

4

质量管理教程

质量变异 及其统计规律

CHAPTER OUTLINE

- 4.1 质量的变异
- 4.2 质量数据的收集整理
- 4.3 质量变异数字特征的度量
- 4.4 质量变异的统计规律
- 4.5 中心极限定理



QUALITY MANAGEMENT

4.1 质量的变异

- 质量的统计观点
 - 质量具有变异性
 - 质量变异具有统计规律性



2

影响工序质量的基本因素

■ 影响工序质量的六个基本因素（6M）

- 人(Manpower)
- 机器(Machinery)
- 材料(Material)
- 方法(Method)
- 测量(Measurement)
- 环境(Mother-natured)



3

随机性变异与系统性变异

■ 6M所导致的变异有两类

- 随机性变异：变异的出现是随机的，无规律
- 系统性变异（或称系统性变异，特殊性变异）：变异的出现是有规律的，可以追溯变异的原因

随机性变异

1. PCB板上随机出现焊点缺陷
2. 航班到达时间正常在8:00-8:30之间
3. 材料配比间的随机波动
4. 车削加工轴的直径正常波动

系统性变异

1. PCB板上固定位置出现焊点缺陷
2. 航班由于大雾到达时间延误5个小时
3. 工人将材料配方看错造成配比严重错误
4. 刀具磨损造成直径偏大



4

4.2 质量数据的收集整理

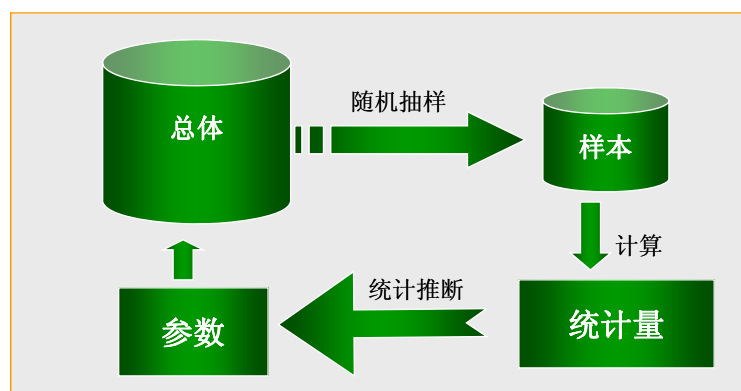
■ 质量数据的分类

- 计量型：连续且可测量的数据
- 计数型：不可测量，只能根据属性判定
 - 计件值
 - 计点值



基本术语

- 统计：是研究数据的科学，可分为描述性统计和推断性统计



- **总体**：研究对象的全体；可分为有限总体与无限总体
- **样本**：总体的一个部分，若抽样过程中总体中的每一个个体都有相等的机会被抽取，称为随机抽样
- **参数**：对总体特性的描述，对于固定的总体，**是常数**
- **统计量**：根据样本计算出的样本统计变量，对于不同的样本，统计量**是随机变量**



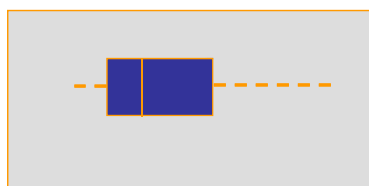
7

质量数据的整理与展示

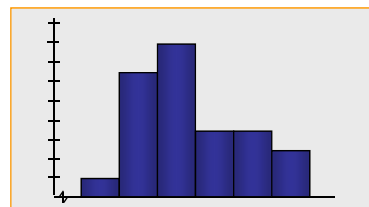
```

5 2 7
6 2 2 2 2 5 6 7 8 8 8 9 9 9
7 1 1 2 2 3 4 4 5 5 5 6 7 8 9 9 9
8 0 0 2 3 5 8 9
9 1 3 7 7 7 8 9
10 1 4 5 5 9

```



月工资	频数
1100—1200	2
1000—1100	5
900—1000	8
800—900	17
700—800	10
600—700	5
500—600	3



8

茎叶图

52 57 62 62 62 62 65 66 67 68 68 68 69
69 69 71 71 72 72 73 74 74 75 75 75 76
77 78 79 79 79 80 80 82 83 85 88 89 91
93 97 97 97 98 99 101 104 105 105 109

```

5 | 2 7
6 | 2 2 2 2 5 6 7 8 8 8 9 9 9
7 | 1 1 2 2 3 4 4 5 5 5 6 7 8 9 9 9
8 | 0 0 2 3 5 8 9
9 | 1 3 7 7 7 8 9
10 | 1 4 5 5 9

```



9

盒图

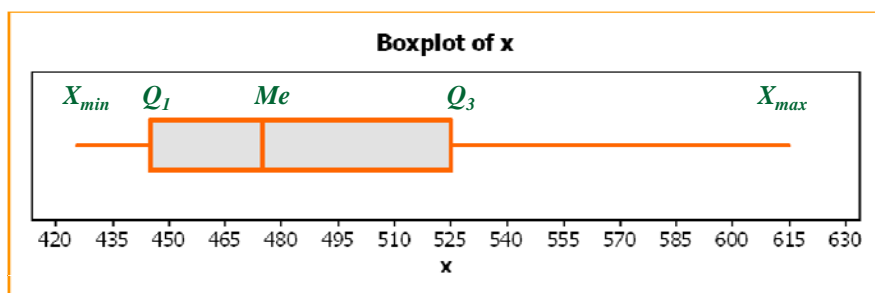
- $X_{min} = 425$ $X_{max} = 615$
- $Q_1 = X_{0.25} = X_{(n+1)/4} = 450$
- $Me = 475$
- $Q_3 = X_{0.75} = X_{3(n+1)/4} = 525$
- 数据按从小到大排列如下表

425	430	430	435	435	435	435	435	440	440
440	440	440	445	445	445	445	445	450	450
450	450	450	450	450	460	460	460	465	465
465	470	470	472	475	475	475	480	480	480
480	485	490	490	490	500	500	500	500	510
510	515	525	525	525	535	549	550	570	570
575	575	580	590	600	600	600	600	615	615



10

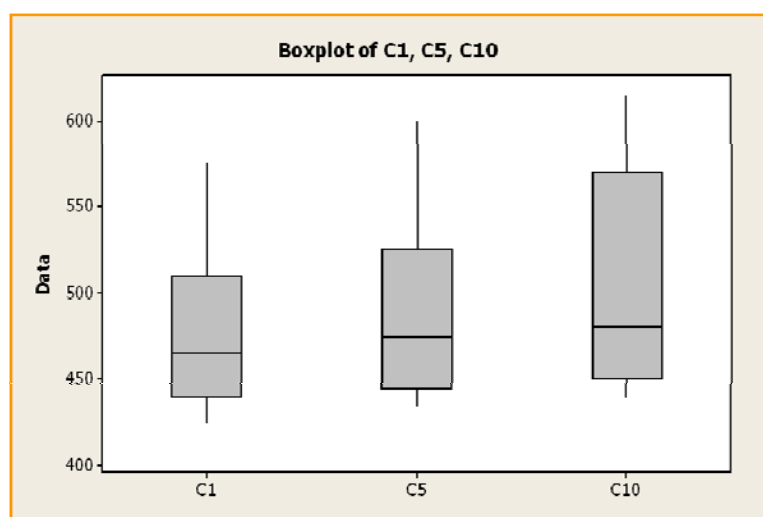
盒图的绘制



天津大学

11

多变量的比较



天津大学

12

直方图

- 建立频数分布表
- 根据频数分布表绘制直方图
- 分析直方图并能根据对直方图的分析采取纠偏措施



13

4.3 质量变异数字特征的度量

- 数据集中程度的度量
 - 平均数
 - 中位数
 - 众数
- 数据的离散程度
 - 极差
 - 方差
 - 标准差

μ σ %
 \bar{x}



14

数据集中心程度的度量

平均数

$$\mu = \frac{\sum X_i}{N} \quad (\text{总体})$$

$$\bar{X} = \frac{\sum X_i}{n} \quad (\text{样本})$$

$$\bar{X} = \frac{\sum f_i X_i}{n} \quad (\text{加权式})$$

中位数

众数



天津大学

15

数据的离散程度

■ 极差 $R = \text{最大值} - \text{最小值} = X_{\max} - X_{\min}$

■ 方差 $\sigma^2 = \frac{\sum (X_i - \mu)^2}{N} \quad (\text{总体})$

$$s^2 = \frac{\sum (X_i - \bar{X})^2}{n-1} \quad (\text{样本})$$

标准差

$$\sigma = \sqrt{\frac{\sum (X_i - \mu)^2}{N}} \quad (\text{总体})$$

$$s = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}} \quad (\text{样本})$$



天津大学

16

小样本条件下

$$\frac{R}{d_2} = \hat{\sigma}$$

d_2 是和样本含量有关的系数



17

四分位间距与变异系数

- 四分位间距 (IQR, Interquartile Range)

$$IQR = X_{0.75} - X_{0.25} = Q_3 - Q_1$$

- 变异系数 (CV, Coefficient of Variation)

$$CV = \frac{\sigma}{|\mu|}$$

$$\gamma = \frac{S}{|X|}$$



18

四分位点值的计算

$$Q_1 = X_{0.25} = X^{(q1)}$$

$$Q_3 = X_{0.75} = X^{(q3)}$$

$$q1 = \frac{(n+1)}{4}$$

$$q3 = \frac{3(n+1)}{4}$$

可以采用线性插值

的方法计算 $X^{(q1)}, X^{(q3)}$



19

数据分布的形状

偏度 (Skewness)

$$\beta_3 = \frac{n}{(n-1)(n-2)} \sum_{i=1}^n (X_i - \bar{X})^3 / S^3$$

>0, 右偏; <0左偏;

注意: 等于零不等于是对称分布。

峰度 (Kurtosis)

$$\beta_4 = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \frac{\sum_{i=1}^n (x - \bar{x})^4}{S^4} - \frac{3(n-1)^2}{(n-2)(n-3)}$$

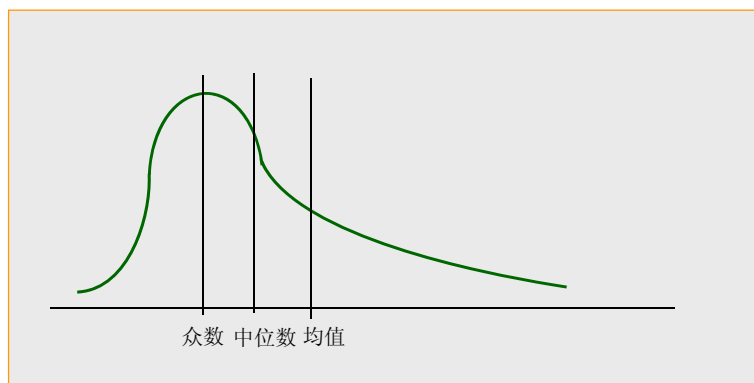
>0, 尖峰; 比正态分布更陡, 尾平;

<0, 比正态分布平, 比正态分布尾细

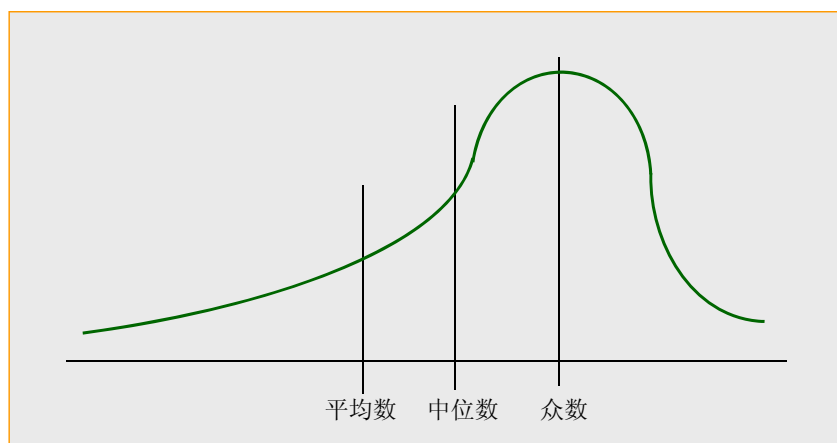


20

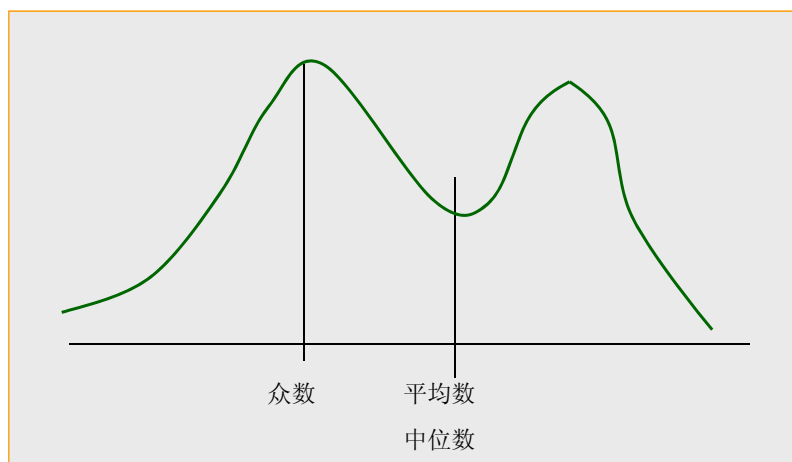
右偏态情形下分布集中程度与离散程度间的关系



左偏态情形下分布集中程度与离散程度间的关系



双峰型分布下分布集中程度与离散程度间的关系



4.4 质量变异的统计规律

- 概率的计算
- 二项式分布（贝努力分布）
- 泊松分布
- 正态分布

超几何分布

问题：若有一批产品，批量 $N=10000$ 件，已知其中有不良品 $D=500$ 件；若从中随机抽取 $n=100$ 件，问其中有 $d=3$ 件不良品的概率是多少？



25

$$P(x = d) = \frac{C_D^d C_{N-D}^{n-d}}{C_N^n}$$

$$E(x) = n \frac{D}{N}$$

$$D(x) = \left(\frac{N-n}{N-1} \right) n \frac{D}{N} \left(1 - \frac{D}{N} \right)$$



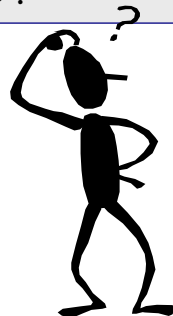
26

两个离散分布

- 二项式分布 (Binomial Distribution)
- 泊松分布 (Poisson Distribution)

若根据大量抽样测得某产品的单位缺陷 (DPU) = 1, 今从大量的该产品中抽取一件, 问该产品没有缺陷的概率为多少?

若供应商提供一批产品, 已知不良率为 p , 从中随机抽取 n 个, 问恰好有 d 个不良品的概率?



天津大学

27

二项式分布 (Binomial Distribution)

- - 实验次数固定
- - 每次实验相互独立
- - 每次实验结果只有二个
- - 每次实验概率保持不变

$$P(x = d) = C_n^d p^d (1 - p)^{n-d}$$

注: $C_n^d = \frac{n!}{d!(n-d)!}$



天津大学

28

泊松分布(Poisson Distribution)

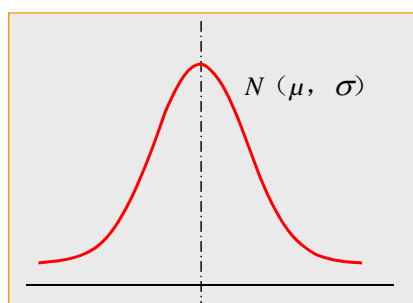
- -缺陷的出现是随机的
- -缺陷出现的概率与“面积”大小成正比
- -在趋于无穷小的一块“面积”上，出现两个或两个以上缺陷的机会为零

$$P(x = d) = \frac{\lambda^d e^{-\lambda}}{d!}$$



29

正态分布

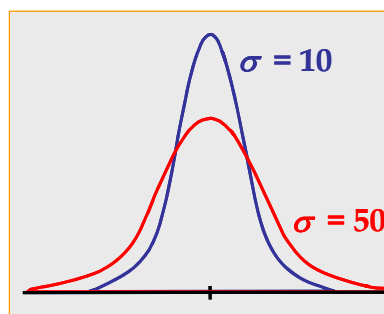


$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$$

其中: $\pi = 3.14159$
 $e = 2.71828$

特点:

- 1、对称
- 2、钟型
- 3、具有两个参数: μ 和 σ



30

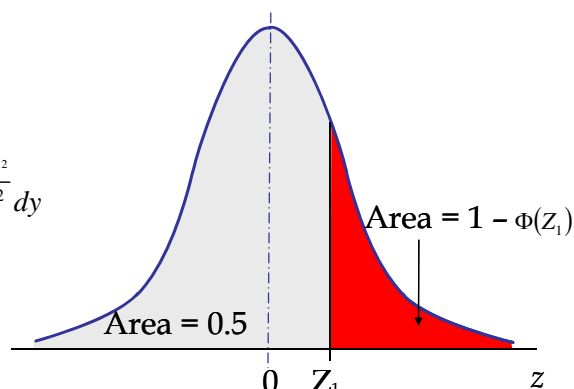
标准正态分布

- 随机变量X服从均值为0，标准差为1的正态分布
- 可将一般的正态分布转化为标准正态分布：

$$z = \frac{x - \mu}{\sigma}$$

- 查标准正态分布表

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{y^2}{2}} dy$$

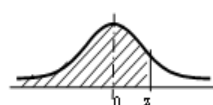


天津大学

31

查标准正态分布表

$$\Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54379	0.54776	0.55172	0.55567	0.55962	0.56396	0.56749	0.57142	0.57534
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62551	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68438	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.72226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75803	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78523
0.8	0.78814	0.79103	0.79389	0.79673	0.79954	0.80234	0.80510	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83397	0.83646	0.83891
1.0	0.84134	0.84375	0.84613	0.84849	0.85083	0.85314				0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87285	0.87493				0.88297
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435				0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149				0.91773
1.4	0.91924	0.92073	0.92219	0.92364	0.92506	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053				0.95448
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994				0.96327
1.8	0.96407	0.96485	0.96562	0.96637	0.96711	0.96783				0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670



天津大学

32

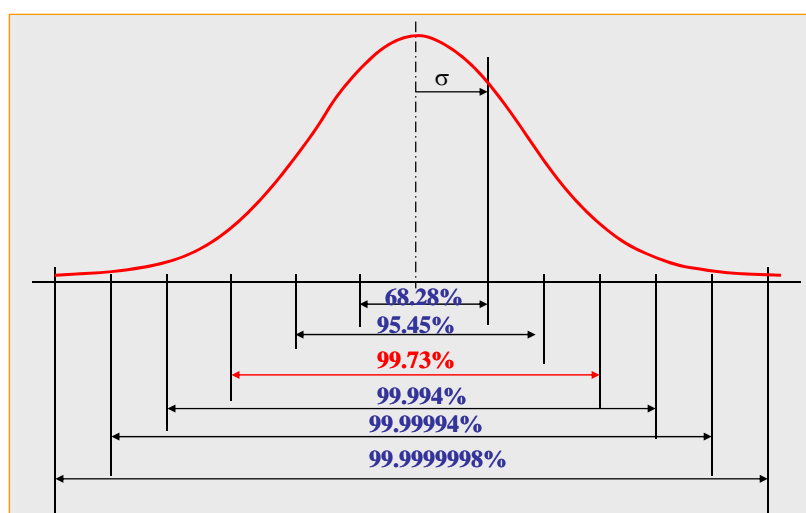
推断正态分布的参数

	总体参数		样本统计量
集中程度	μ	←	\bar{X}
离散程度	σ	←	S



33

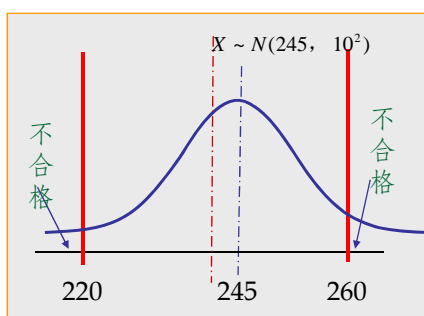
正态分布的概率



34

思考题:

某铜管生产工序技术要求铜管的抗拉强度为220-260kg, 假定该工序所生产的铜管抗拉强度服从正态分布, 均值为245kg, 标准差为10kg, 问该工序的缺陷率水平?



35

4.5 中心极限定理

1、若 X_1, X_2, \dots, X_n 是独立同分布的随机变量

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

当 n 较大时 \bar{X} 趋于正态分布

2、均值 (\bar{X}) 分布的标准差 $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$

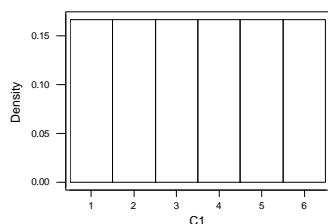
3、均值 (\bar{X}) 分布的中心与总体分布中心相同



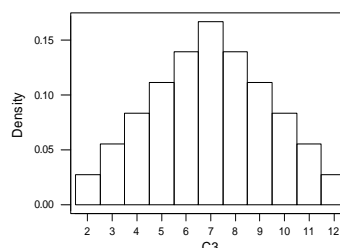
36

中心极限定理得演示

1. 通过掷骰子演示中心极限定理
2. 软件模拟中心极限定理



一个骰子



两个骰子



37

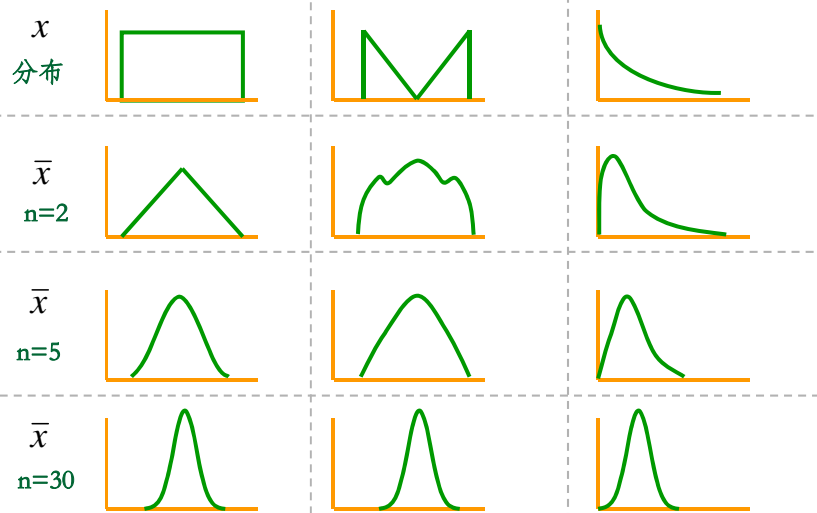
Minitab辅助模拟

- 在Minitab中定义 X_i 为 $[0, 1]$ 区间上的均匀分布，演示中心极限定理。 $C1-C10$ 均为 $[0, 1]$ 区间上的均匀分布产生的随机数， $C11$ 、 $C12$ 、 $C13$ 分别为前2、5、10列的均值。



38

样本均值的分布



39

本章完，谢谢各位！

地 址：天津大学管理学院
邮 编：300072
Email: shi@tju.edu.cn

天津大学质量管理课程组