

统计基础

常州企业管理中心 www.CZQYGL.COM

XQ Huang

Mar-20-03

什么是统计？

统计 就是通过**样本**特性**推断**总体特性的过程



类似于赌博，有一定的风险。

可信度受取样方法，样本大小等因素的影响。

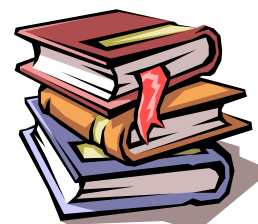
人类认知外部的演变过程



魔术



艺术



科学（统计）

统计是科学的

统计学术语和定义

总体---要研究的具有相同特性的对象的全部。

总体中（单位产品）的个数用 “**N**” 表示

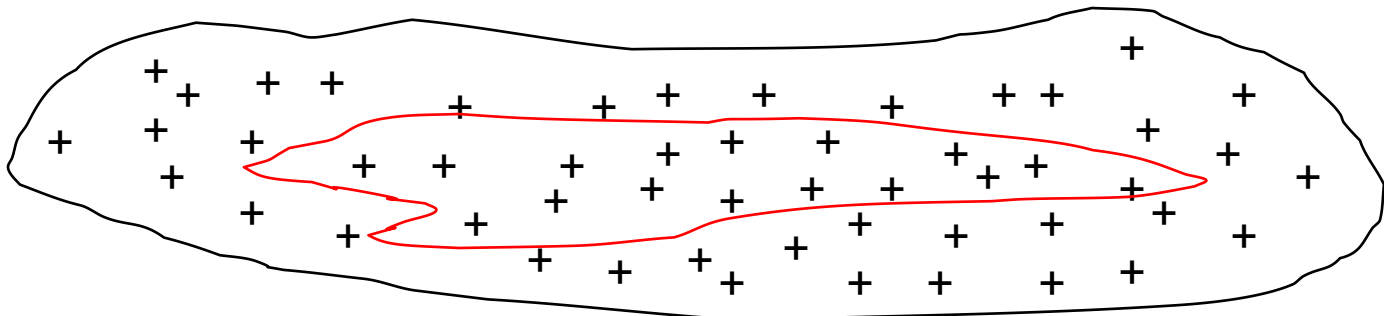
样本---总体中的一个子集。

通常是按一定的抽样方法从总体中抽出用于测量、分析，从而推断总体特性的一组样品。其个数用 “**n**” 表示

样本数（样本大小）---样本中样品的个数 n 。（下图中 $n = ?$ ）

样本越大，越能反应总体。但测量分析的工作量越大。

你能举例说明样本与总体吗？



样本是总体的一部分，是总体的代表。我们通常测量样本来推测总体。

统计学术语和定义

平均值--- $\bar{X}=(X_1+X_2+\cdots+X_n)/n$ 。

方差---数据与其平均值之间差值的平方的平均值。

总体的方差用 δ^2 ; 样本的方差用 S^2 表示。

标准差(均方差)---方差的平方根. 它是衡量数据分散程度的一个指标

$$S = \sqrt{((X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2) / (n-1)}$$

例子

给定一个样本{2, 6, 4,}

$$\bar{X} = (2+6+4)/3 = 4.0$$

$$S = \sqrt{((2-4)^2 + (6-4)^2 + (4-4)^2) / (3-1)} = 2$$



标准差越大,数据越分散

统计学术语和定义

极差(Range)--- $R=(X_{\max}+X_{\min})$

极差是最大值与最小值之间的差异.

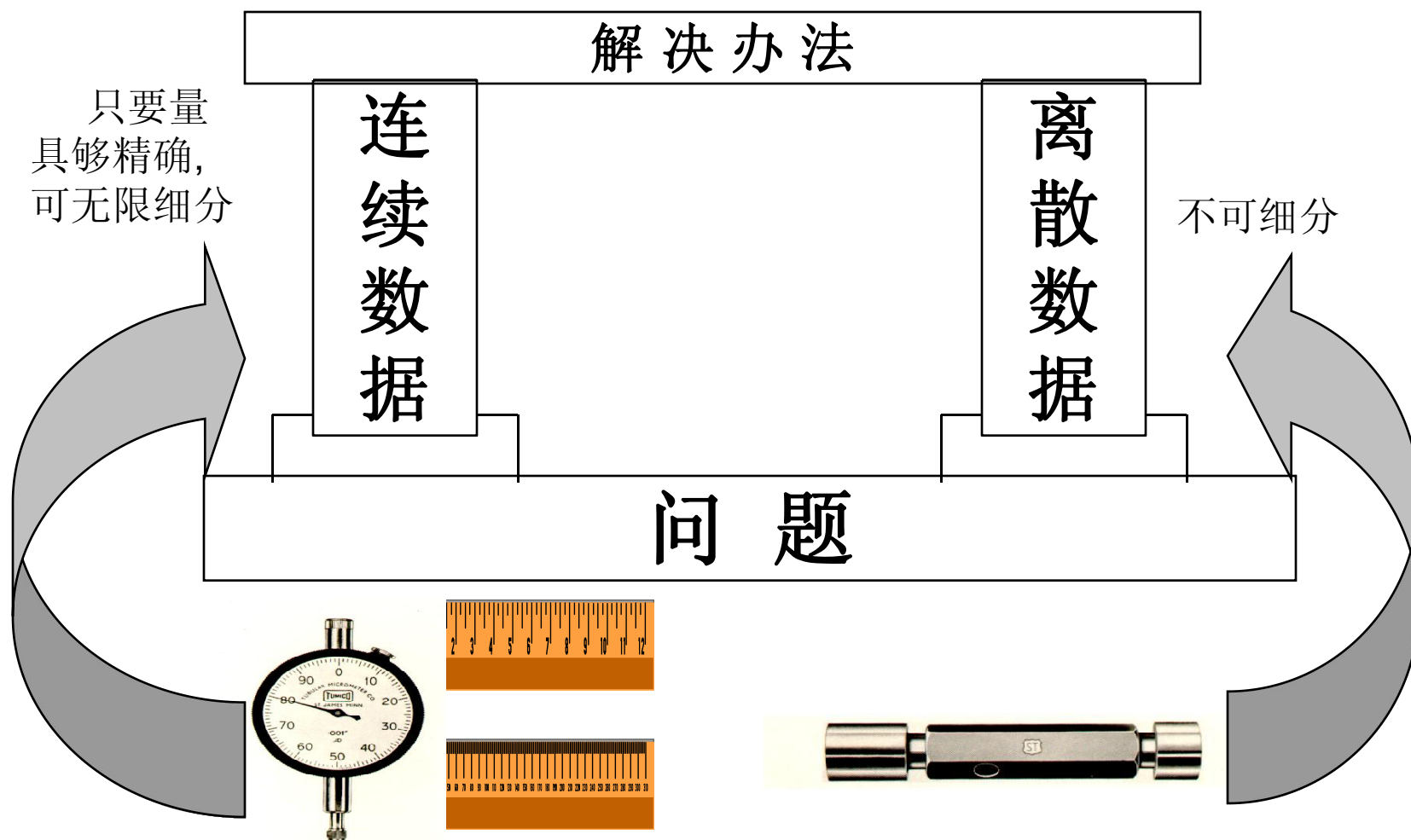


为什么通常用标准差,而不用极差衡量数据的分散程度?



标准差反应总体,极差反应个别

数据类型



不同的数据类型有不同的处理方法

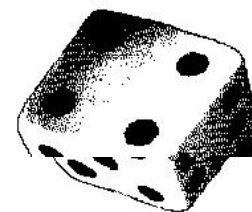
数据类型

你们在自己的区域能接触到什么类型的
数据？

(请各举出五个例子)

概率（可能性）

- 最大概率是多少？
- 什么意思？
- 最小概率是多少？
- 什么意思？
- 0.5的概率意味什么？



概率即可能性

概率（可能性）-练习

答案

- 最大概率是多少？1.0 = 100%
什么意思？
- 最小概率是多少？0.0 = 0%
什么意思？
- 0.5的概率意味什么？50% (掷硬币)



小概率事件是不可能发生的

可能性（概率）

客户要求

什么是制程能力？

什么是符合规格要求的可能性（概率）？

制程能力和可能性的关系？

实际问题的描述 ...

以用两个骰子掷数为例，假如客户的要求是掷到3，4，5，6，7，8，9，10，11中的一个。



风险计算- 统计问题描述

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

	1	2	3	4	5	6
1	0.0278	0.0278	0.0278	0.0278	0.0278	0.0278
2	0.0278	0.0278	0.0278	0.0278	0.0278	0.0278
3	0.0278	0.0278	0.0278	0.0278	0.0278	0.0278
4	0.0278	0.0278	0.0278	0.0278	0.0278	0.0278
5	0.0278	0.0278	0.0278	0.0278	0.0278	0.0278
6	0.0278	0.0278	0.0278	0.0278	0.0278	0.0278

值	组合	概率
2	1	0.0278
3	2	0.0556
4	3	0.0833
5	4	0.1111
6	5	0.1389
7	6	0.1667
8	5	0.1389
9	4	0.1111
10	3	0.0833
11	2	0.0556
12	1	0.0278
合计	36	1

掷到“2”的可能性

1

in

36

=

0.0278

掷到“12”的可能性

1

in

36

=

0.0278

不符合客户要求的可能性



0.0556

第一个骰子掷到任意一个值的概率=1/6=0.1667

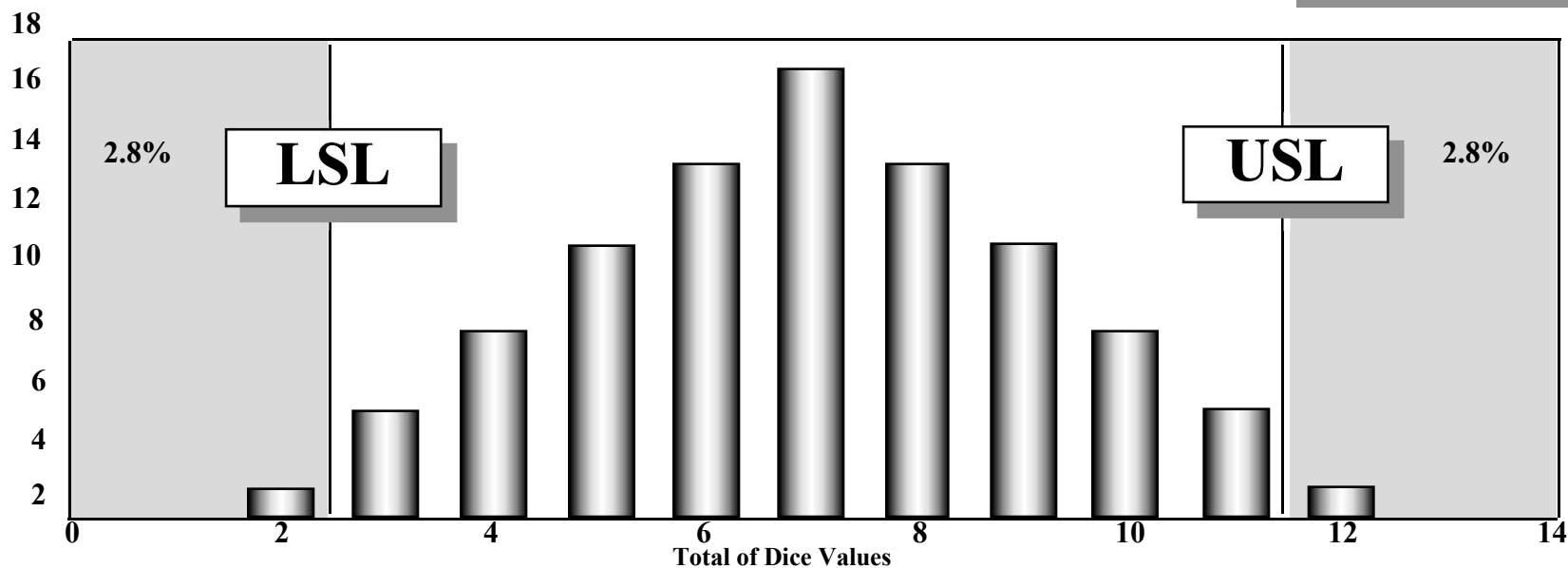
第二个骰子掷到任意一个值的概率=1/6=0.1667

两个骰子掷到任意一个组合的概率=1/6*1/6=1/36=0.0278

用图形将结果表达出来

假如客户的要求是掷到3, 4, 5, 6, 7, 8, 9, 10, 11中的一个

Value	Combinations	Probability
2	1	.0278
3	2	.0556
4	3	.0833
5	4	.1111
6	5	.1389
7	6	.1667
8	5	.1389
9	4	.1111
10	3	.0833
11	2	.0556
12	1	.0278
Total	36	1.0000



满足客户要求的区域占 94.4%

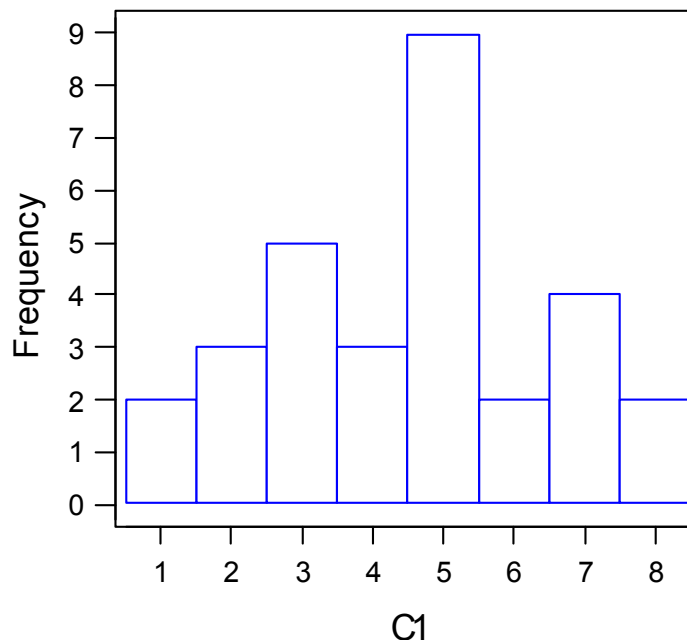
...因此满足客户要求的概率是 94.4 %

统计分布图

- 我们可以通过统计把一个制程或一个总体的状态用图形的方式表示出来。

- 图形的类别:

- 点图
- 柱状图
- 正态分布图



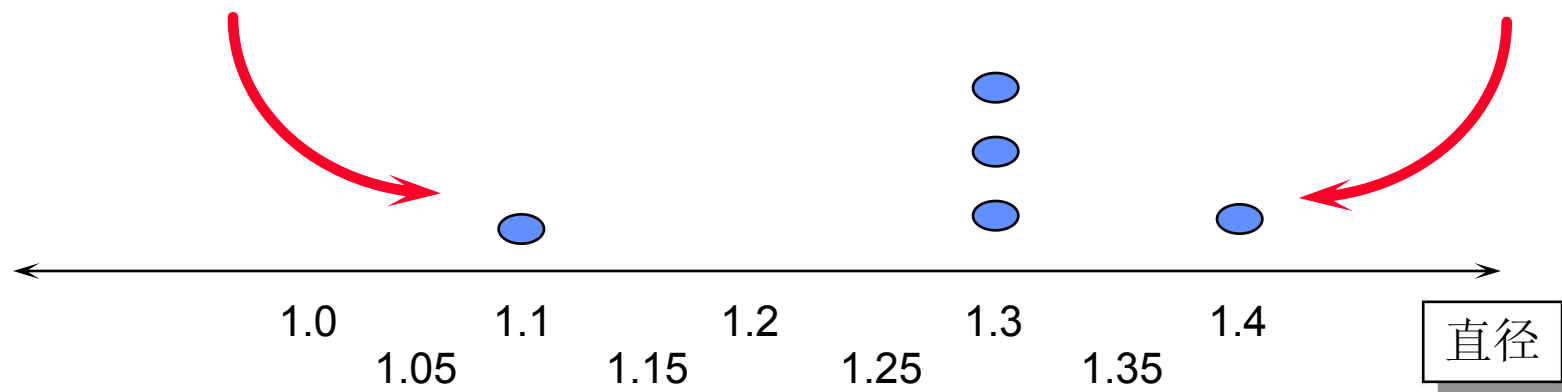
3	7	5
3	7	5
3	6	4
2	6	1
1	5	2
4	4	5
5	3	5
8	2	3
7	5	8
7	5	5

图形有助于清楚的看出研究对象的特性

点图

第二个观察点

第一个观察点



用无心磨床磨轴.

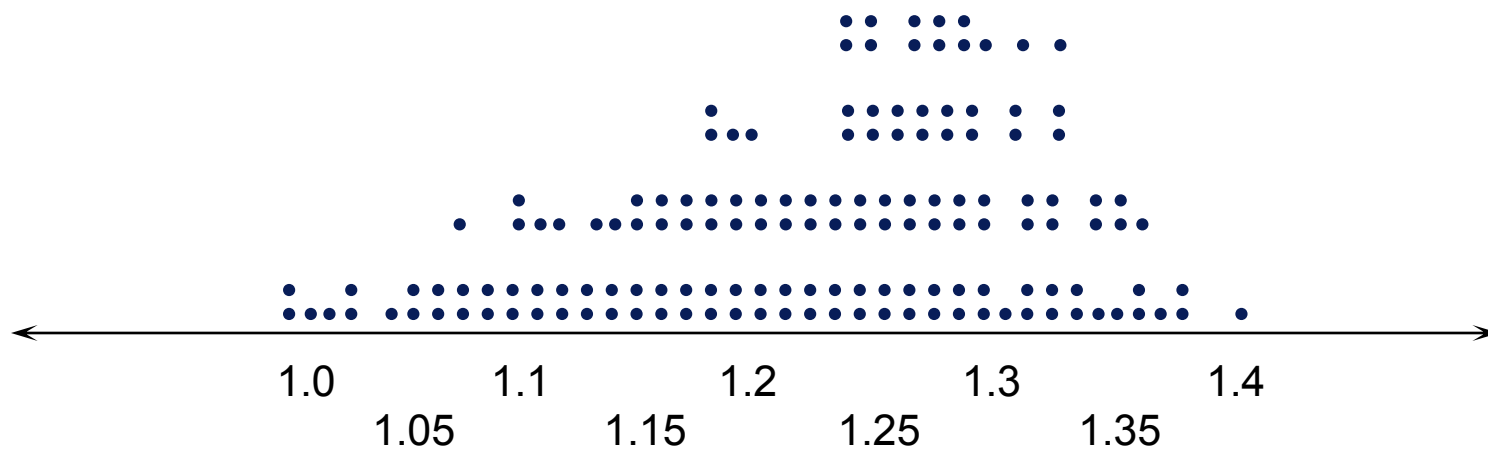
直径范围是 **1.0 to 1.4 inches**. 每测一支轴就按比率在图上描一个点

例如:

第一个观察点 = **1.4 inches**

第二个观察点 = **1.1 inches**

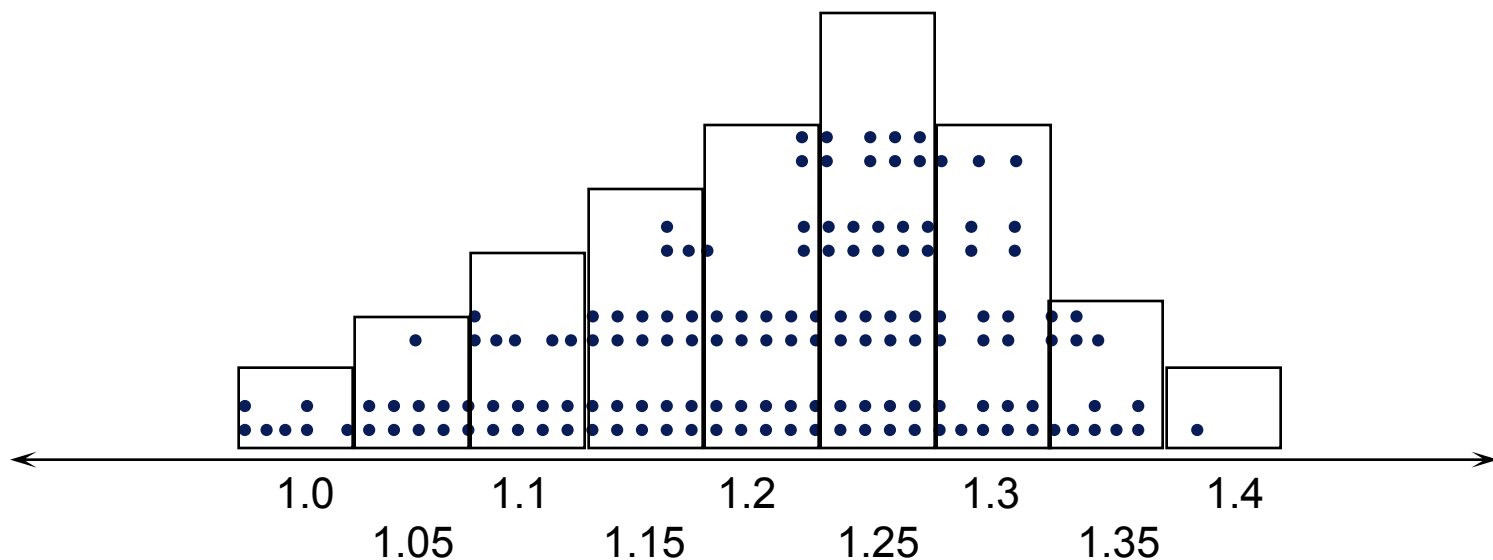
点图



假设我们测了**150**支轴，数据如上。

请你描述一下我们磨轴的工序？

点图 & 柱形图

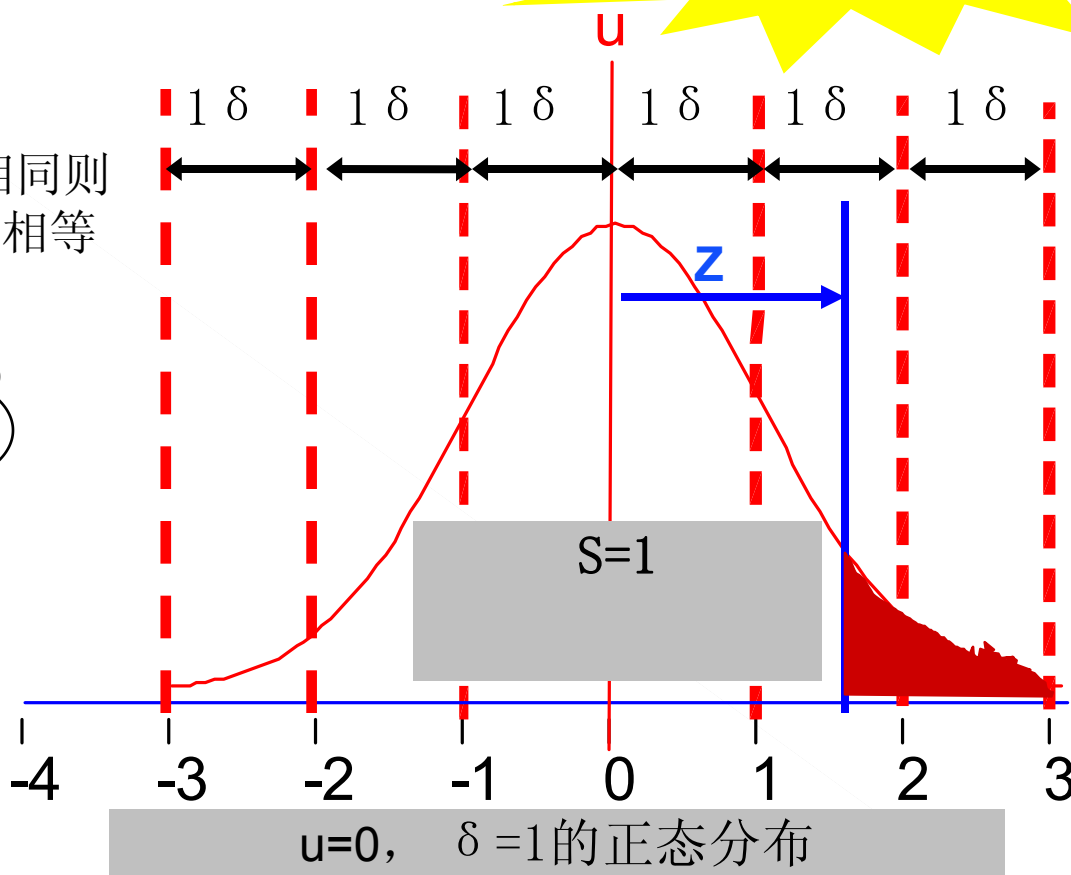


用等间距的柱形框将点包起来，柱形框的高度与其宽度范围内的点数成正比，就成了柱状图。

正态分布图

- 曲线底下的总面积等于1。
- 其形状和位置由 u 和 δ 决定。
- 对不同的正态曲线，如 Z 值相同则蓝线右边部分曲线底下的面积相等

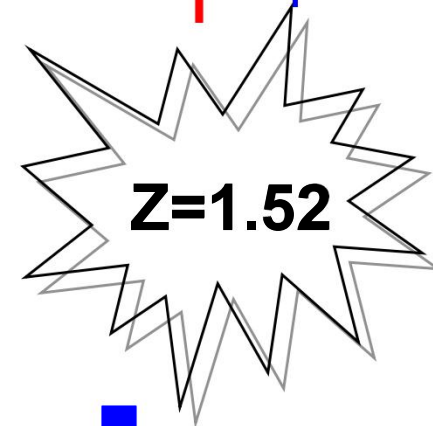
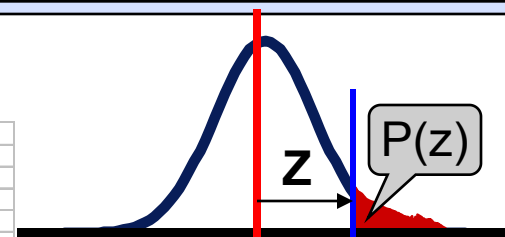
$$Z = (X - u) / \delta$$



正态分布曲线是一条反映概率大小的曲线

正态分布表

Z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	5.00E-01	4.96E-01	4.92E-01	4.88E-01	4.84E-01	4.80E-01	4.76E-01	4.72E-01	4.68E-01	4.64E-01
0.1	4.60E-01	4.56E-01	4.52E-01	4.48E-01	4.44E-01	4.40E-01	4.36E-01	4.33E-01	4.29E-01	4.25E-01
0.2	4.21E-01	4.17E-01	4.13E-01	4.09E-01	4.05E-01	4.01E-01	3.97E-01	3.94E-01	3.90E-01	3.86E-01
0.3	3.82E-01	3.78E-01	3.75E-01	3.71E-01	3.67E-01	3.63E-01	3.59E-01	3.56E-01	3.52E-01	3.48E-01
0.4	3.45E-01	3.41E-01	3.37E-01	3.34E-01	3.30E-01	3.26E-01	3.23E-01	3.19E-01	3.16E-01	3.12E-01
0.5	3.09E-01	3.05E-01	3.02E-01	2.98E-01	2.95E-01	2.91E-01	2.88E-01	2.84E-01	2.81E-01	2.78E-01
0.6	2.74E-01	2.71E-01	2.68E-01	2.64E-01	2.61E-01	2.58E-01	2.55E-01	2.51E-01	2.48E-01	2.45E-01
0.7	2.42E-01	2.39E-01	2.36E-01	2.33E-01	2.30E-01	2.27E-01	2.24E-01	2.21E-01	2.18E-01	2.15E-01
0.8	2.12E-01	2.09E-01	2.06E-01	2.03E-01	2.01E-01	1.98E-01	1.95E-01	1.92E-01	1.89E-01	1.87E-01
0.9	1.84E-01	1.81E-01	1.79E-01	1.76E-01	1.74E-01	1.71E-01	1.69E-01	1.66E-01	1.64E-01	1.61E-01
1	1.59E-01	1.56E-01	1.54E-01	1.52E-01	1.49E-01	1.47E-01	1.45E-01	1.42E-01	1.40E-01	1.38E-01
1.1	1.36E-01	1.34E-01	1.31E-01	1.29E-01	1.27E-01	1.25E-01	1.23E-01	1.21E-01	1.19E-01	1.17E-01
1.2	1.15E-01	1.13E-01	1.11E-01	1.09E-01	1.08E-01	1.06E-01	1.04E-01	1.02E-01	1.00E-01	9.85E-02
1.3	9.68E-02	9.51E-02	9.34E-02	9.18E-02	9.01E-02	8.85E-02	8.69E-02	8.53E-02	8.38E-02	8.23E-02
1.4	8.08E-02	7.93E-02	7.78E-02	7.64E-02	7.49E-02	7.35E-02	7.21E-02	7.08E-02	6.94E-02	6.81E-02
1.5	6.68E-02	6.55E-02	6.43E-02	6.30E-02	6.18E-02	6.06E-02	5.94E-02	5.82E-02	5.71E-02	5.6E-02
1.6	5.48E-02	5.37E-02	5.26E-02	5.16E-02	5.05E-02	4.95E-02	4.85E-02	4.75E-02	4.65E-02	4.55E-02
1.7	4.46E-02	4.36E-02	4.27E-02	4.18E-02	4.09E-02	4.01E-02	3.82E-02	3.84E-02	3.75E-02	3.67E-02
1.8	3.59E-02	3.52E-02	3.44E-02	3.36E-02	3.29E-02	3.22E-02	3.14E-02	3.07E-02	3.01E-02	2.94E-02
1.9	2.87E-02	2.81E-02	2.74E-02	2.68E-02	2.62E-02	2.56E-02	2.50E-02	2.44E-02	2.39E-02	2.33E-02
2	2.28E-02	2.22E-02	2.17E-02	2.12E-02	2.07E-02	2.02E-02	1.97E-02	1.92E-02	1.88E-02	1.83E-02
2.1	1.79E-02	1.74E-02	1.70E-02	1.66E-02	1.62E-02	1.58E-02	1.54E-02	1.50E-02	1.46E-02	1.43E-02
2.2	1.39E-02	1.36E-02	1.32E-02	1.29E-02	1.26E-02	1.22E-02	1.19E-02	1.16E-02	1.13E-02	1.10E-02
2.3	1.07E-02	1.04E-02	1.02E-02	9.90E-03	9.64E-03	9.39E-03	9.14E-03	8.89E-03	8.66E-03	8.42E-03
2.4	8.20E-03	7.98E-03	7.76E-03	7.55E-03	7.34E-03	7.14E-03	6.95E-03	6.76E-03	6.57E-03	6.39E-03
2.5	6.21E-03	6.04E-03	5.87E-03	5.70E-03	5.54E-03	5.39E-03	5.23E-03	5.09E-03	4.94E-03	4.80E-03
2.6	4.66E-03	4.53E-03	4.40E-03	4.27E-03	4.15E-03	4.02E-03	3.90E-03	3.79E-03	3.68E-03	3.57E-03
2.7	3.47E-03	3.36E-03	3.26E-03	3.17E-03	3.07E-03	2.98E-03	2.89E-03	2.80E-03	2.72E-03	2.64E-03
2.8	2.56E-03	2.48E-03	2.40E-03	2.33E-03	2.26E-03	2.19E-03	2.12E-03	2.05E-03	1.99E-03	1.93E-03
2.9	1.87E-03	1.81E-03	1.75E-03	1.70E-03	1.64E-03	1.59E-03	1.54E-03	1.49E-03	1.44E-03	1.40E-03
3	1.35E-03	1.31E-03	1.26E-03	1.22E-03	1.18E-03	1.14E-03	1.11E-03	1.07E-03	1.04E-03	1.00E-03
3.1	9.68E-04	9.35E-04	9.04E-04	8.74E-04	8.45E-04	8.16E-04	7.89E-04	7.62E-04	7.36E-04	7.11E-04
3.2	6.87E-04	6.64E-04	6.41E-04	6.19E-04	5.98E-04	5.77E-04	5.57E-04	5.38E-04	5.19E-04	5.01E-04
3.3	4.84E-04	4.67E-04	4.50E-04	4.34E-04	4.19E-04	4.04E-04	3.90E-04	3.76E-04	3.63E-04	3.50E-04
3.4	3.37E-04	3.25E-04	3.13E-04	3.02E-04	2.91E-04	2.80E-04	2.70E-04	2.60E-04	2.51E-04	2.42E-04
3.5	2.33E-04	2.24E-04	2.16E-04	2.08E-04	2.00E-04	1.93E-04	1.86E-04	1.79E-04	1.72E-04	1.66E-04
3.6	1.59E-04	1.53E-04	1.47E-04	1.42E-04	1.36E-04	1.31E-04	1.26E-04	1.21E-04	1.17E-04	1.12E-04
3.7	1.08E-04	1.04E-04	9.97E-05	9.59E-05	9.21E-05	8.86E-05	8.51E-05	8.18E-05	7.85E-05	7.55E-05
3.8	7.25E-05	6.96E-05	6.69E-05	6.42E-05	6.17E-05	5.92E-05	5.68E-05	5.46E-05	5.24E-05	5.03E-05
3.9	4.82E-05	4.63E-05	4.44E-05	4.26E-05	4.09E-05	3.92E-05	3.76E-05	3.61E-05	3.46E-05	3.32E-05
4	3.18E-05	3.05E-05	2.92E-05	2.80E-05	2.68E-05	2.57E-05	2.47E-05	2.36E-05	2.26E-05	2.17E-05
4.1	2.08E-05	1.99E-05	1.91E-05	1.82E-05	1.75E-05	1.67E-05	1.60E-05	1.53E-05	1.47E-05	1.40E-05
4.2	1.34E-05	1.29E-05	1.23E-05	1.18E-05	1.13E-05	1.08E-05	1.03E-05	9.86E-06	9.43E-06	9.01E-06
4.3	8.62E-06	8.24E-06	7.88E-06	7.53E-06	7.20E-06	6.88E-06	6.57E-06	6.28E-06	6.00E-06	5.73E-06
4.4	5.48E-06	5.23E-06	5.00E-06	4.77E-06	4.56E-06	4.35E-06	4.16E-06	3.97E-06	3.79E-06	3.62E-06
4.5	3.45E-06	3.29E-06	3.14E-06	3.00E-06	2.86E-06	2.73E-06	2.60E-06	2.48E-06	2.37E-06	2.26E-06
4.6	2.15E-06	2.05E-06	1.96E-06	1.87E-06	1.78E-06	1.70E-06	1.62E-06	1.54E-06	1.47E-06	1.40E-06
4.7	1.33E-06	1.27E-06	1.21E-06	1.15E-06	1.10E-06	1.05E-06	9.96E-07	9.48E-07	9.03E-07	8.59E-07
4.8	8.18E-07	7.79E-07	7.41E-07	7.05E-07	6.71E-07	6.39E-07	6.08E-07	5.78E-07	5.50E-07	5.23E-07
4.9	4.98E-07	4.73E-07	4.50E-07	4.28E-07	4.07E-07	3.87E-07	3.68E-07	3.50E-07	3.32E-07	3.16E-07

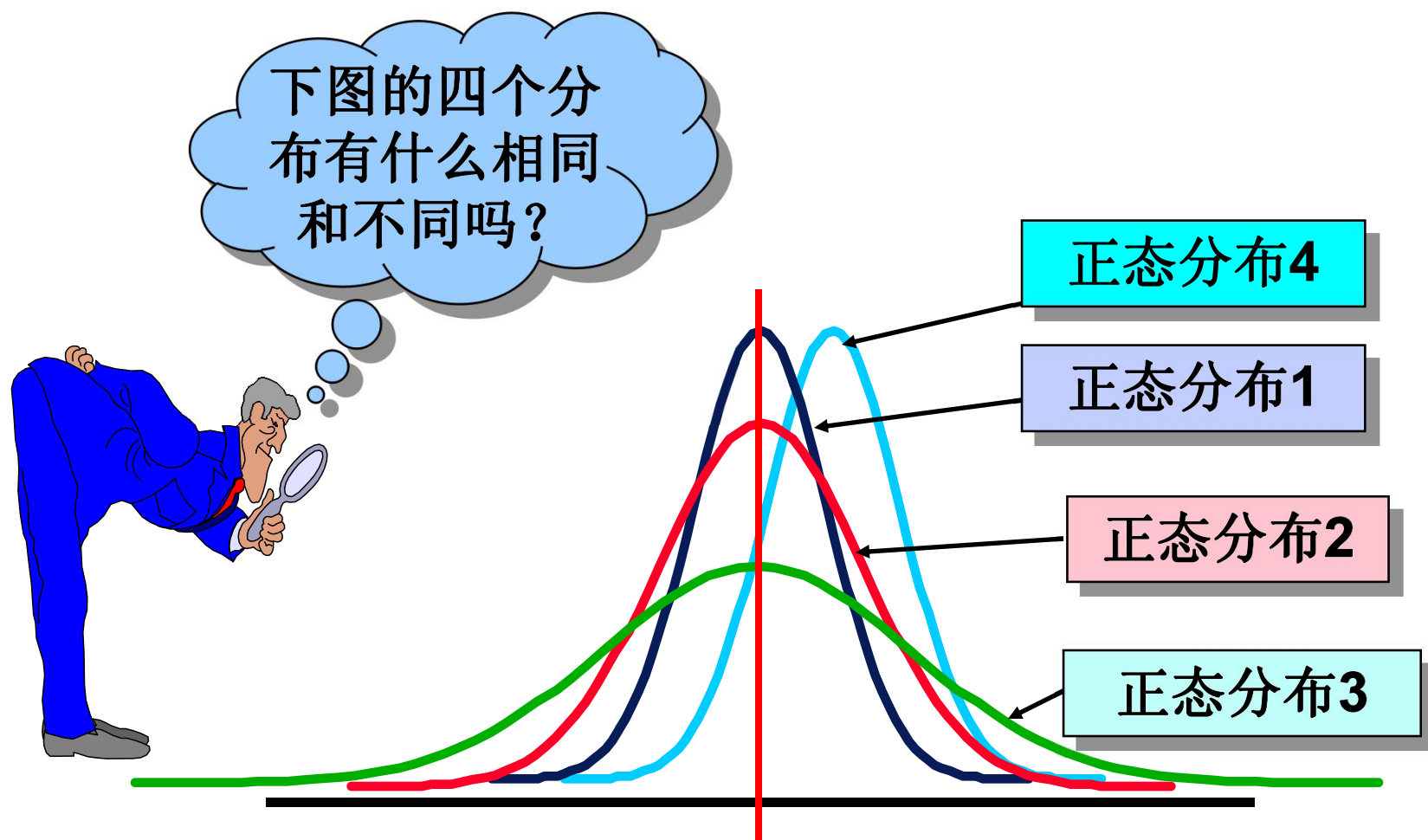


查表可知

 $P(z)=0.0643$
Microsoft Excel
工作表

P(z)由Z值唯一决定.

正态分布图



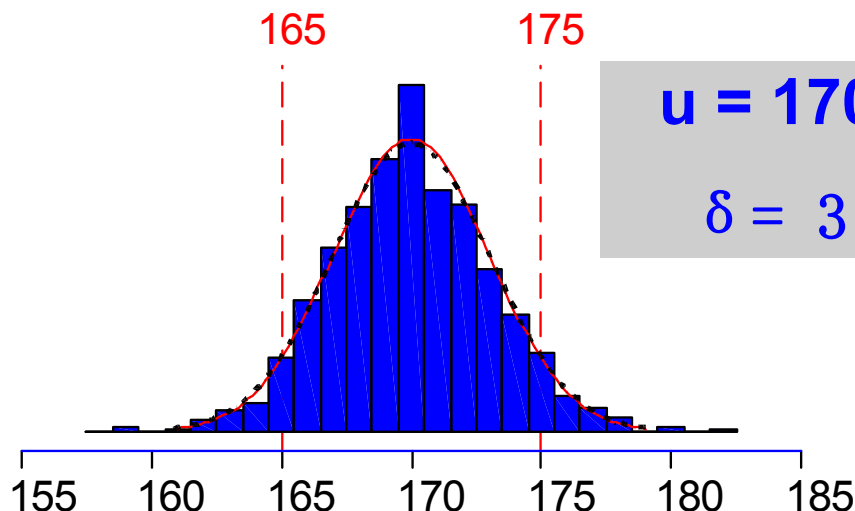
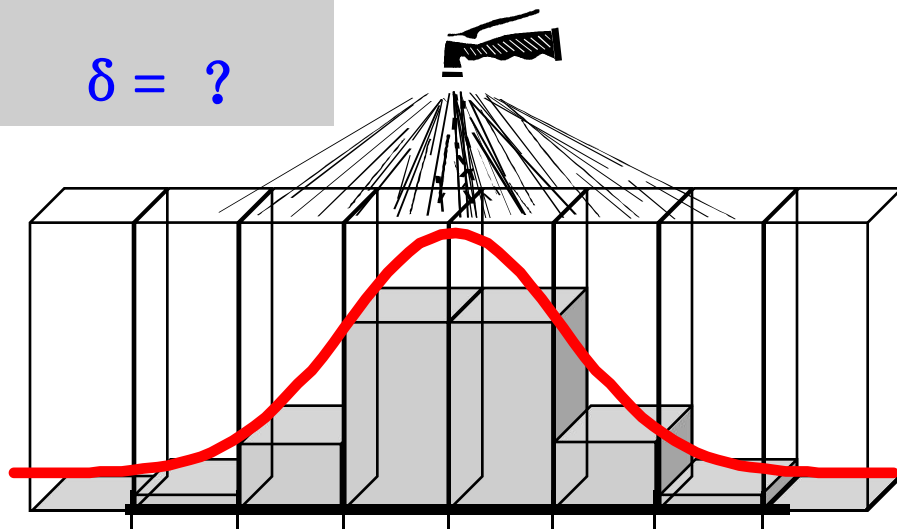
其形状和位置由 μ 和 σ 决定。

正态分布的例子

国人平均身高

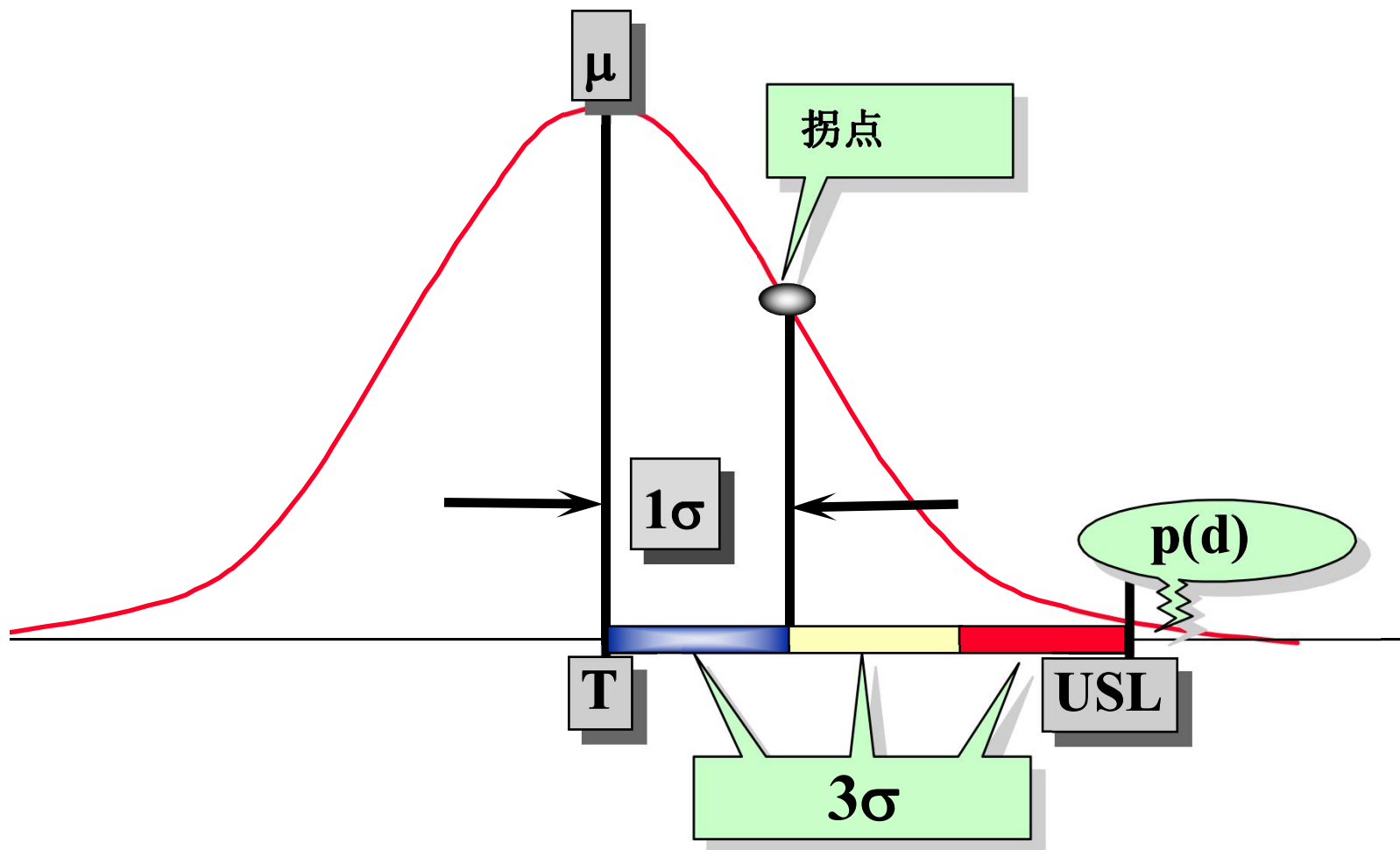
$$\mu = ?$$

$$\sigma = ?$$

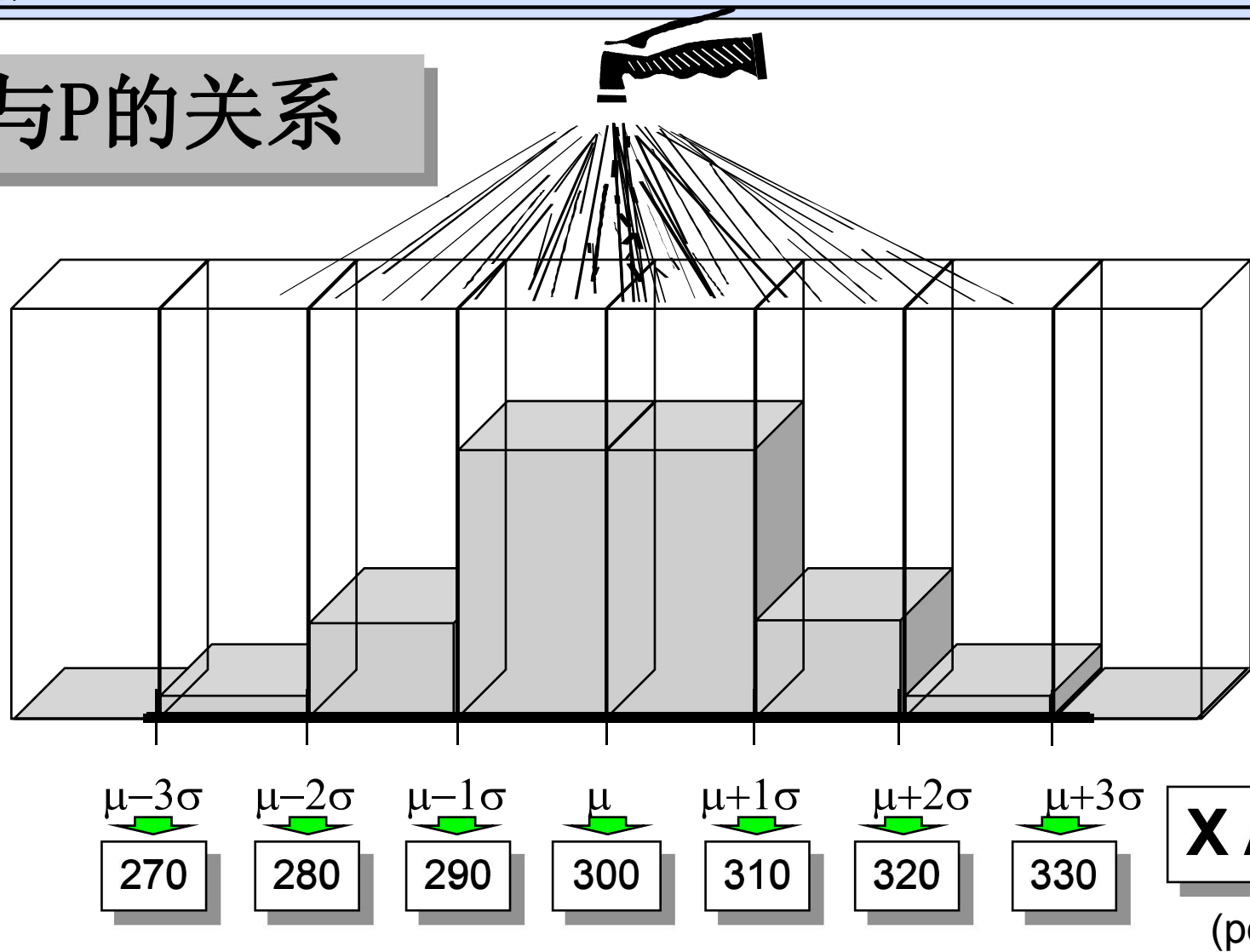
Microsoft Excel
工作表身高在**175**以上的占多大比率:?身高在**165**以下的占多大比率:?身高在**165~175**之间的占多大比率:?

正态分布只是属于众多分布中的一种，但由经验得知，正常情况下，大部分的制程都满足正态分布。

正态分布图



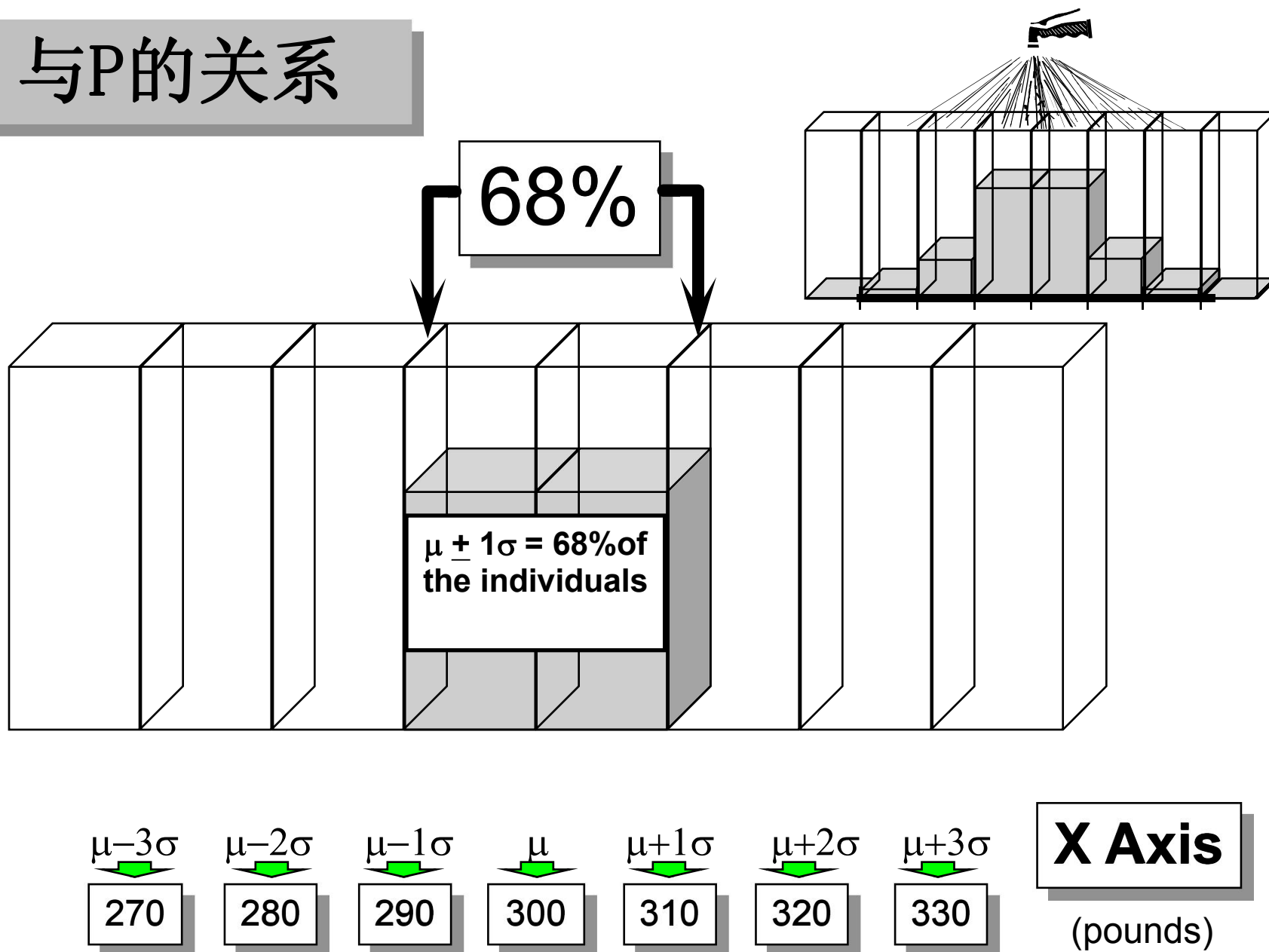
δ 与P的关系



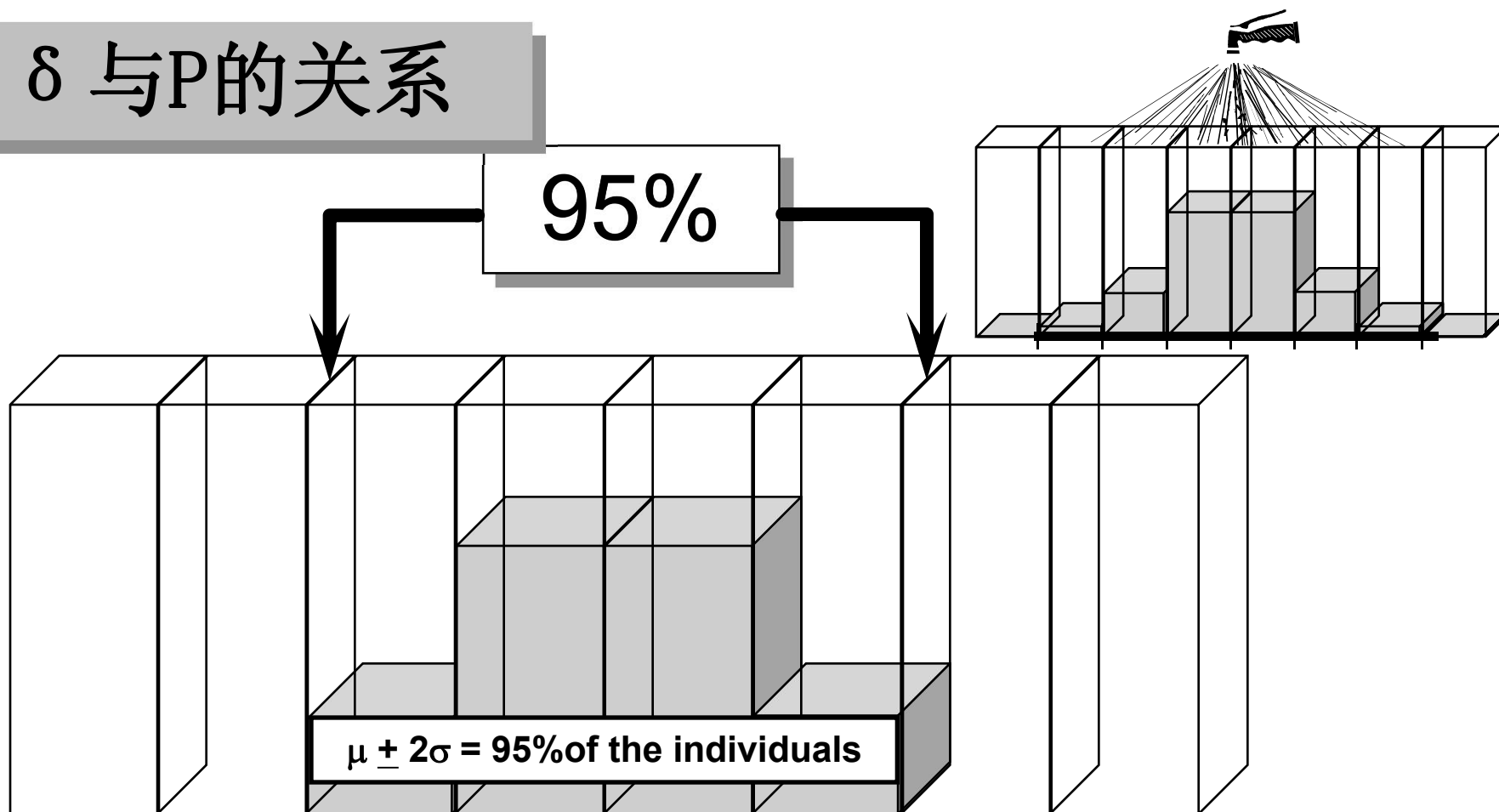
练习

假设 $\mu=300$, $\sigma=10$

σ 与P的关系



δ 与P的关系

 $\mu - 3\sigma$

270

 $\mu - 2\sigma$

280

 $\mu - 1\sigma$

290

 μ

300

 $\mu + 1\sigma$

310

 $\mu + 2\sigma$

320

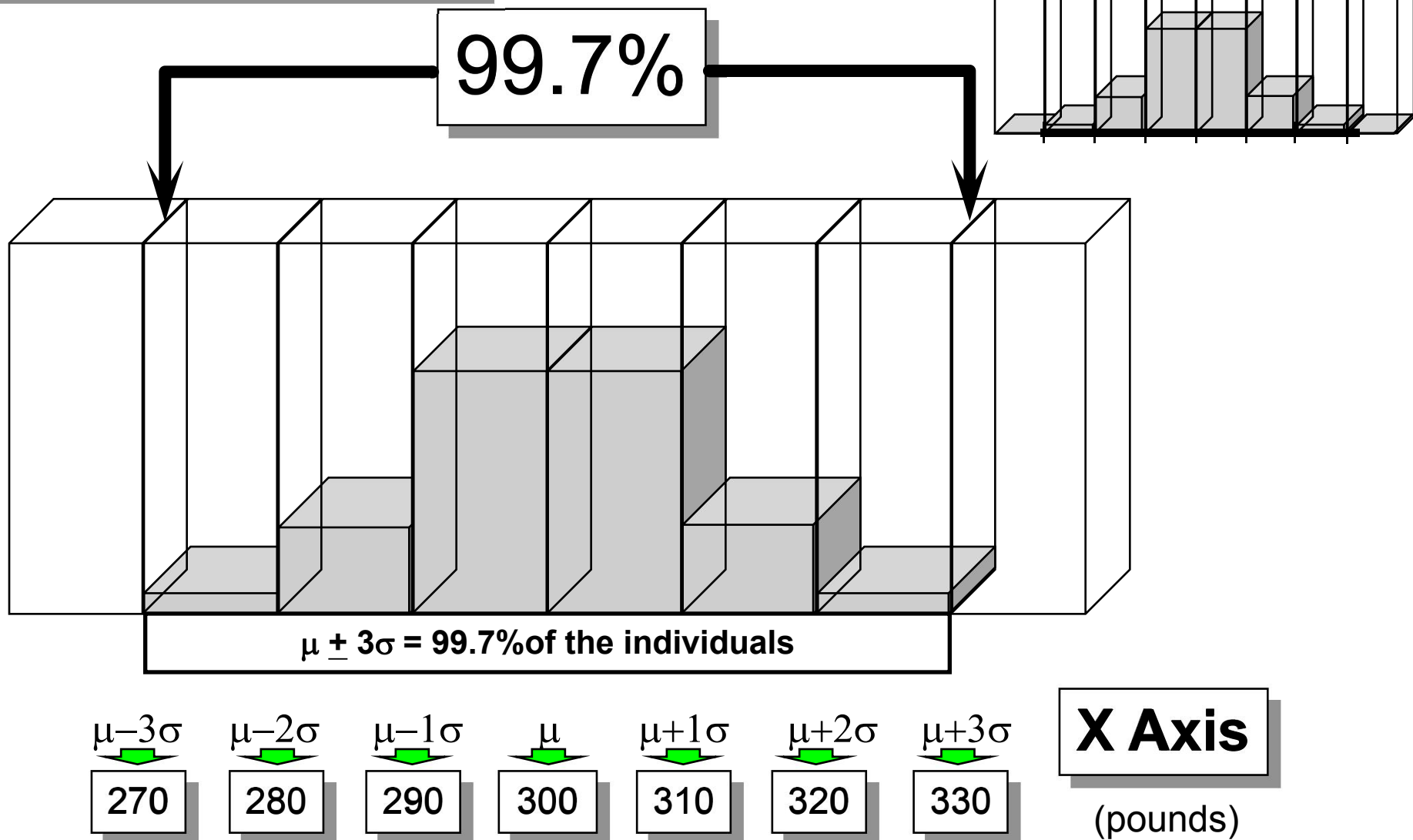
 $\mu + 3\sigma$

330

X Axis

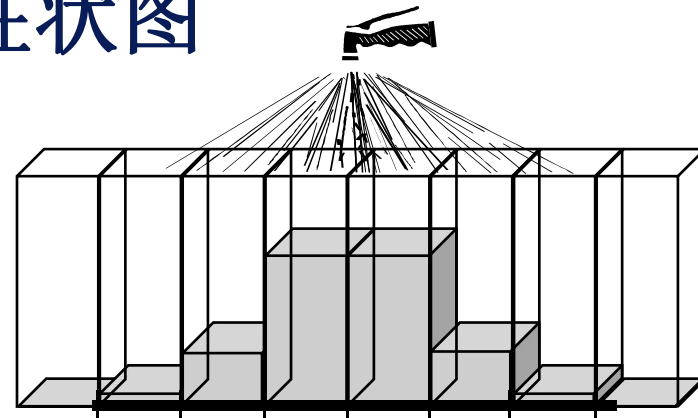
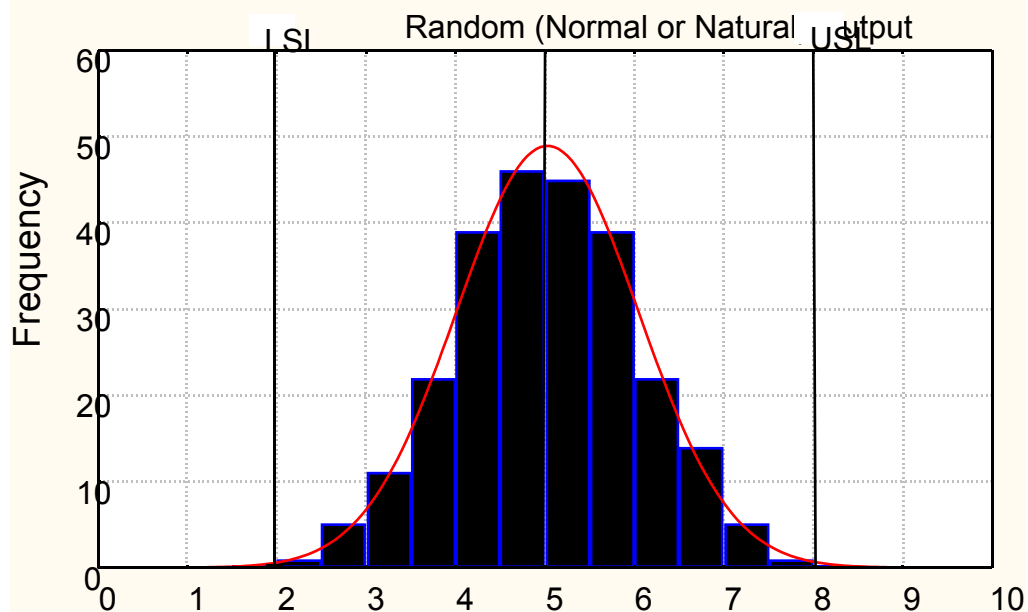
(pounds)

δ 与P的关系



几种典型的柱状图

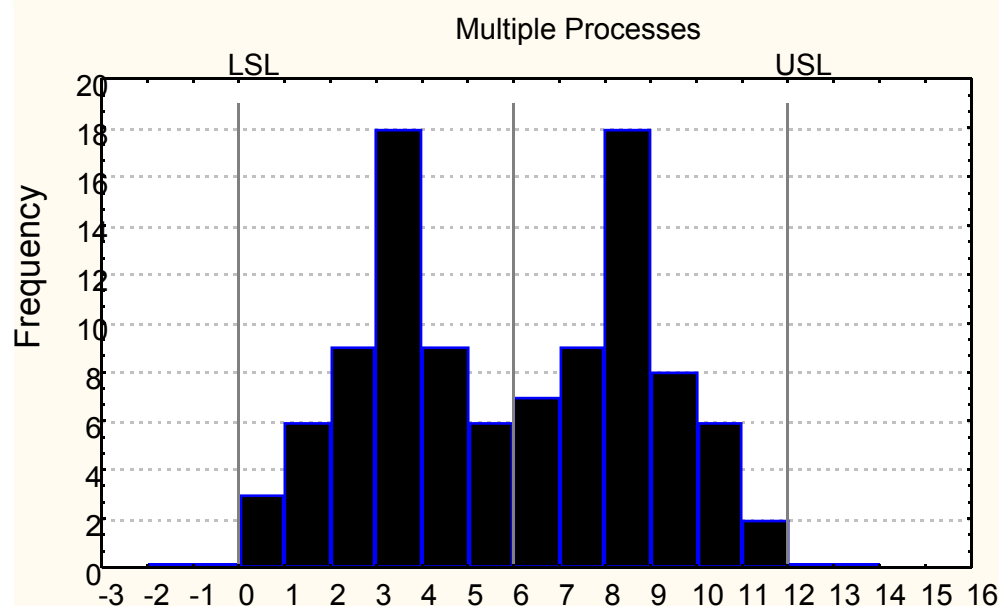
正常型



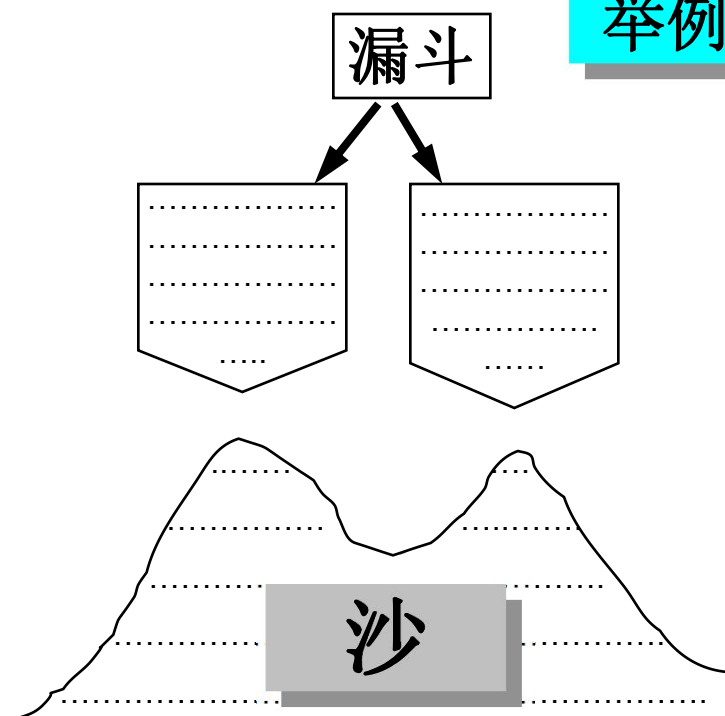
工序无异常

几种典型的柱状图

双峰型



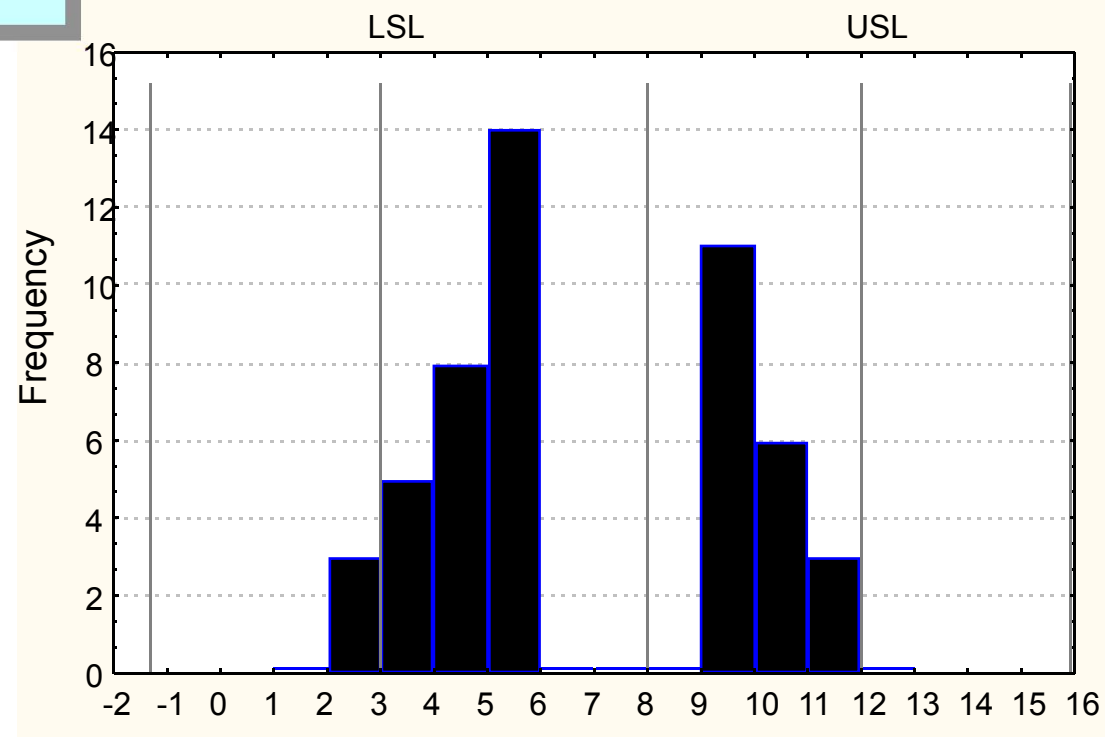
举例



通常是由样本来自不同的班次、机台、组别等引起

几种点典型的柱状图

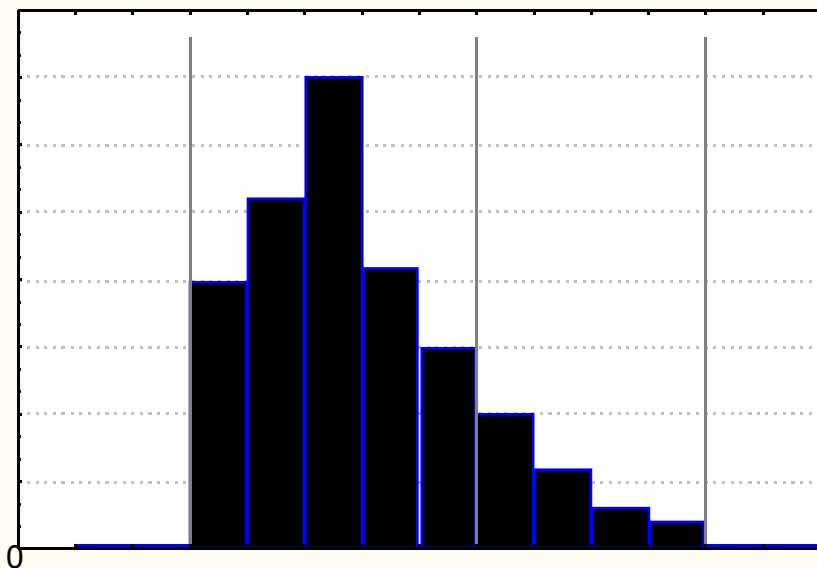
离岛型



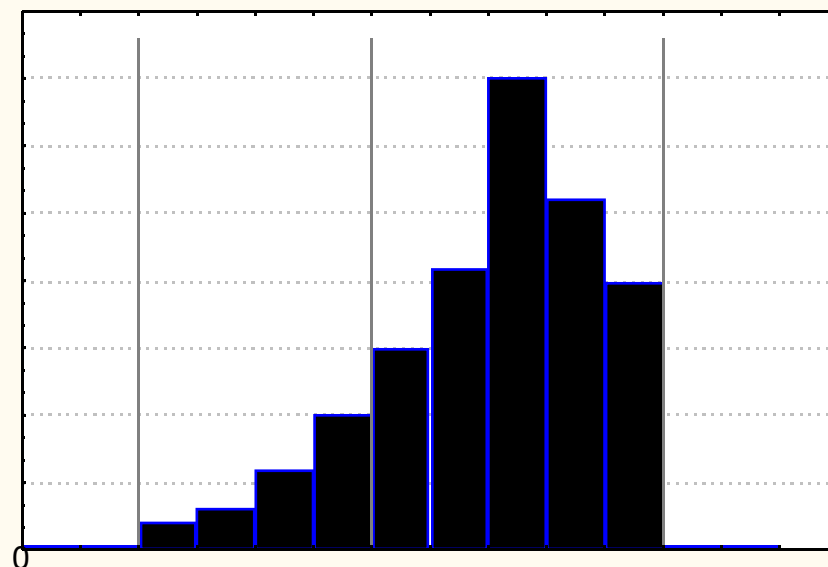
类似于双峰型

几种典型的柱状图

左倾型

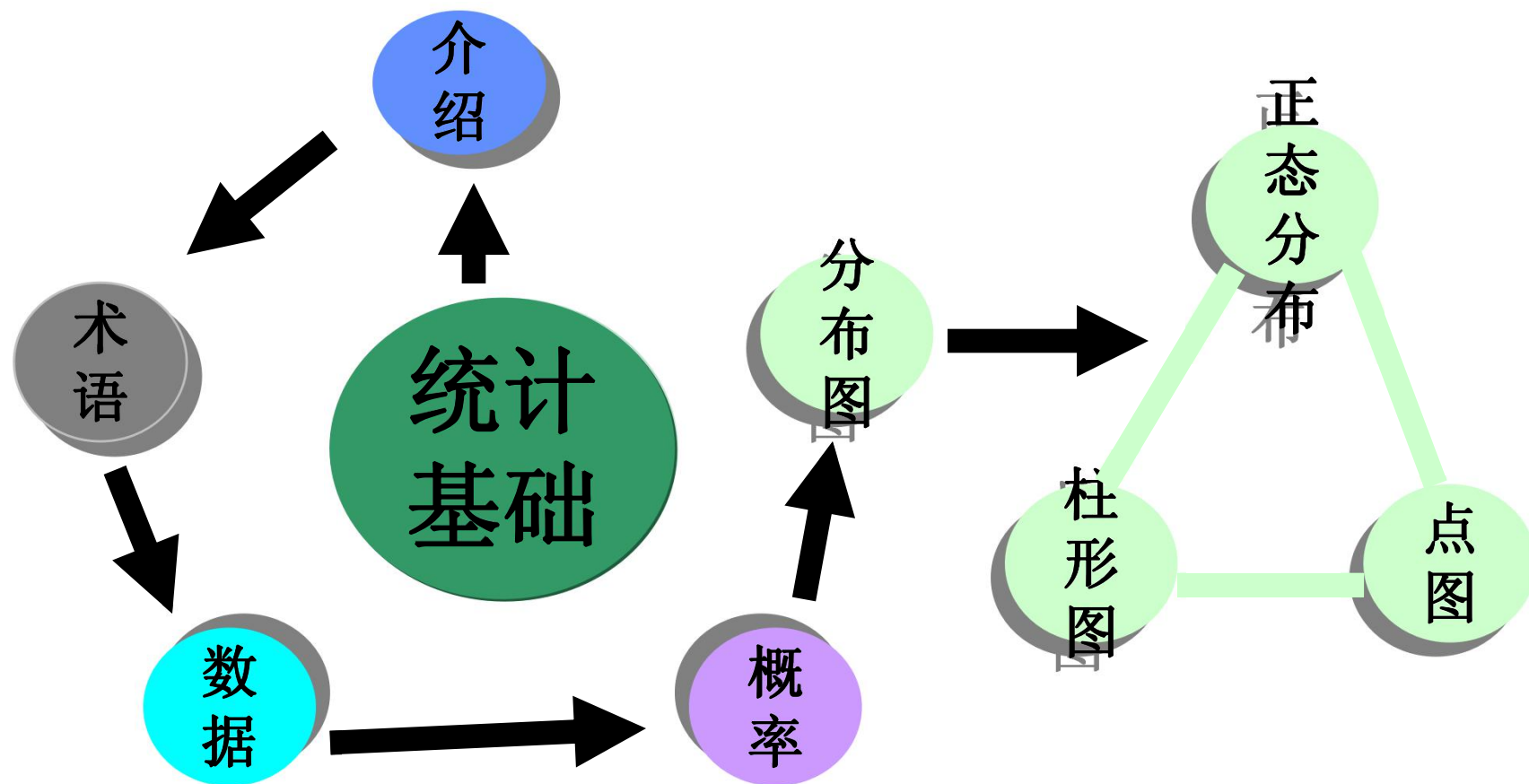


右倾型



工序有异常

课程主要内容总结



课程完