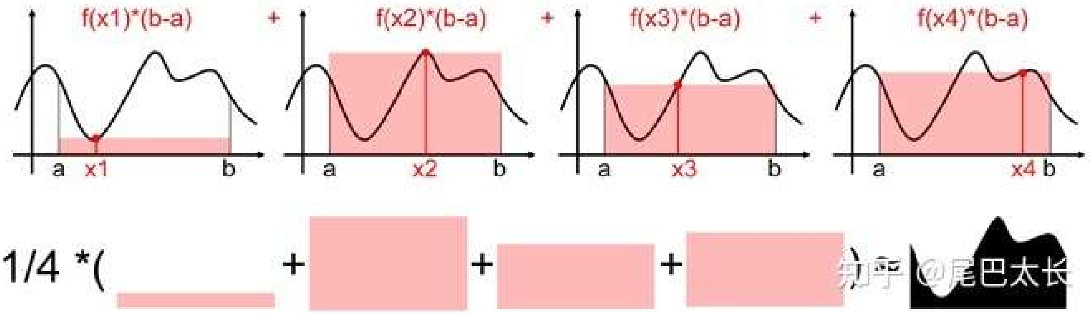
https://zhuanlan.zhihu.com/p/396653562

先看第一图



这幅图表明了什么意思呢？我们知道，计算[a,b]内的定积分就是求曲线 f(x)、直线 x=a,x=b 以及 x 轴围成的形状的面积，因此，如果我们在曲线上随机地选取 N 个点，计算如图所示的粉红色长方形面积之和，再求个平均，其实就得到了定积分的近似值。点的数量取得越多，这个平均值就越逼近定积分的真实值。

用公式写出来就是：

$$\frac{1}{N}[(b-a)\times f(X_1)+(b-a)\times f(X_2)+\cdots+(b-a)\times f(X_N)]=\frac{b-a}{N}\sum_{i=1}^{N}f(X_i)$$

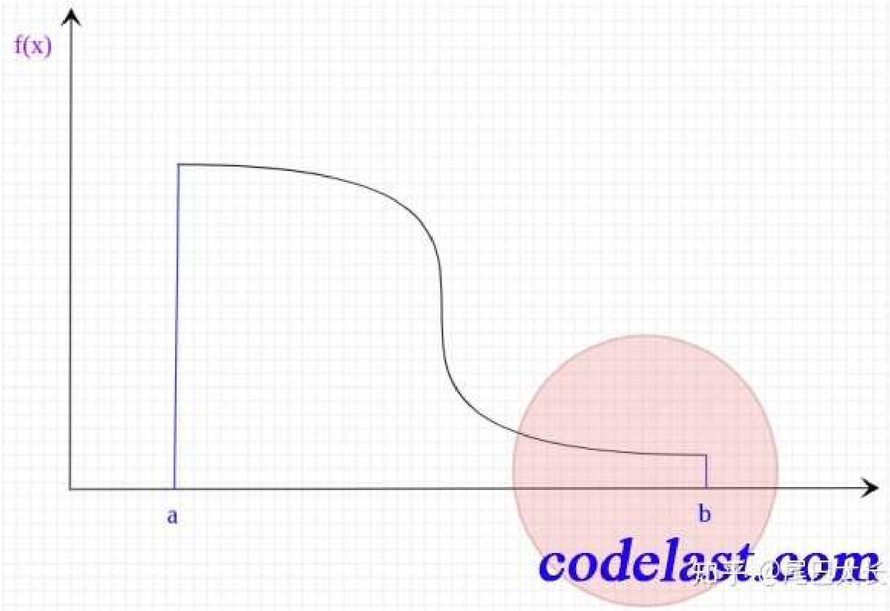
表达式为：

$$F^N=\frac{1}{N}\sum_{i=1}^{N}\frac{f(X_i)}{p(X_i)}$$

这个式子没有积分符号∫，但是它却叫做“积分”公式，这是因为这个式子求的是积分的近似值——当 N 越大的时候，计算出的值就越接近定积分的真实值。

在公式中，有一个奇怪的东西，就是 p(Xi)，它表示 Xi 这个点，在某个分布下取 Xi 这个值的概率。那么这个分布是什么呢？比如说，它能不能是简单的均匀分布？后面我们会看到，这个分布是我们自己选取的。

既然是要随机采样 N 个点，那么是不是随便用什么样的策略去采样，都可以达到同样的效果呢？这里用一幅图来说明，采样也是要讲究策略的，否则效果会很差：



由于定积分值就是曲线下的面积，显然，如果我们采样的点恰巧大部分处于圆圈内，那么这些点下的面积之和必然比较小，此时，我们按前面所说的计算矩形面积的方法算得的积分值，是远远不能反映积分的真实值的。也就是说，圆圈处的点，对积分值的贡献小，靠近 x=ax=a 处的曲线上的点，对积分值的贡献大。

所以，在实际采样的时候，靠近圆圈处的点应该少采一些，非圆圈处的点应该多采一些。

这就是重要性采样（Importance Sampling）的概念由来了——采样要按“重要性”来进行，不应该“平等对待”。

如果采样恰到好处的话，可能只需要进行很少的采样（计算若干个点的函数值），就可以求出误差很小的积分值。