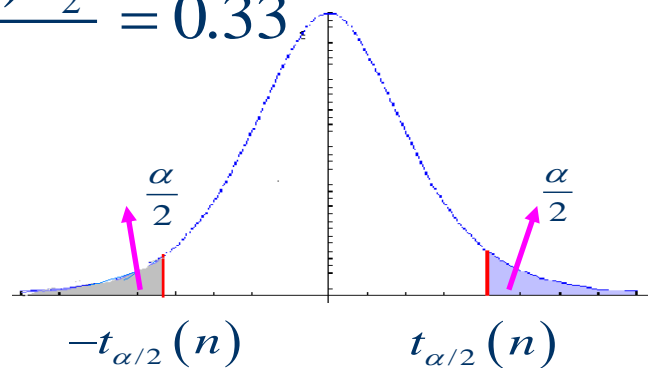
$$n_1 = 8, n_2 = 6, S_w = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} = 0.33$$

建立假设： $H_0: \mu_1 = \mu_2, H_1: \mu_1 \neq \mu_2$

$$T = \frac{\bar{X} - \bar{Y}}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{1.51 - 1.66}{0.33 \sqrt{\frac{1}{8} + \frac{1}{6}}} = -0.842$$

$$t_{\alpha/2}(n_1 + n_2 - 2) = t_{0.05}(12) = 1.7823$$

$$\Rightarrow |T| < t_{\alpha/2}(n_1 + n_2 - 2)$$



在  $\alpha=0.1$  下接受原假设, 认为病人血液中两种药浓度无显著差异

# 假设检验的计算

**例：** 在平炉进行一项试验，以确定改变操作方法的建议是否会增加钢的得率，试验是在同一只平炉上进行的。每炼一炉钢时除操作方法外，其它条件都尽可能做到相同。先用标准方法炼一炉，然后用建议的新方法炼一炉，以后交替进行，各炼了10炉，其得率分别为

标准方法： 78.1 72.4 76.2 74.3 77.4 78.4 76.0 75.5 76.7 77

新方法： 79.1 81.0 77.3 79.1 80.0 79.1 79.1 77.3 80.2 82.1

设这两个样本相互独立，且分别来自正态总体 $N(\mu_1, \sigma^2)$ 和 $N(\mu_2, \sigma^2)$ ， $\mu_1, \mu_2, \sigma^2$ 均未知。问建议的新操作方法能否提高得率？（取 $\alpha=0.05$ ）

# 假设检验的计算

**解：**建立假设： $H_0: \mu_1 \geq \mu_2$ ,  $H_1: \mu_1 < \mu_2$ 。 则

$$n_1 = 10, \bar{X} = 76.23, S_1^2 = 3.325; \quad n_2 = 10, \bar{Y} = 79.43, S_2^2 = 2.225$$

$$S_w = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} = \sqrt{2.775} = 1.666$$

对于给定的 $\alpha$ , 拒绝域为:  $W = \{T < -t_\alpha(n_1 + n_2 - 2)\}$

$$T = \frac{\bar{X} - \bar{Y}}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{76.23 - 79.43}{1.666 \sqrt{\frac{1}{10} + \frac{1}{10}}} = -4.295 \quad \longrightarrow \quad T < -t_\alpha(n_1 + n_2 - 2)$$

$$t_\alpha(n_1 + n_2 - 2) = t_{0.05}(18) = 1.7341$$

在 $\alpha=0.05$ 下拒绝原假设,认为建议的新操作方法较原来的方法为优.



# 两个正态总体的假设检验

➤ 设 $X$ 和 $Y$ 是两个相互独立的正态总体,  $X \sim N(\mu_1, \sigma_1^2)$ ,  $Y \sim N(\mu_2, \sigma_2^2)$ ,  $(X_1, X_2, \dots, X_{n_1})$ 和 $(Y_1, Y_2, \dots, Y_{n_2})$ 为分别来自总体 $X$ 和 $Y$ 的样本, 样本均值与样本方差为:

$$\bar{X} = \frac{1}{n_1} \sum_{i=1}^{n_1} X_i, \quad S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_i - \bar{X})^2$$

$$\bar{Y} = \frac{1}{n_2} \sum_{i=1}^{n_2} Y_i, \quad S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (Y_i - \bar{Y})^2$$

□ 关于总体方差的假设检验:

①  $H_0: \sigma_1^2 = \sigma_2^2$

②  $H_0: \sigma_1^2 \leq \sigma_2^2$

③  $H_0: \sigma_1^2 \geq \sigma_2^2$

# 两个正态总体方差的假设检验

◆  $H_0: \sigma_1^2 = \sigma_2^2$ ,  $H_1: \sigma_1^2 \neq \sigma_2^2$

思路分析:

因两总体  $N(\mu_1, \sigma_1^2)$  和  $N(\mu_2, \sigma_2^2)$  的样本方差  $S_1^2$  和  $S_2^2$  分别为  $\sigma_1^2$  和  $\sigma_2^2$  的无偏估计。所以, 直观上讲,  $S_1^2/S_2^2$  是  $\sigma_1^2/\sigma_2^2$  的一个好的估计。

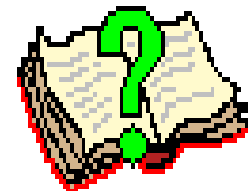
当  $H_0: \sigma_1^2 = \sigma_2^2$  成立时,  $\sigma_1^2/\sigma_2^2 = 1$ , 作为其估计,  $S_1^2/S_2^2$  也应与1相差不大。当该值过分地大或过分地小时, 都应拒绝原假设。

合理的思路是: 找两个界限  $c_1$  和  $c_2$ , 使得

- 当  $c_1 < S_1^2/S_2^2 < c_2$  时, 接受  $H_0$ ;
- 当  $S_1^2/S_2^2 \leq c_1$ , 或  $S_1^2/S_2^2 \geq c_2$  时, 拒绝  $H_0$ 。

# 两个正态总体方差的假设检验

问题是 $c_1$ 与 $c_2$ 如何确定?

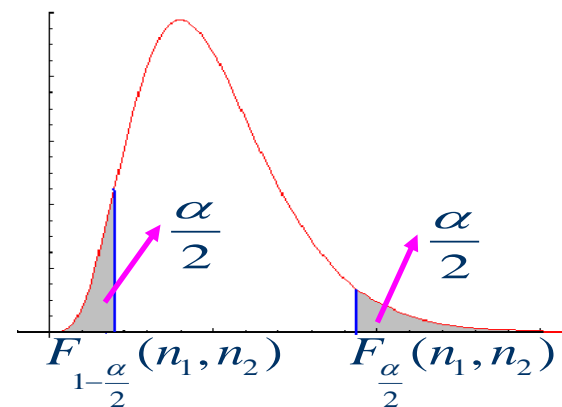


由定理6.6可知

$$\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F(n_1 - 1, n_2 - 1)$$

当  $H_0: \sigma_1^2 = \sigma_2^2$  成立时, 构造检验统计量

$$F = S_1^2 / S_2^2 \sim F(n_1 - 1, n_2 - 1)$$



故  $H_0$  的拒绝域为

$$W = \left\{ \frac{S_1^2}{S_2^2} \leq F_{1-\alpha/2}(n_1 - 1, n_2 - 1) \text{ 或 } \frac{S_1^2}{S_2^2} \geq F_{\alpha/2}(n_1 - 1, n_2 - 1) \right\}$$

检验统计量落入拒绝域的概率为 $\alpha$

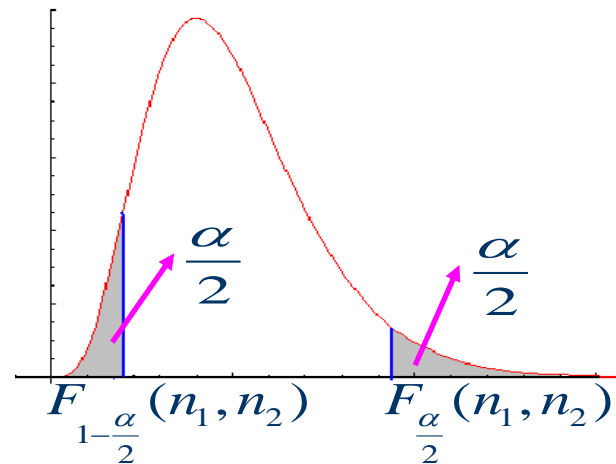
上述检验法称为 **F 检验法**

# 两个正态总体方差的假设检验

①  $H_0: \sigma_1^2 = \sigma_0^2; H_1: \sigma_1^2 \neq \sigma_0^2$

当 $H_0$ 成立时，构造检验统计量：

$$F = \frac{S_1^2}{S_2^2} \sim F(n_1 - 1, n_2 - 1)$$



对于给定的显著性水平 $\alpha$ ，拒绝域为：

$$W = \{F \leq F_{1-\alpha/2}(n_1 - 1, n_2 - 1) \text{ 或 } F \geq F_{\alpha/2}(n_1 - 1, n_2 - 1)\}$$

检验统计量落入拒绝域的概率为 $\alpha$

# 两个正态总体方差的假设检验

②  $H_0: \sigma_1^2 \leq \sigma_2^2; H_1: \sigma_1^2 > \sigma_2^2$

由定理6.6可知:

$$\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F(n_1 - 1, n_2 - 1)$$

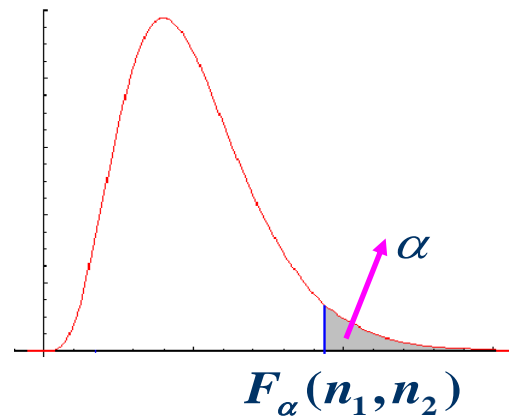
当 $H_0$ 成立时, 构造检验统计量:  $F = \frac{S_1^2}{S_2^2}$

$$P\left\{\frac{S_1^2}{S_2^2} > F_\alpha(n_1 - 1, n_2 - 1)\right\} \leq P\left\{\frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} > F_\alpha(n_1 - 1, n_2 - 1)\right\} = \alpha$$

对于给定的显著性水平 $\alpha$ , 拒绝域为:

$$W = \{F > F_\alpha(n_1 - 1, n_2 - 1)\}$$

检验统计量落入拒绝域的概率 $\leq \alpha$



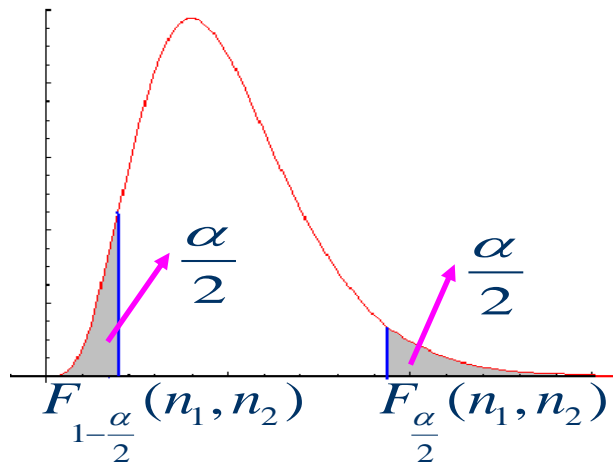
# 两个正态总体方差的假设检验

③  $H_0: \sigma_1^2 \geq \sigma_2^2; H_1: \sigma_1^2 < \sigma_2^2$

由定理6.6可知:

$$\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F(n_1 - 1, n_2 - 1)$$

当 $H_0$ 成立时, 构造检验统计量:  $F = \frac{S_1^2}{S_2^2}$



$$P\left\{\frac{S_1^2}{S_2^2} < F_{1-\alpha}(n_1 - 1, n_2 - 1)\right\} \leq P\left\{\frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} < F_{1-\alpha}(n_1 - 1, n_2 - 1)\right\} = \alpha$$

对于给定的显著性水平 $\alpha$ , 拒绝域为:

$$W = \left\{ F < F_{1-\alpha}(n_1 - 1, n_2 - 1) \right\}$$

检验统计量落入拒绝域的概率 $\leq \alpha$

# 两个正态总体方差的假设检验

构造检验统计量:  $F = \frac{S_1^2}{S_2^2}$

①  $H_0: \sigma_1^2 = \sigma_2^2$

对于给定的显著性水平 $\alpha$ ，拒绝域为：

$$W = \{F \leq F_{1-\alpha/2}(n_1-1, n_2-1) \text{ 或 } F \geq F_{\alpha/2}(n_1-1, n_2-1)\}$$

②  $H_0: \sigma_1^2 \leq \sigma_2^2$

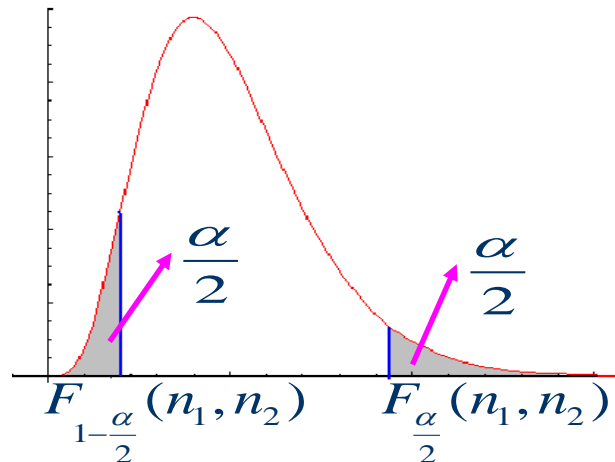
对于给定的显著性水平 $\alpha$ ，拒绝域为：

$$W = \{F > F_{\alpha}(n_1-1, n_2-1)\}$$

③  $H_0: \sigma_1^2 \geq \sigma_2^2$

对于给定的显著性水平 $\alpha$ ，拒绝域为：

$$W = \{F < F_{1-\alpha}(n_1-1, n_2-1)\}$$



# 假设检验的计算

**例：** 甲乙两厂生产同一种电阻，现从甲乙两厂的产品中分别随机地抽取12个和10个样品,测得它们的电阻值后，计算出样本方差分别为 $S_1^2=1.40$ ， $S_2^2=4.38$ 。假设两厂生产的电阻的阻值分别服从正态分布 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$ 。

在显著性水平  $\alpha = 0.10$  下, 是否可接受:

(1)  $\sigma_1^2 = \sigma_2^2$ ;                      (2)  $\sigma_1^2 \leq \sigma_2^2$



# 假设检验的计算

**解：** 检验统计量为 $F = S_1^2/S_2^2$

$$F_{1-\alpha}(n_1, n_2) = \frac{1}{F_{\alpha}(n_2, n_1)}$$

$$n_1=12, n_2=10, \alpha=0.10, S_1^2=1.40, S_2^2=4.38, S_1^2/S_2^2=0.32$$

(1) 建立假设： $H_0: \sigma_1^2 = \sigma_2^2$ ;  $H_1: \sigma_1^2 \neq \sigma_2^2$

拒绝域为： $W = \{F \leq F_{1-\alpha/2}(n_1-1, n_2-1) \text{ 或 } F \geq F_{\alpha/2}(n_1-1, n_2-1)\}$

$$F_{1-\alpha/2}(n_1-1, n_2-1) = F_{0.95}(11, 9) = \frac{1}{F_{0.05}(9, 11)} = \frac{1}{2.90} = 0.34$$

因 $F = S_1^2/S_2^2 = 0.32 < 0.34$ ，所以，无须再考虑 $F_{\alpha/2}(n_1-1, n_2-1)$ 的值，就可得到拒绝 $\sigma_1^2 = \sigma_2^2$ 的结论。

# 假设检验的计算

**解：** 检验统计量为 $F = S_1^2/S_2^2$

$$n_1=12, n_2=10, \alpha=0.10, S_1^2=1.40, S_2^2=4.38, S_1^2/S_2^2=0.32$$

(2) 建立假设： $H_0: \sigma_1^2 \leq \sigma_2^2$ ;  $H_1: \sigma_1^2 > \sigma_2^2$

拒绝域为： $V = \{ F > F_{\alpha}(n_1 - 1, n_2 - 1) \}$

查F分布表，因查不到 $F_{0.10}(11,9)$ ，改用 $F_{0.10}(12,9)$ 和 $F_{0.10}(10,9)$ 的平均值近似之，得

$$\begin{aligned} F_{11,9}(0.10) &\approx [F_{0.10}(10,9) + F_{0.10}(12,9)]/2 \\ &\approx [2.42 + 2.38]/2 = 2.40 \end{aligned}$$

因 $F = S_1^2/S_2^2 = 0.32 < 2.40$ ，故接受假设 $\sigma_1^2 \leq \sigma_2^2$ 的假设

# 小 结

本章介绍了假设检验的基本概念和步骤；  
并在正态总体的条件下，对一个总体及两个总体的均值和方差进行了假设检验。

# 正态总体均值的假设检验

	原假设 $H_0$	检验统计量	备择假设 $H_1$	拒绝域
1	$\mu \leq \mu_0$ $\mu \geq \mu_0$ $\mu = \mu_0$ $(\sigma^2 \text{已知})$	$U = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \neq \mu_0$	$U > u_\alpha$ $U < -u_\alpha$ $ U  \geq u_{\alpha/2}$
2	$\mu \leq \mu_0$ $\mu \geq \mu_0$ $\mu = \mu_0$ $(\sigma^2 \text{未知})$	$T = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \neq \mu_0$	$T > t_\alpha(n-1)$ $T < -t_\alpha(n-1)$ $ T  \geq t_{\alpha/2}(n-1)$
3	$\mu_1 \leq \mu_2$ $\mu_1 \geq \mu_2$ $\mu_1 = \mu_2$ $(\sigma_1^2, \sigma_2^2 \text{已知})$	$U = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \neq \mu_0$	$U > u_\alpha$ $U < -u_\alpha$ $ U  \geq u_{\alpha/2}$
4	$\mu_1 \leq \mu_2$ $\mu_1 \geq \mu_2$ $\mu_1 = \mu_2$ $(\sigma_1^2 = \sigma_2^2 = \sigma^2 \text{未知})$	$T = \frac{\bar{X} - \bar{Y}}{S_\omega \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ $S_\omega^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 2)S_2^2}{n_1 + n_2 - 2}$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \neq \mu_0$	$T > t_\alpha(n_1 + n_2 - 2)$ $T < -t_\alpha(n_1 + n_2 - 2)$ $ T  \geq t_{\alpha/2}(n_1 + n_2 - 2)$

# 正态总体方差的假设检验

	原假设 $H_0$	检验统计量	备择假设 $H_1$	拒绝域
5	$\sigma^2 \leq \sigma_0^2$ $\sigma^2 \geq \sigma_0^2$ $\sigma^2 = \sigma_0^2$ ( $\mu$ 未知)	$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$	$\sigma^2 > \sigma_0^2$ $\sigma^2 < \sigma_0^2$ $\sigma^2 \neq \sigma_0^2$	$\chi^2 > \chi_\alpha^2(n-1)$ $\chi^2 < \chi_{1-\alpha}^2(n-1)$ $\chi^2 \geq \chi_{\alpha/2}^2(n-1)$ 或 $\chi^2 \leq \chi_{1-\alpha/2}^2(n-1)$
6	$\sigma_1^2 \leq \sigma_2^2$ $\sigma_1^2 \geq \sigma_2^2$ $\sigma_1^2 = \sigma_2^2$ ( $\mu_1, \mu_2$ 未知)	$F = \frac{S_1^2}{S_2^2}$	$\sigma_1^2 > \sigma_2^2$ $\sigma_1^2 < \sigma_2^2$ $\sigma_1^2 \neq \sigma_2^2$	$F > F_\alpha(n_1-1, n_2-1)$ $F < F_{1-\alpha}(n_1-1, n_2-1)$ $F \geq F_{\alpha/2}(n_1-1, n_2-1)$ 或 $F \leq F_{1-\alpha/2}(n_1-1, n_2-1)$