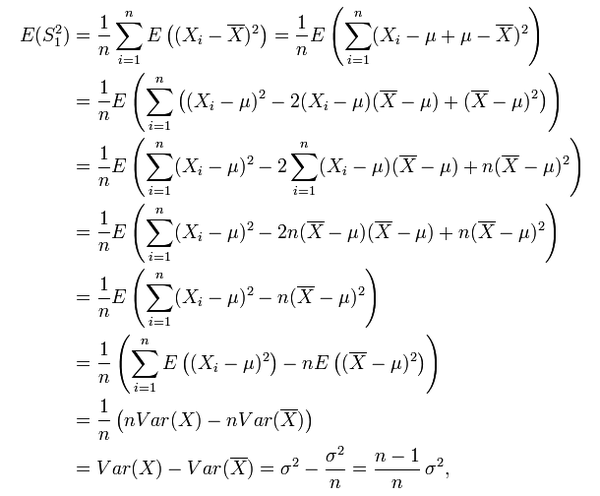
本來，按照定義，方差的 estimator 應該是這個：  
https://pic4.zhimg.com/933d8c8c9473c3c0deb2d4d267fc368f_b.jpg但，這個 estimator 有 bias，因為：  
  
而 (n-1)/n \* σ² **!=** σ² ，所以，為了避免使用有 bias 的 estimator，我們通常使用它的修正值 S²：  
https://pic4.zhimg.com/04dba7fc15e216a83de2ea138993118f_b.jpg

看下面这个等式的数学证明：  
\mathbb{E}\Big[\frac{1}{n-1} \sum_{i=1}^n\Big(X_i -\bar{X}\Big)^2 \Big]=\sigma^2.  
但是这个答案显然不够直观（教材里面统计学家像变魔法似的不知怎么就得到了上面这个等式）。  
下面我将提供一个略微更友善一点的解释。  
==================================================================  
===================== 答案的分割线 ===================================  
==================================================================  
首先，我们假定随机变量X的数学期望\mu是已知的，然而方差\sigma^2未知。在这个条件下，根据方差的定义我们有  
\mathbb{E}\Big[\big(X_i -\mu\big)^2 \Big]=\sigma^2, \quad\forall i=1,\ldots,n,  
  
由此可得  
**\mathbb{E}\Big[\frac{1}{n} \sum_{i=1}^n\Big(X_i -\mu\Big)^2 \Big]=\sigma^2.**  
**因此\frac{1}{n} \sum_{i=1}^n\Big(X_i -\mu\Big)^2 是方差\sigma^2的一个无偏估计，注意式中的分母不偏不倚正好是n！**  
这个结果符合直觉，并且在数学上也是显而易见的。  
  
现在，我们考虑随机变量X的数学期望\mu是未知的情形。这时，我们会倾向于无脑直接用样本均值\bar{X}替换掉上面式子中的\mu。这样做有什么后果呢？后果就是，  
**如果直接使用\frac{1}{n} \sum_{i=1}^n\Big(X_i -\bar{X}\Big)^2 作为估计，那么你会倾向于低估方差！**  
这是因为：  
  
换言之，除非正好\bar{X}=\mu，否则我们一定有  
\frac{1}{n}\sum_{i=1}^n(X_i-\bar{X})^2 <\frac{1}{n}\sum_{i=1}^n(X_i-\mu)^2 ,  
而不等式右边的那位才是的对方差的“正确”估计！  
这个不等式说明了，为什么直接使用\frac{1}{n} \sum_{i=1}^n\Big(X_i -\bar{X}\Big)^2 会导致对方差的低估。  
  
那么，在不知道随机变量真实数学期望的前提下，如何“正确”的估计方差呢？答案是把上式中的分母n换成n-1，通过这种方法把原来的偏小的估计“放大”一点点，我们就能获得对方差的正确估计了：