

# 《传播统计学》

## 正态分布

教师：林志良

邮箱：[linzhl@nfu.edu.cn](mailto:linzhl@nfu.edu.cn)

个人网站：[www.zhilianglin.com](http://www.zhilianglin.com)

# 目录

- 正态分布
- 标准正态分布及Z分数

# 正态分布

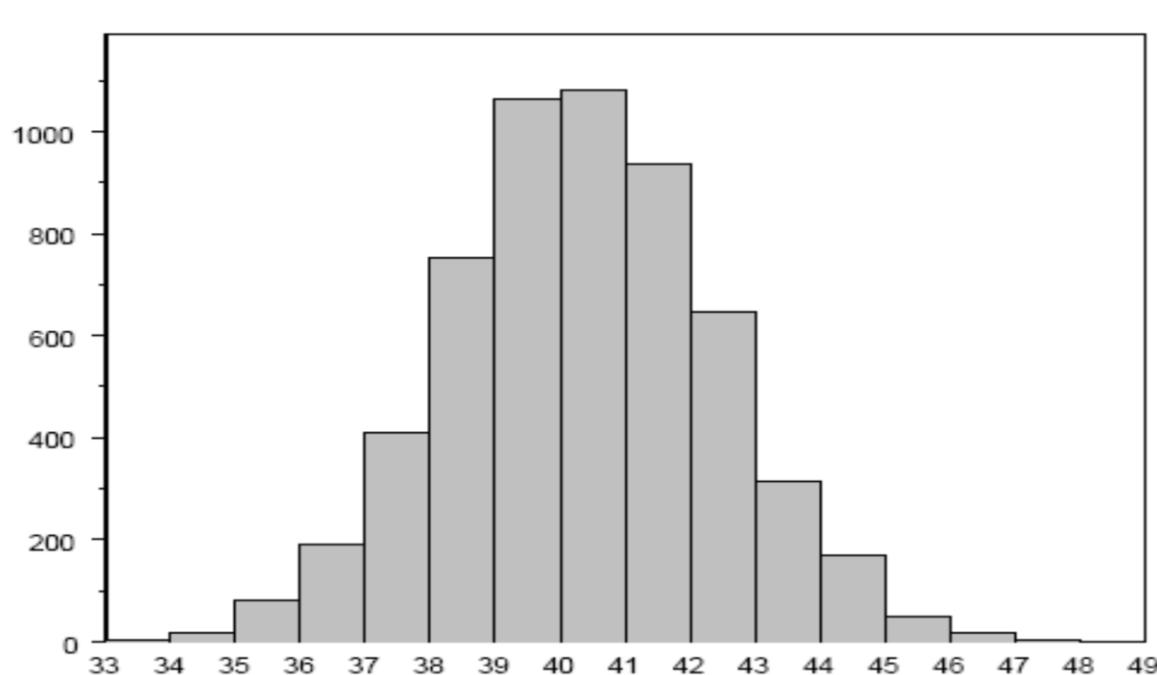
## 5732位苏格兰军人胸围数据

胸围(英寸)	频数	胸围(英寸)	频数
33	3	41	935
34	19	42	646
35	81	43	313
36	189	44	168
37	409	45	50
38	753	46	18
39	1062	47	3
40	1082	48	1

# 正态分布

胸围数据的直方图，平均值及标准差

分别为39.85及2.07

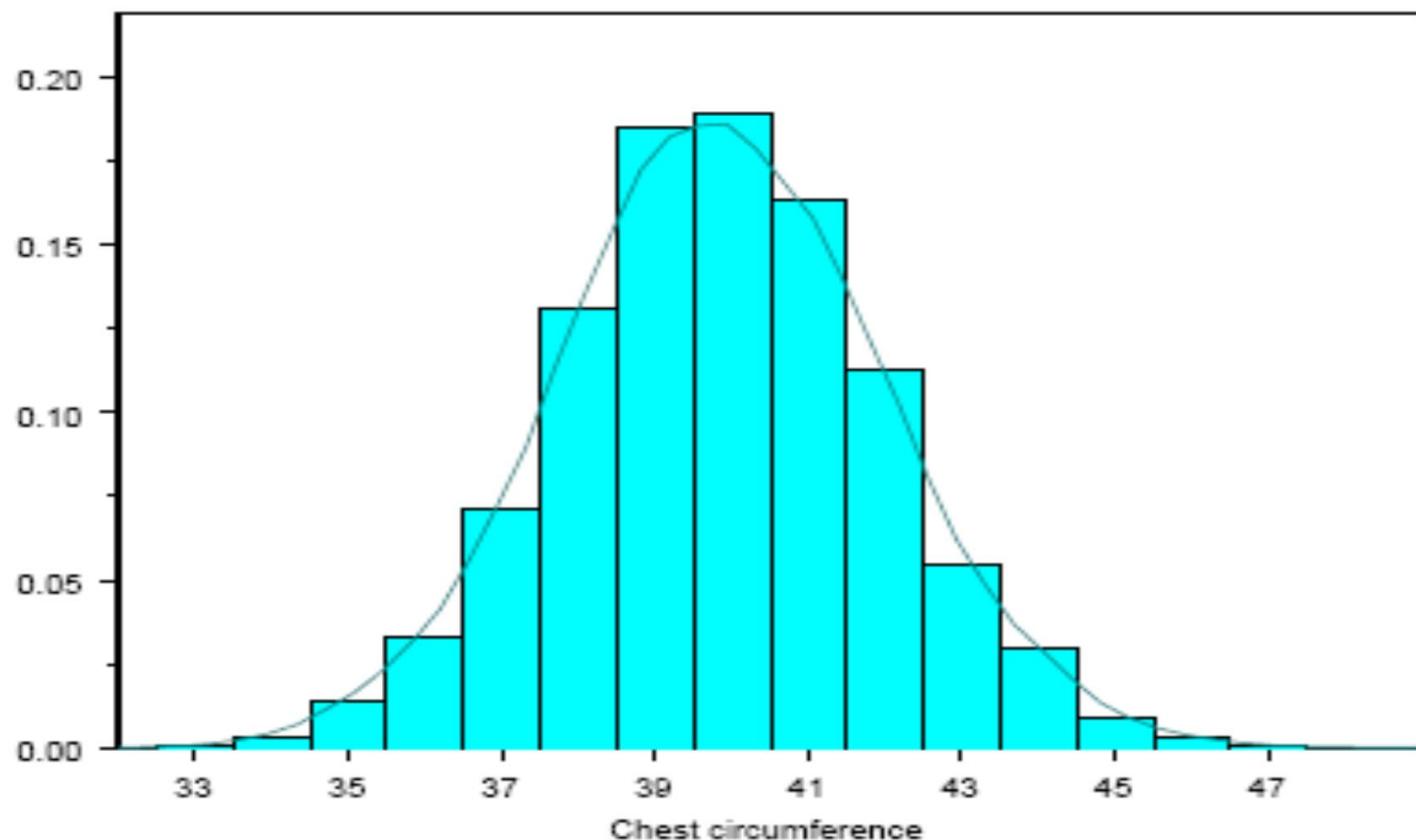


- 总体而言，该直方图呈现“钟型”，对称于其平均值。

# 正态分布

胸围数据的直方图，平均值及标准差

分别为39.85及2.07

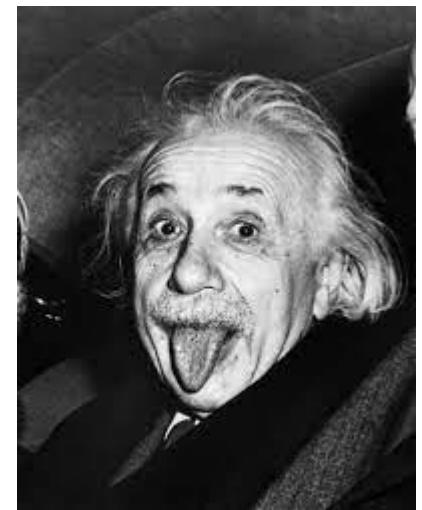
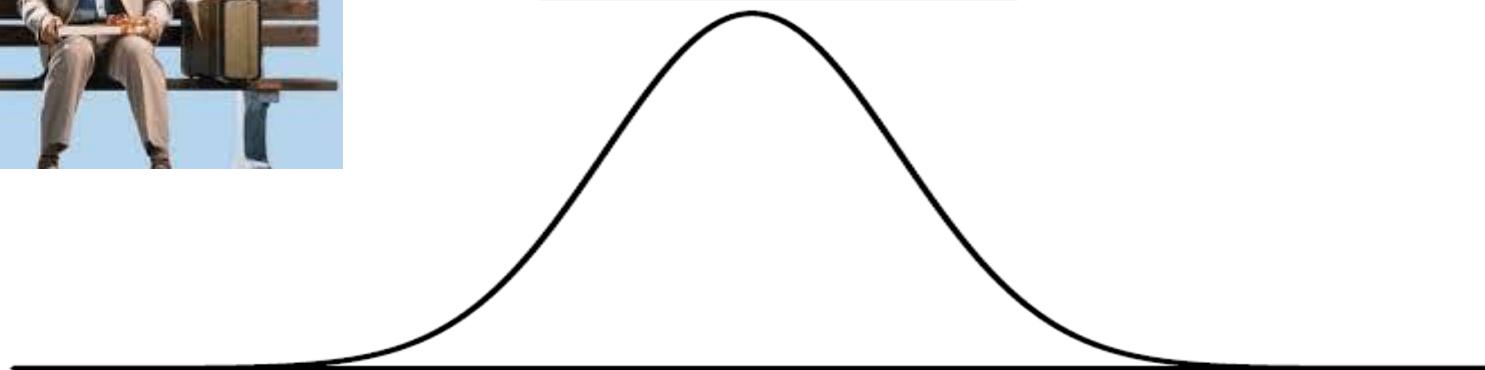




# 正态分布

## 正态分布 ( Normal Distribution )

- 大部份的值都极集中在平均值两侧，极端值不多，图形以平均值为中心左右对称且成钟形分布。



# 正态分布

## 正态分布的函数表达式

$$f(X) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{(X-\mu)}{\sigma}\right)^2}$$

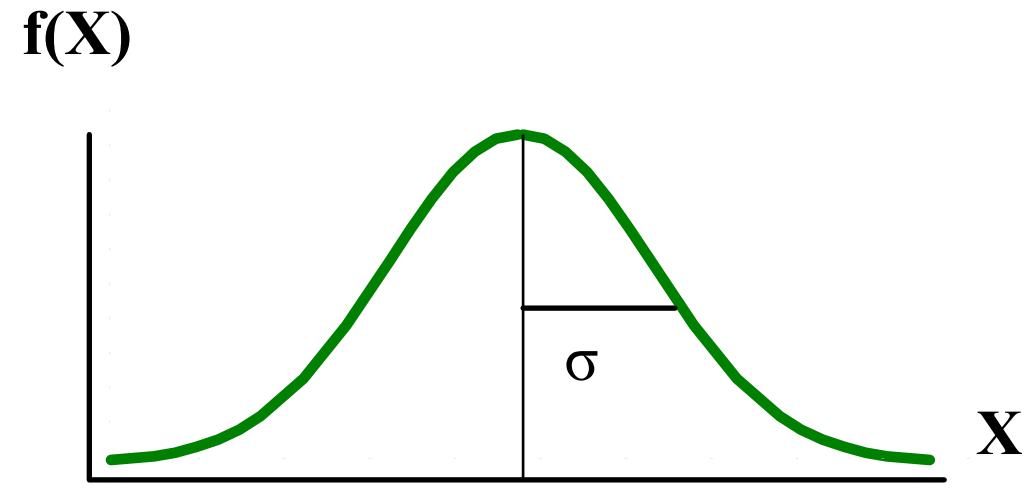
- 也可记作  $X \sim N(\mu, \sigma^2)$



# 正态分布

## 正态分布的特点

- 钟形分布
- 对称
- 均值、中位数、众数为同一个数



正态分布的**位置**由**平均值** $\mu$ 决定

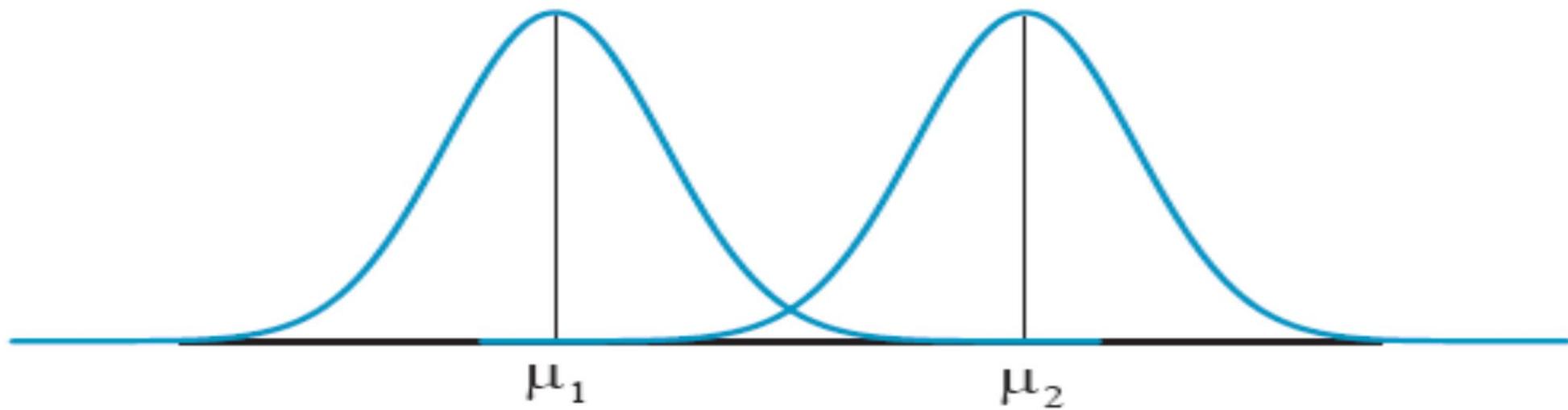
正态分布的**形状**由**标准差** $\sigma$ 决定

X的取值范围为负无穷到正无穷 ( $-\infty \sim +\infty$ )

# 正态分布

正态分布的位置因平均数而异

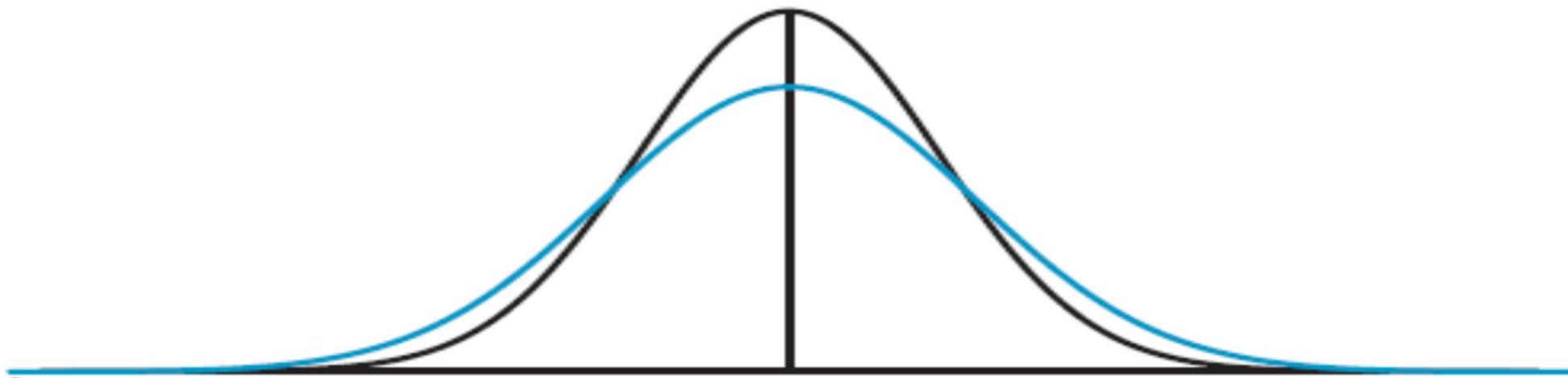
- 标准差相同而平均数不同的正态分布，形状相同，但位置不一样。



# 正态分布

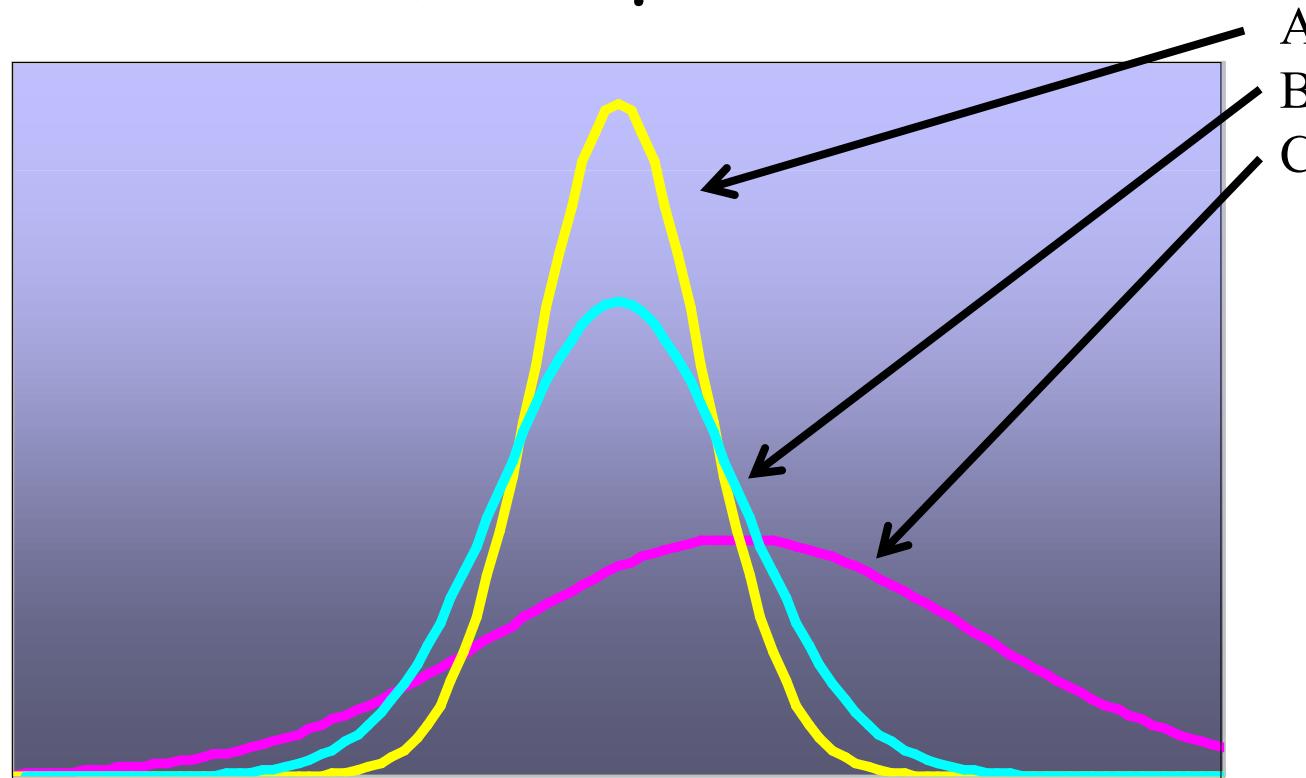
**正态分布的形状由标准差决定**

- 平均数相同标准差不同的正态分布，中心位置相同，但高度形状不一样



# 正态分布

正态分布因 $\mu$ 和 $\sigma$ 不同而不同



- A和B具有相同的平均值，不同的标准差
- B和C平均值和标准差都不一样



# 标准正态分布

## 标准正态分布

(Standard Normal Distribution)

- 任何正态分布（无论其 $\mu$ 和 $\sigma$ 的取值为多少）都能转化成标准正态分布（Z分布）
- 为了得到Z分数，我们需要将原数据X的单位进行转化
- 均值为0，标准差为1的正态分布曲线被称为标准正态分布。



# 标准正态分布

## 转化成标准正态分布

- 将原始数据 $X$ 转化为标准正态分布（Z分布）的公式为原始数据 $X$ 减去均值 $\mu$ 再除以标准差 $\sigma$ 。

$$Z = \frac{X - \mu}{\sigma}$$

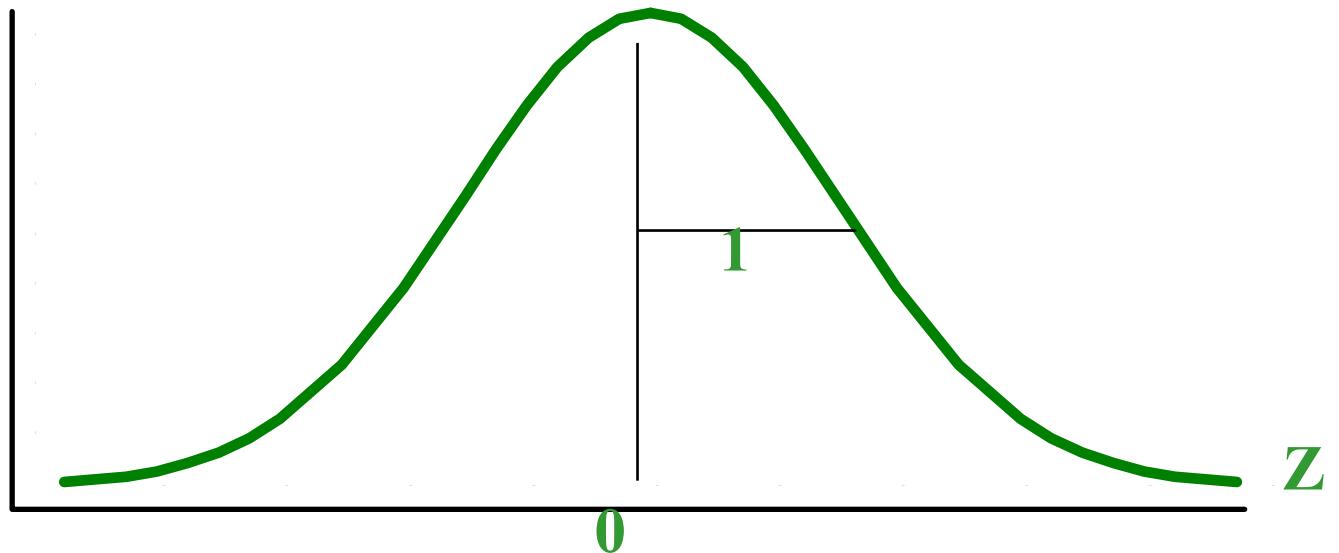
Z分布的均值为0，标准差为1

# 标准正态分布

- 也称为Z分布
- 均值为0
- 标准差为1

## 标准正态分布

$f(Z)$



# 标准正态分布

## 例子

- 如果 $X$ 是服从均值为100，标准差为50的正态分布。

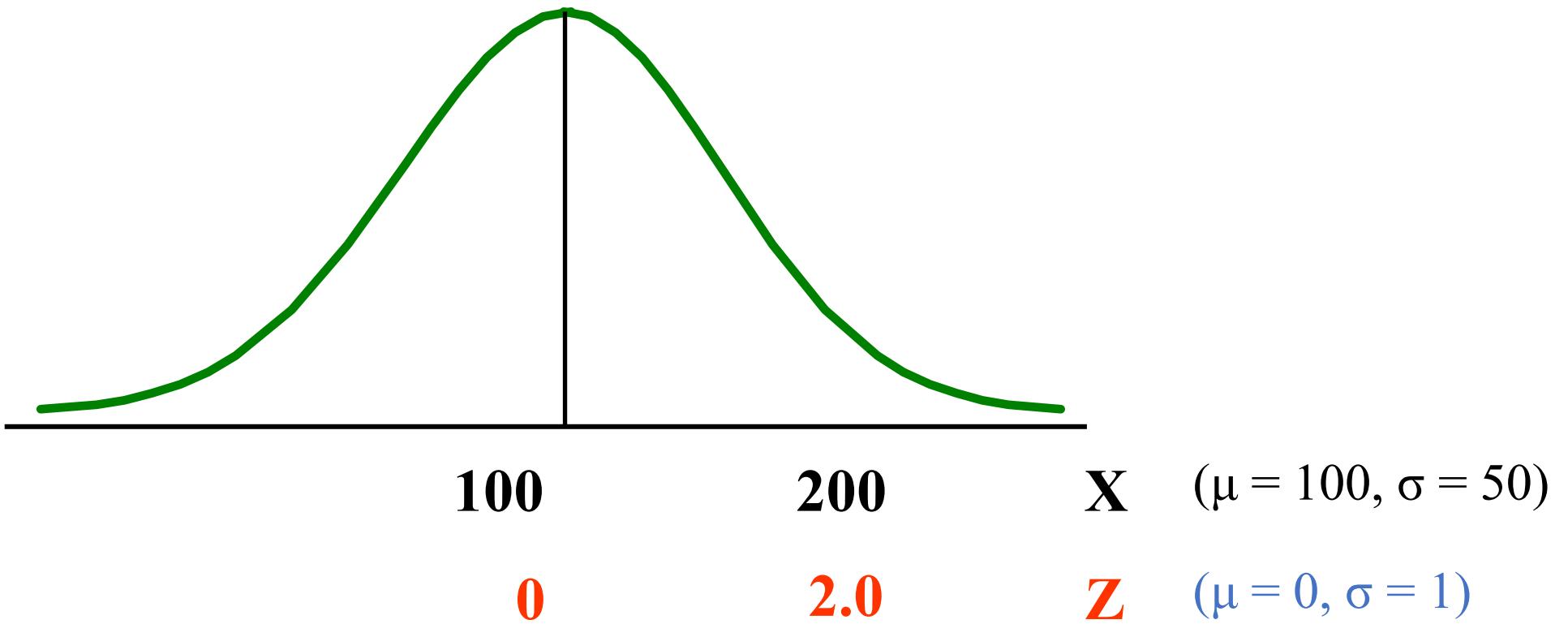
给定 $X=200$ ，计算其Z分数 (Z-score) :

$$Z = \frac{X - \mu}{\sigma} = \frac{200 - 100}{50} = 2.0$$

- 这意味着 $X=200$ 离均值的距离为两个标准差

# 标准正态分布

比较X和Z的单位



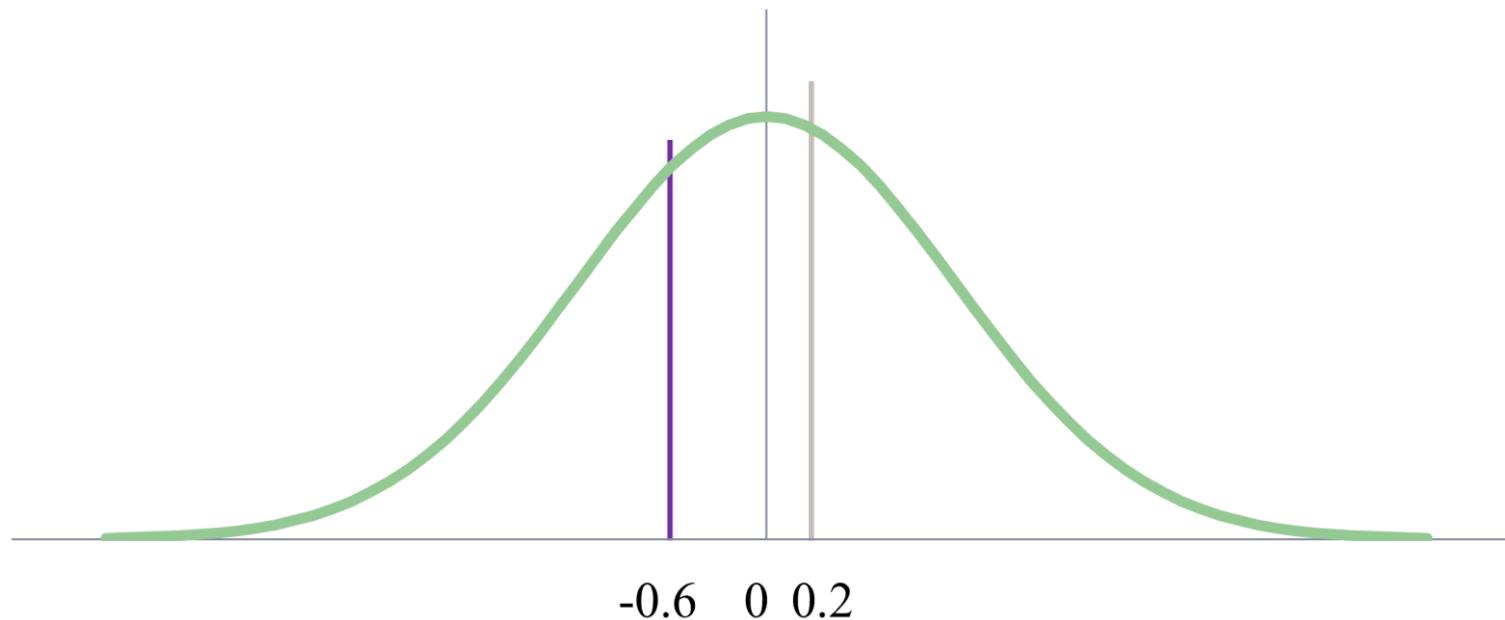
# 标准正态分布

**例子：统计学 VS 经济学 哪一门课的班级排名比较好**

- 小明的统计学67分，平均为65分，标准差10分  
经济学72分，平均为75分，标准差5分
- 则统计学的标准分为  $(67-65) / 10 = 0.2$
- 经济学的标准分为  $(72-75) / 5 = -0.6$
- 所以小明的统计学在班上的排名比较好

# 标准正态分布

例子：统计学 VS 经济学

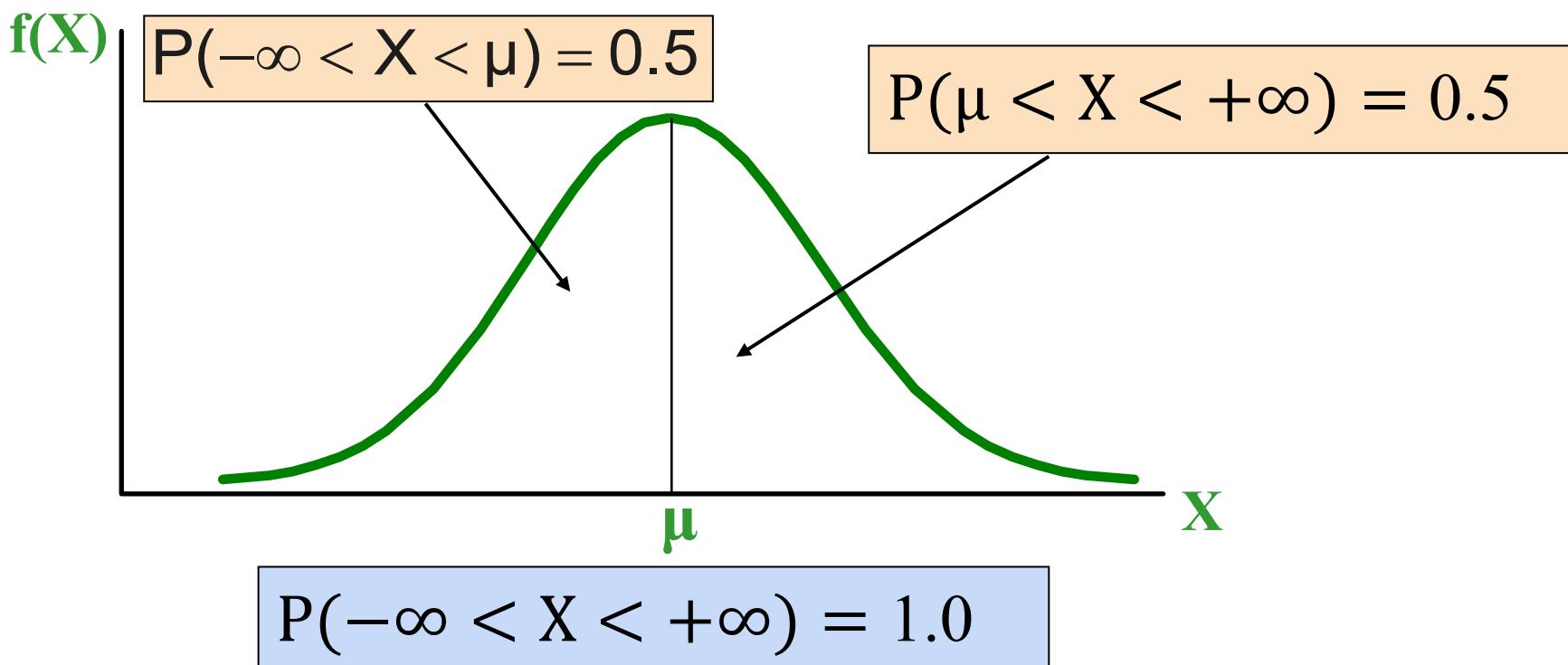




# 标准正态分布

计算概率

- 正态分布曲线下的总面积为1，因为曲线是对称的，因此均值左右两半的曲线下的面积皆为0.5

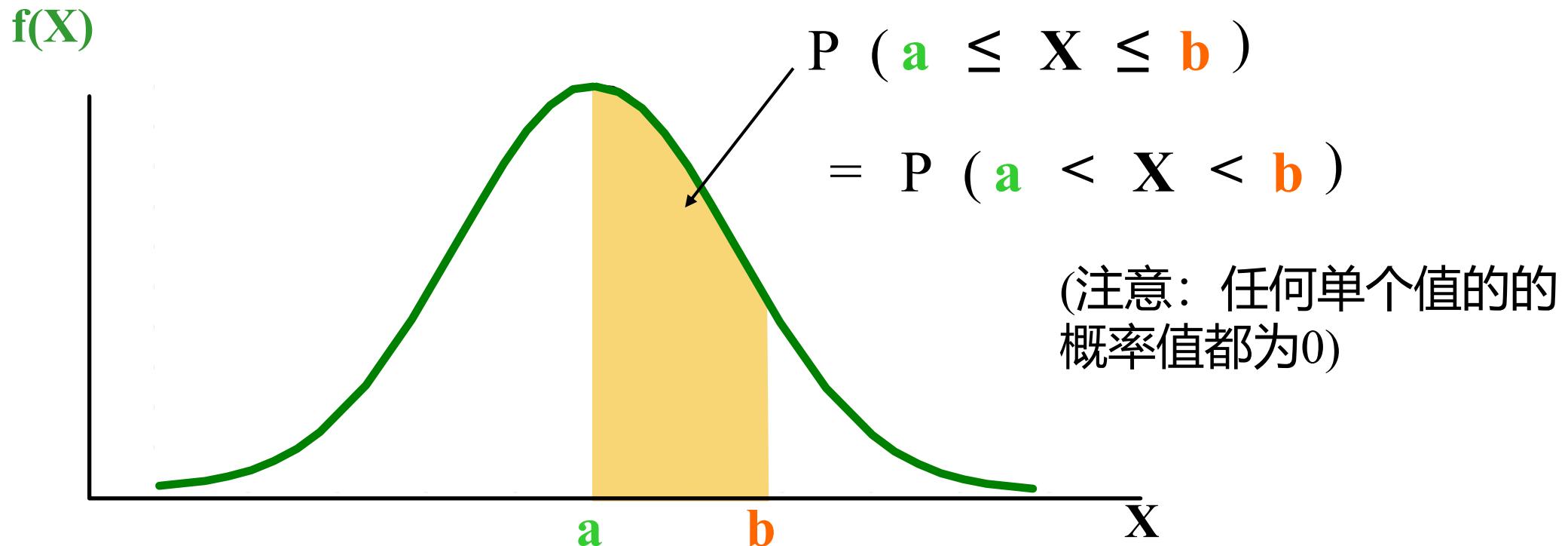




# 标准正态分布

## 计算概率

- 概率是通过曲线下的面积计算的

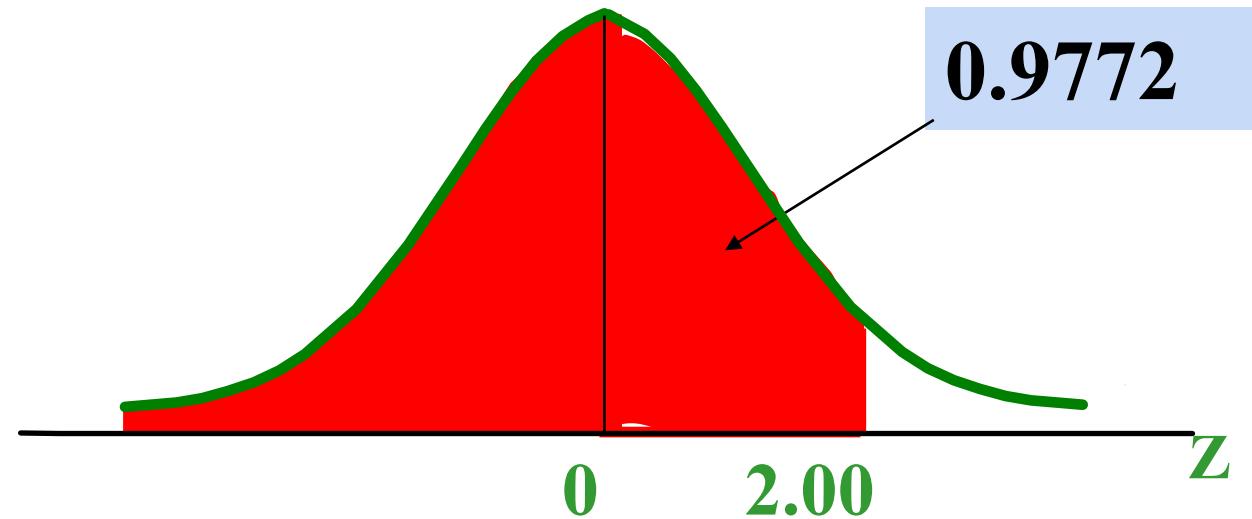


# 标准正态分布

## 标准正态分布表

- 通过查询正态分布表，我们可以得知小于某个特定的Z值（从负无穷到Z）的概率。

$$P(Z < 2.00) = 0.9772$$



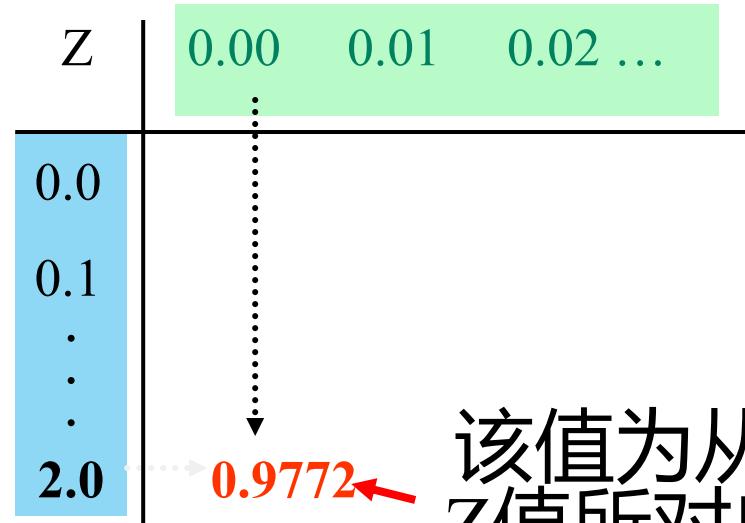


# 标准正态分布

## 标准正态分布表

行显示的是Z值小数点的第二位

列显示的是  
Z值的小数  
点第一位



该值为从 $-\infty$ 到给定  
Z值所对应的概率

$$P(Z < 2.00) = 0.9772$$



# 标准正态分布

## 计算正态分布概率的一般程序

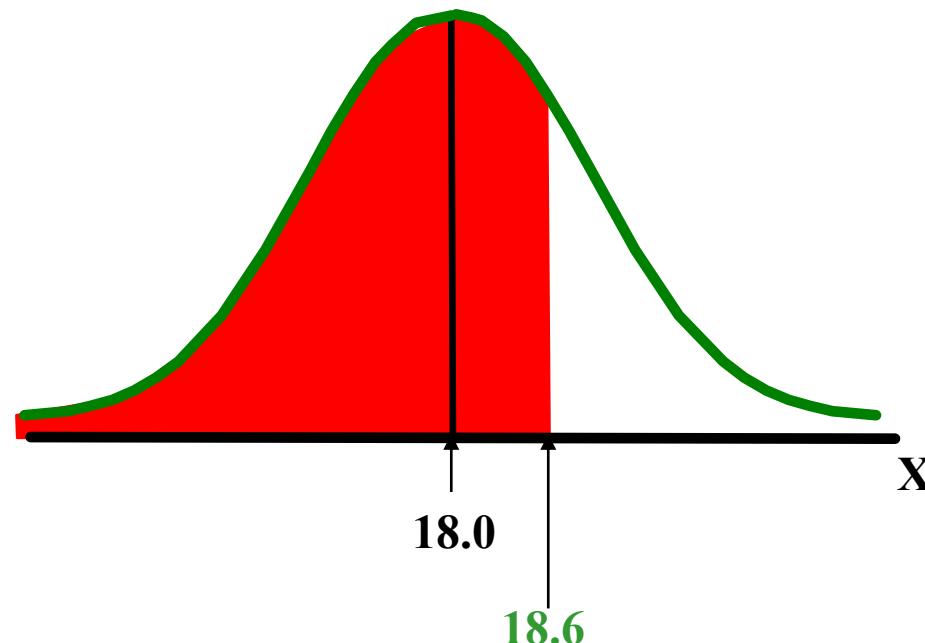
当 $X$ 服从正态分布时，计算 $P(a < X < b)$ ：

- 画出 $X$ 的正态分布曲线
- 将 $X$ 值转化成 $Z$ 值
- 查询标准正态分布表

## 标准正态分布

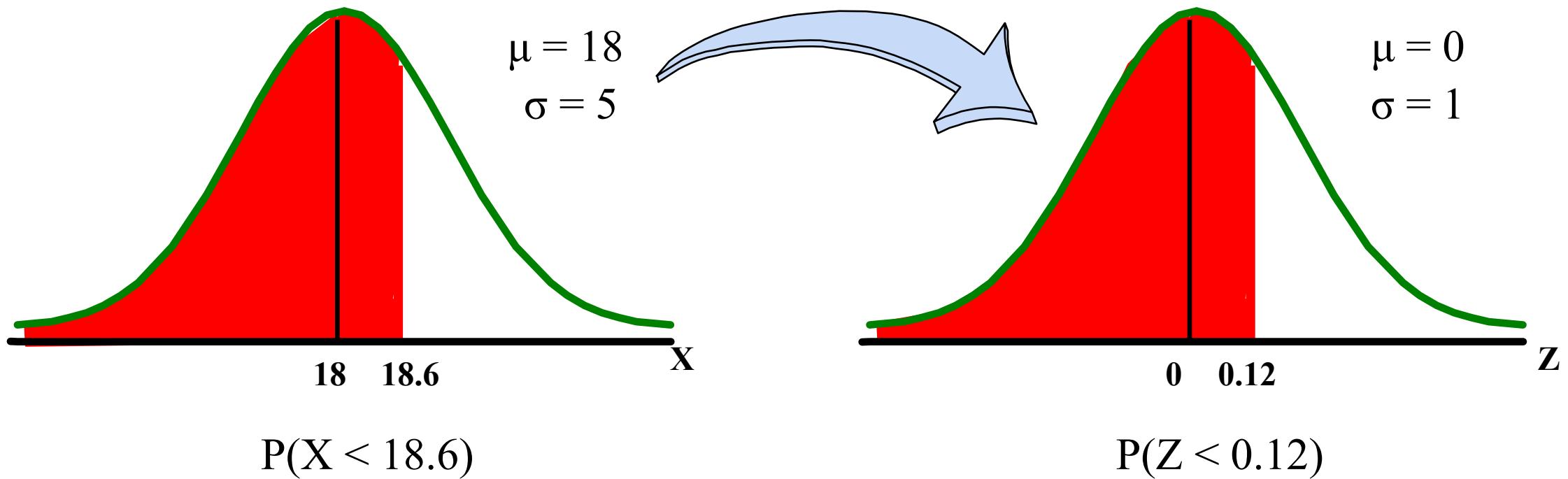
- $X$ 代表的是网络下载图片所需要的时间（单位为秒）
- 假如 $X$ 服从均值为18.0，标准差为5.0的正态分布。

求 $P(X < 18.6)$



# 标准正态分布

$$Z = \frac{X - \mu}{\sigma} = \frac{18.6 - 18.0}{5.0} = 0.12$$



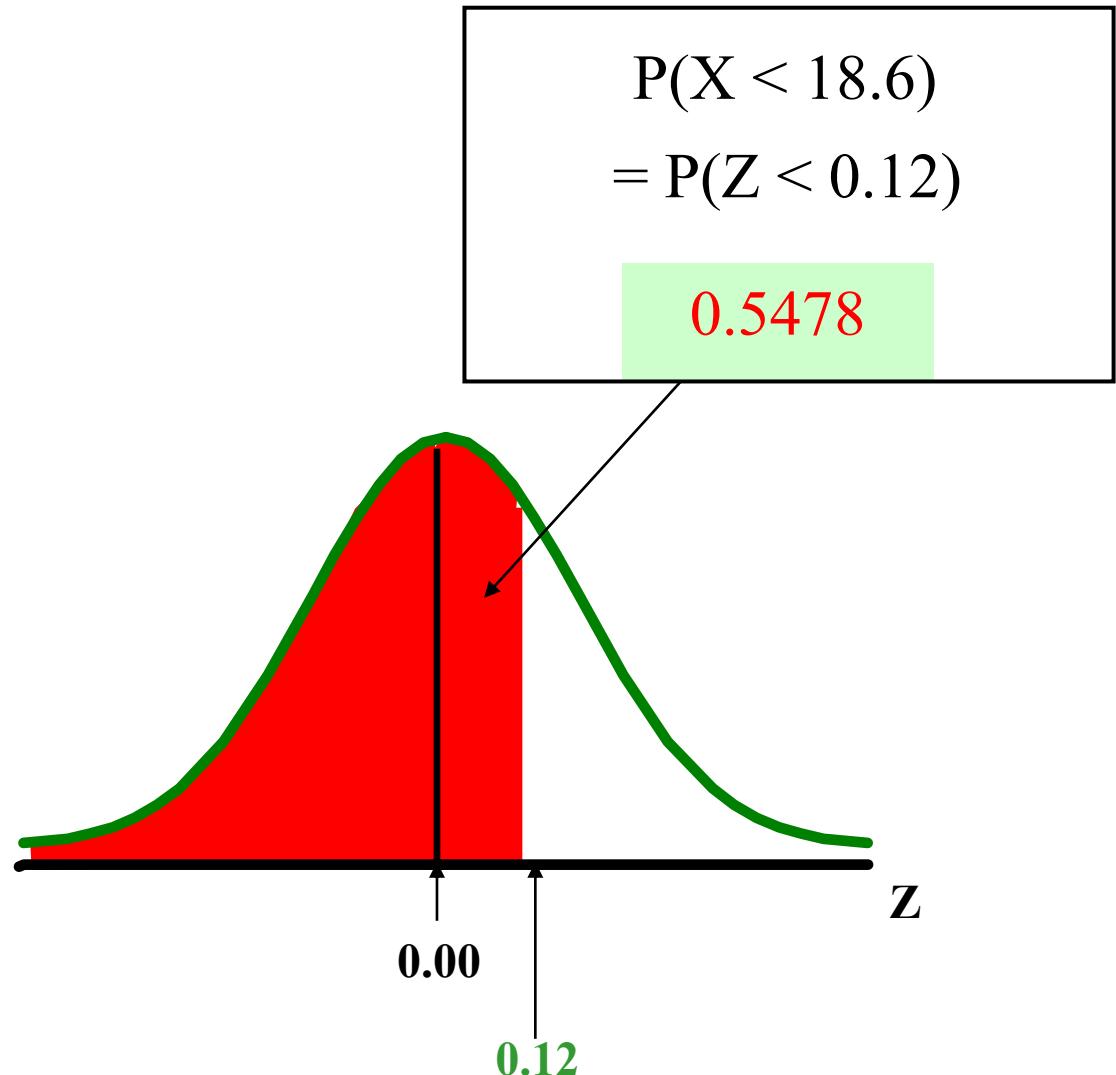
# 标准正态分布

标准正态分布表 (部分)

Z	.00	.01	.02
0.0	.5000	.5040	.5080
0.1	.5398	.5438	.5478
0.2	.5793	.5832	.5871
0.3	.6179	.6217	.6255

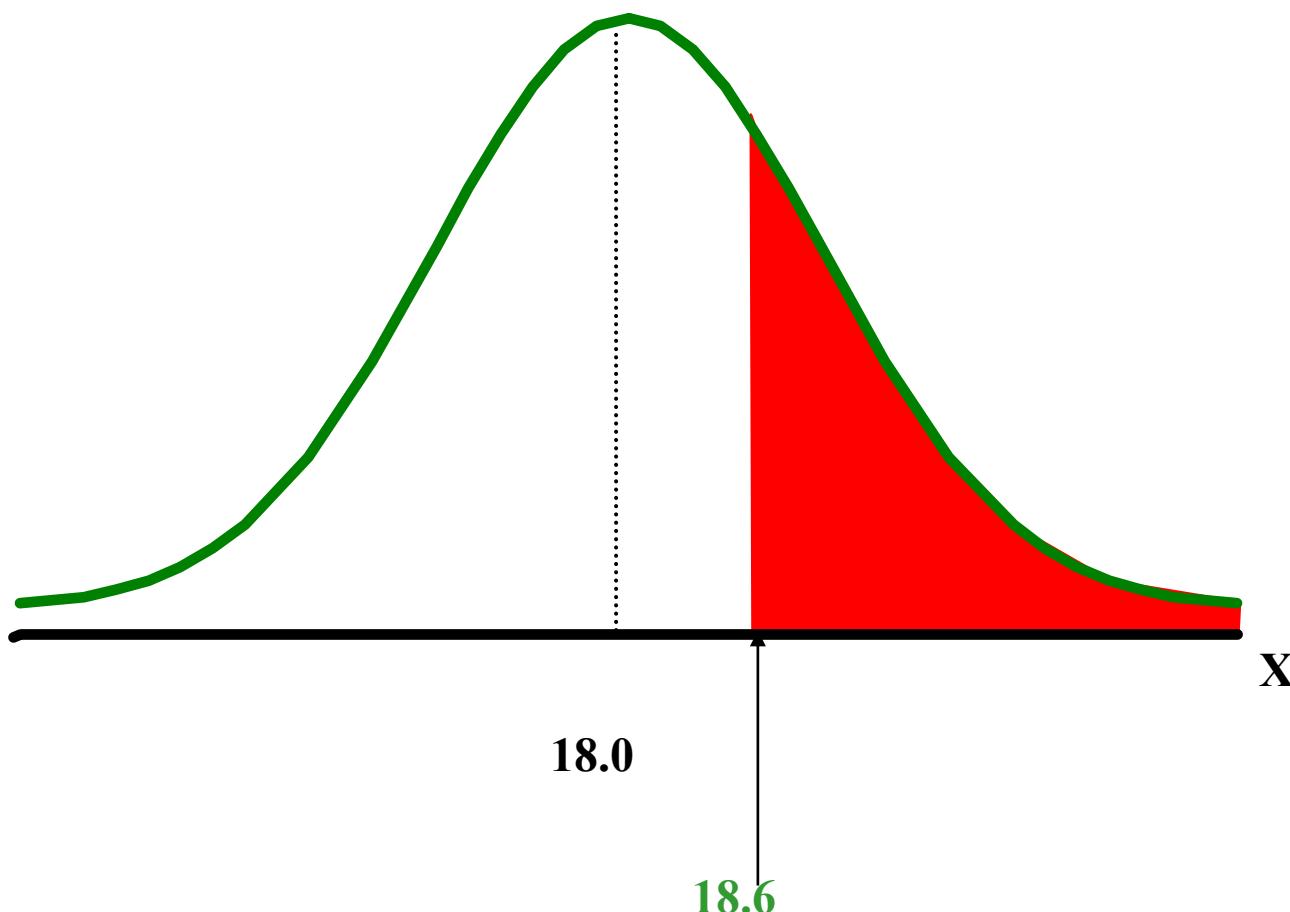
$$\begin{aligned}P(X < 18.6) \\= P(Z < 0.12)\end{aligned}$$

0.5478



# 标准正态分布

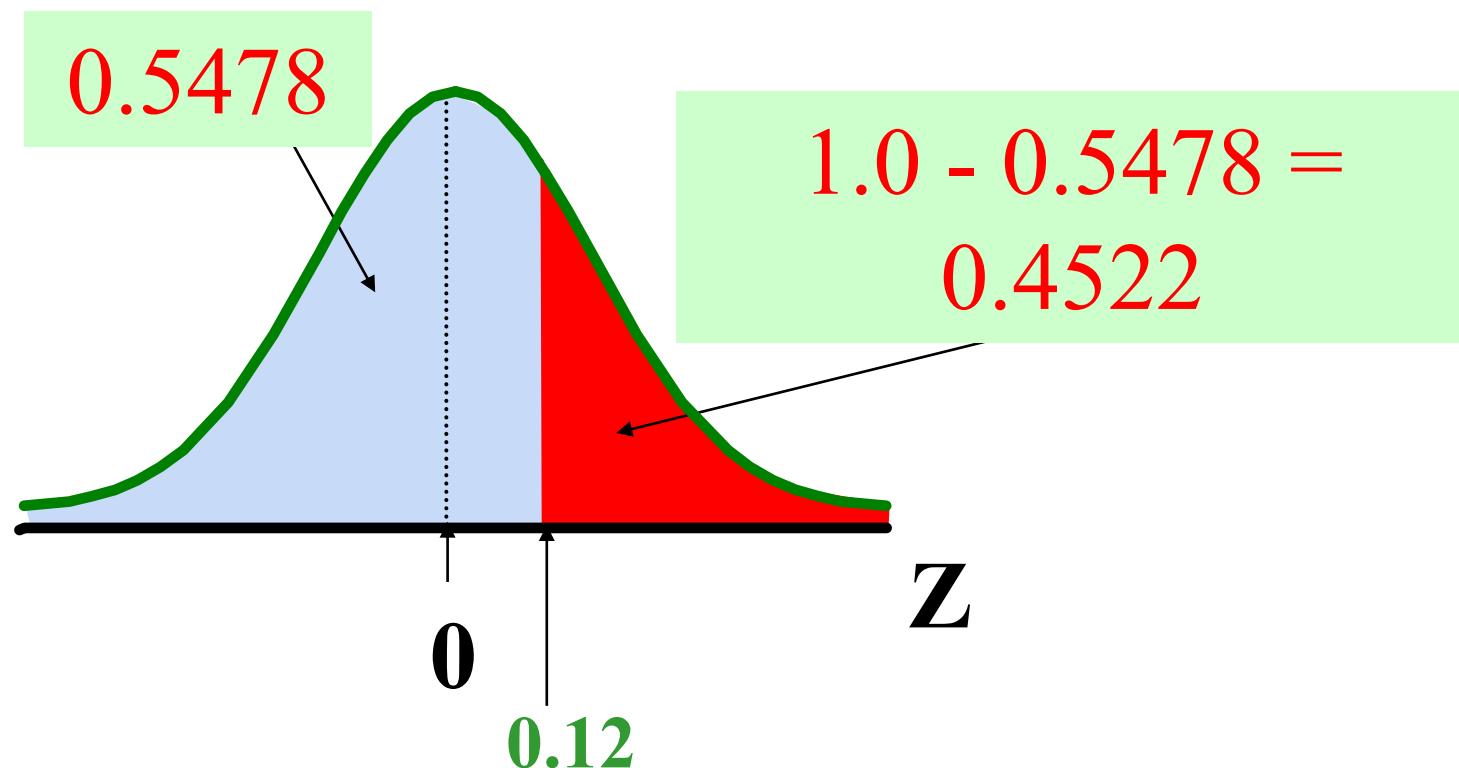
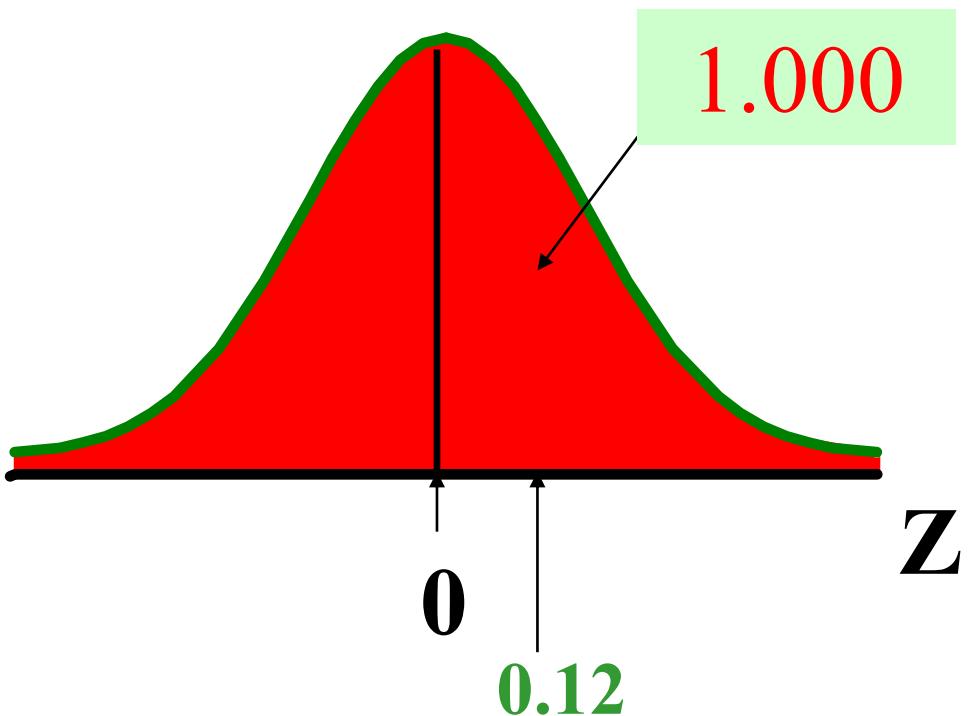
求 $P(X > 18.6)$



# 标准正态分布

$$P(X > 18.6) = P(Z > 0.12) = 1.0 - P(Z \leq 0.12)$$

$$= 1.0 - 0.5478 = 0.4522$$



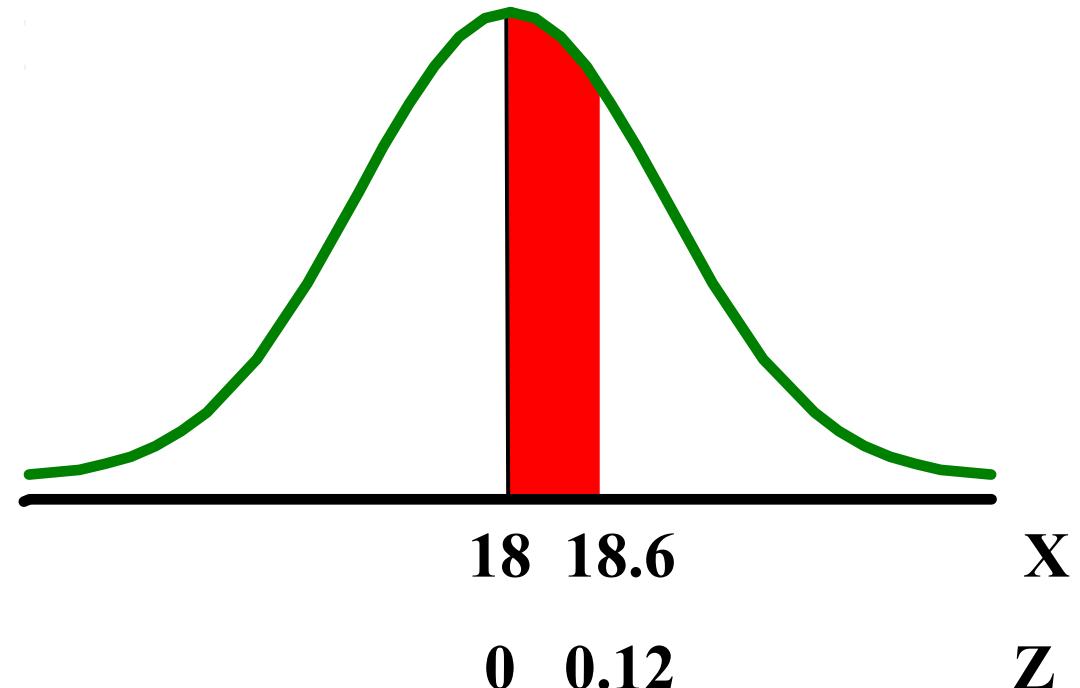
# 标准正态分布

求 $P(18 < X < 18.6)$

计算Z值

$$Z = \frac{X - \mu}{\sigma} = \frac{18 - 18}{5} = 0$$

$$Z = \frac{X - \mu}{\sigma} = \frac{18.6 - 18}{5} = 0.12$$



$$\begin{aligned}P(18 < X < 18.6) \\= P(0 < Z < 0.12)\end{aligned}$$

# 标准正态分布

标准正态分布表 (部分)

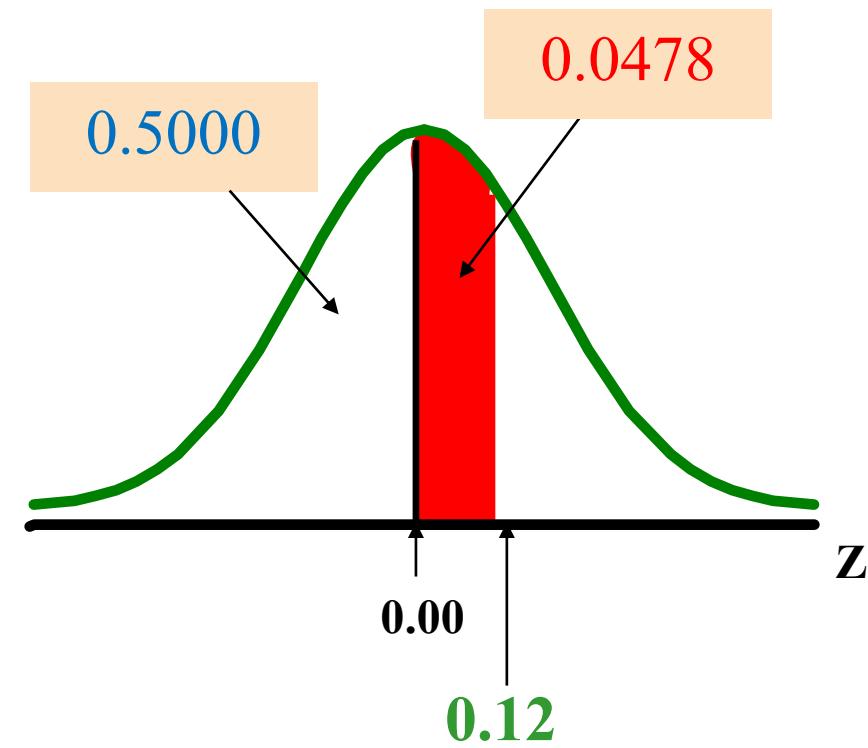
Z	.00	.01	.02
0.0	.5000	.5040	.5080
0.1	.5398	.5438	.5478
0.2	.5793	.5832	.5871
0.3	.6179	.6217	.6255

$$P(18 < X < 18.6)$$

$$= P(0 < Z < 0.12)$$

$$= P(Z < 0.12) - P(Z \leq 0)$$

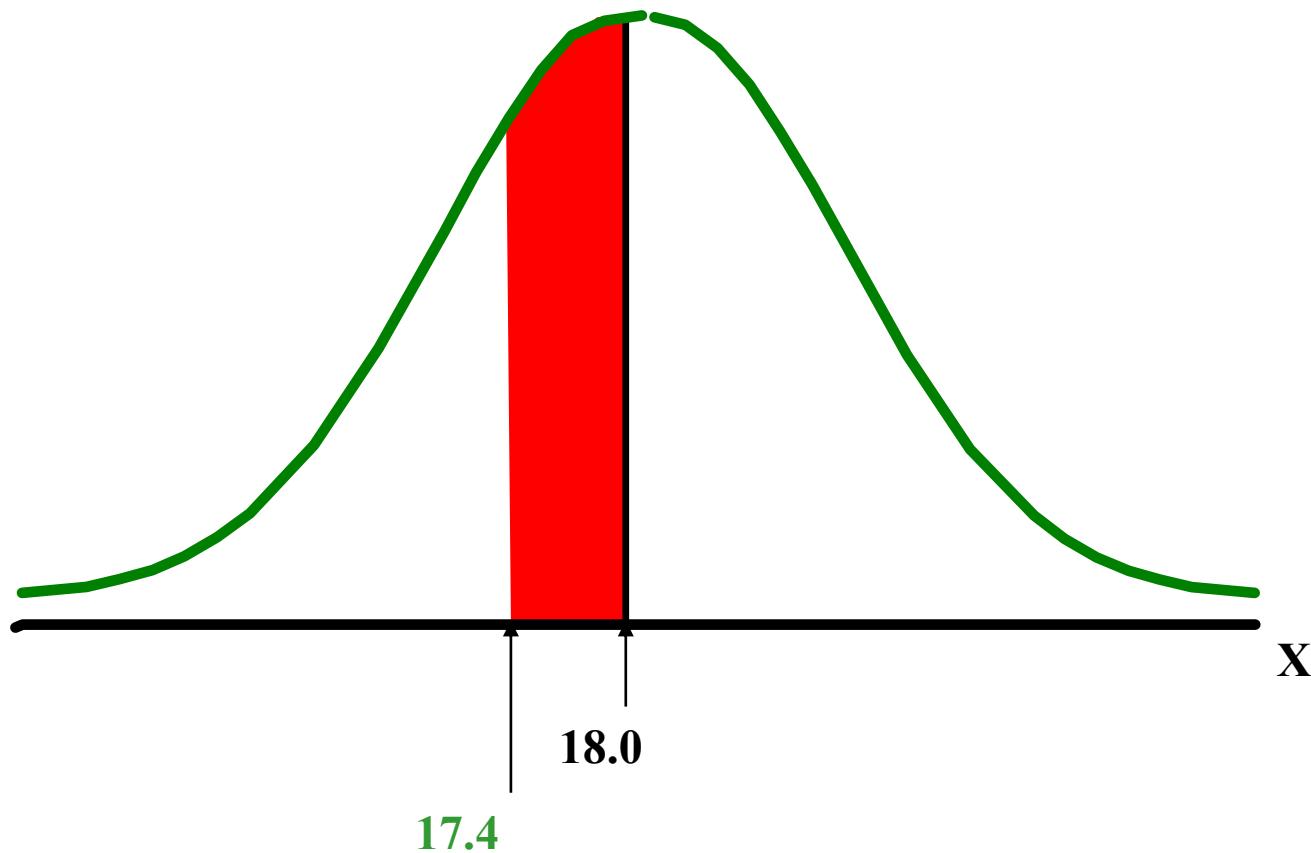
$$= 0.5478 - 0.5000 = 0.0478$$



# 标准正态分布



求 $P(17.4 < X < 18)$



# 标准正态分布



## 例子：算算智商

- 假如人群的智商服从正态分布，且均值 $\mu=100$ ，标准差 $\sigma=15$ 。
  - 电影《阿甘正传》的主角阿甘的智商为75，请算一下比他聪明的人有多少？（说出百/千分比）
  - 据说高智商俱乐部门萨俱乐部的入会门槛为智商位于人群的前2%，请算一下智商位于人群的前2%的智商值为多少？

# 标准正态分布

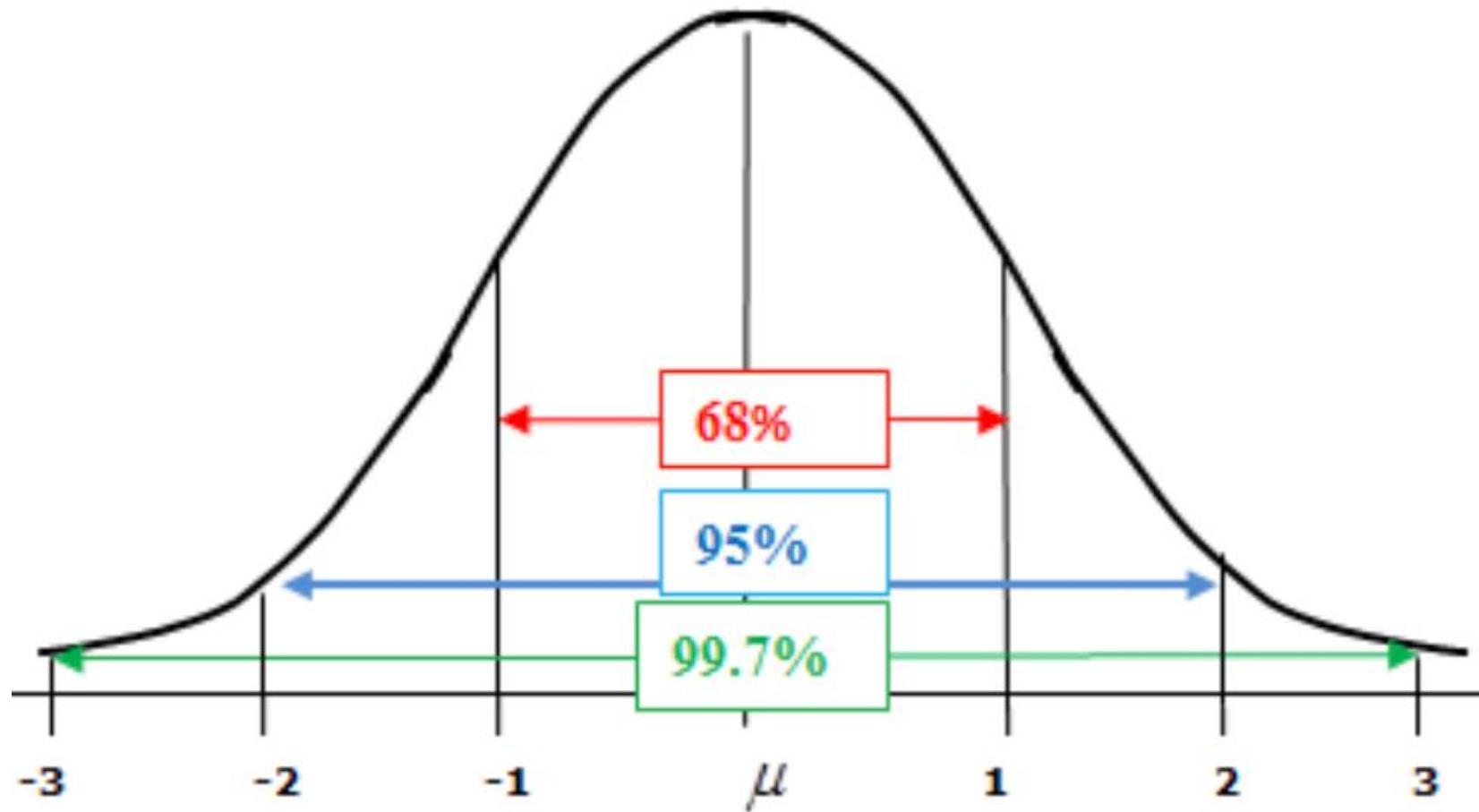


## 例子：挂科人数有多少？

- 某学生选修了一门统计课，TA很担心期末会挂科。据了解，上过该门课的1000名学生的分数成绩合乎正态分布，其平均成绩为73分，标准差10分，想知道被挂的概率有多少？过去分数不及格（60分以下）的人数约有几人？
- 先算标准计分得到 $Z = (60 - 73) / 10 = -1.3$ ，查表得知60分以下人数约占9.68%，也就是说大约有9.68%的概率被挂， $1000 \times 9.68\% = 96.8$ ，过去大概有96个人不及格。



## 68-95-99.7规则



## 例子：IQ测验

成绩为正态分布，平均为100分，标准差为10分，则

- 测验成绩在90分到110分之间，**约占**68%
- 测验成绩在80分到120分之间，**约占**95%
- 测验成绩在70分到130分之间，**约占**99.7%

### 例子：班上同学的身高

某班的学生身高为165.8厘米，标准差为12.38厘米。

- 身高在153.42到178.18厘米之间，约占68%
- 在141.04到190.56厘米之间，约占95%
- 在128.66到202.94厘米之间，约占99.7%

谢谢！