**方差**

**方差：**方差是在概率论和统计方差衡量随机变量或一组数据时离散程度的度量。概率论中方差用来度量随机变量和其数学期望（即均值）之间的偏离程度。统计中的方差（样本方差）是每个样本值与全体样本值的平均数之差的平方值的平均数。在许多实际问题中，研究方差即偏离程度有着重要意义。方差是衡量源数据和期望值相差的度量值。

**方差的定义：**方差在统计描述和概率分布中各有不同的定义，并有不同的公式。

在统计描述中，方差用来计算每一个变量（观察值）与总体均数之间的差异。为避免出现离均差总和为零，离均差平方和受样本含量的影响，统计学采用平均离均差平方和来描述变量的变异程度。总体方差计算公式：

IMG_256

IMG_257为总体方差，IMG_258为变量，IMG_259为总体均值，IMG_260为总体例数。

实际工作中，总体均数难以得到时，应用样本统计量代替总体参数，经校正后，样本方差计算公式：

S^2= ∑(X-IMG_261) ^2 / (n-1) [2] 

S^2为样本方差，X为变量，IMG_262为样本均值，n为样本例数。

在概率分布中，设X是一个离散型[随机变量](https://baike.baidu.com/item/%E9%9A%8F%E6%9C%BA%E5%8F%98%E9%87%8F" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)，若E{[X-E(X)]^2}存在，则称E{[X-E(X)]^2}为X的方差，记为D(X),Var(X)或DX，其中E(X)是X的期望值，X是变量值 [1]，公式中的E是期望值expected value的缩写，意为“变量值与其期望值之差的平方和”的期望值。[2]离散型随机变量方差计算公式：

D(X)=E{[X-E(X)]^2}=E(X^2) - [ E(X)]^2

当D(X)=E{[X-E(X)]^2}称为变量X的方差，而

IMG_263

称为标准差**（**或[均方差](https://baike.baidu.com/item/%E5%9D%87%E6%96%B9%E5%B7%AE" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)**）**。它与X有相同的量纲。标准差是用来衡量一组数据的离散程度的[统计量](https://baike.baidu.com/item/%E7%BB%9F%E8%AE%A1%E9%87%8F" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank) [3]  。

对于连续型随机变量X，若其定义域为(a,b)，概率密度函数为f(x)，连续型随机变量X方差计算公式：

D(X)=IMG_264(x-μ)^2 f(x) dx [2]

方差刻画了随机变量的取值对于其数学期望的[离散程度](https://baike.baidu.com/item/%E7%A6%BB%E6%95%A3%E7%A8%8B%E5%BA%A6" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)。（标准差、方差越大，离散程度越大)

若X的取值比较集中，则方差D(X)较小,若X的取值比较分散，则方差D(X)较大。

因此，D(X)是刻画X取值分散程度的一个量，它是衡量取值分散程度的一个尺度。

**方差的性质：**

1、设C是[常数](https://baike.baidu.com/item/%E5%B8%B8%E6%95%B0" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)，则D（C）=0

2、设X是[随机变量](https://baike.baidu.com/item/%E9%9A%8F%E6%9C%BA%E5%8F%98%E9%87%8F" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)，C是常数，则有

IMG_256

3、设 X 与 Y 是两个随机变量，则

IMG_257

其中协方差

IMG_258

特别的，当X，Y是两个不相关的随机变量则

IMG_259

此性质可以推广到有限多个两两不相关的随机变量之和的情况。

4、D(X)=0的[充分必要条件](https://baike.baidu.com/item/%E5%85%85%E5%88%86%E5%BF%85%E8%A6%81%E6%9D%A1%E4%BB%B6" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)是X以概率1取常数E（X），即

IMG_260

（当且仅当X取常数值E（X）时的概率为1时，D（X）=0。）

注：不能得出X恒等于常数，当x是连续的时候X可以在任意有限个点取不等于常数c的值。

5、D(aX+bY)=a2DX+b2DY+2abCov(X,Y)。

**种类及计算：**

离散型方差 的计算式为：

IMG_256，其中IMG_257 。

而将上式展开后可得：

IMG_258

### 连续型方差

连续型方差的计算式为：

IMG_259 ，其中IMG_260 。将上式展开后可得：

IMG_261

以上两式是一样的，只是写法不同。

证明：由数学期望的性质得

IMG_262

IMG_263

**随机随机变量的期望和方差：**

### 离散型

X服从两点分布，则 IMG_256

X服从[超几何分布](https://baike.baidu.com/item/%E8%B6%85%E5%87%A0%E4%BD%95%E5%88%86%E5%B8%83" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)，即 IMG_257 ，则

IMG_258

X服从[二项分布](https://baike.baidu.com/item/%E4%BA%8C%E9%A1%B9%E5%88%86%E5%B8%83" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)，即IMG_259 ，则

IMG_260

X服从[泊松分布](https://baike.baidu.com/item/%E6%B3%8A%E6%9D%BE%E5%88%86%E5%B8%83" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)，即IMG_261 ,则

IMG_262

### 连续型

X服从[均匀分布](https://baike.baidu.com/item/%E5%9D%87%E5%8C%80%E5%88%86%E5%B8%83" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)，即 IMG_263 **,**则

IMG_264 ,

IMG_265

X服从[指数分布](https://baike.baidu.com/item/%E6%8C%87%E6%95%B0%E5%88%86%E5%B8%83" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)，即IMG_266 , 则

IMG_267

X 服从[正态分布](https://baike.baidu.com/item/%E6%AD%A3%E6%80%81%E5%88%86%E5%B8%83" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)，即IMG_268 , 则

IMG_269

X 服从[标准正态分布](https://baike.baidu.com/item/%E6%A0%87%E5%87%86%E6%AD%A3%E6%80%81%E5%88%86%E5%B8%83" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)，即IMG_270 , 则

IMG_271

**公式：**

方差是实际值与[期望值](https://baike.baidu.com/item/%E6%9C%9F%E6%9C%9B%E5%80%BC" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)之差[平方](https://baike.baidu.com/item/%E5%B9%B3%E6%96%B9" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)的**平均值**，而[标准差](https://baike.baidu.com/item/%E6%A0%87%E5%87%86%E5%B7%AE" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)是方差算术[平方根](https://baike.baidu.com/item/%E5%B9%B3%E6%96%B9%E6%A0%B9" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)。 [5]  在实际计算中，我们用以下公式计算方差。方差是各个数据与[平均数](https://baike.baidu.com/item/%E5%B9%B3%E5%9D%87%E6%95%B0" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)之差的平方的和的平均数,即

IMG_256 ，其中，x表示[样本](https://baike.baidu.com/item/%E6%A0%B7%E6%9C%AC" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)的平均数，n表示样本的数量，xi表示个体，而s^2就表示方差。而当用

IMG_257 作为样本X的方差的估计时，发现其数学期望并不是X的方差，而是X方差的IMG_258 倍，

IMG_259 的数学期望才是X的方差，用它作为X的方差的估计具有“[无偏性](https://baike.baidu.com/item/%E6%97%A0%E5%81%8F%E6%80%A7" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)”，所以我们总是用

IMG_260 来估计X的方差，并且把它叫做“[样本方差](https://baike.baidu.com/item/%E6%A0%B7%E6%9C%AC%E6%96%B9%E5%B7%AE" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)”。

方差是和中心偏离的程度，用来衡量一批数据的波动大小（即这批数据偏离平均数的大小）并把它叫做这组数据的方差，记作S2。 在[样本容量](https://baike.baidu.com/item/%E6%A0%B7%E6%9C%AC%E5%AE%B9%E9%87%8F" \t "https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/_blank)相同的情况下，方差越大，说明数据的波动越大，越不稳定。

公式可以进一步推导为：

IMG_261

。其中x为这组数据中的数据，n为大于0的整数。