

在统计学中，一个概率样本的置信区间（Confidence interval）是对这个样本的某个总体参数的区间估计。

例题：从农场收获的20万个苹果中抽取36个作为样本。样本中苹果重量的均值为112克，标准差为40克。问20万个苹果的重量均值处在100到124克之间的概率是多少？

(1) 求抽样分布

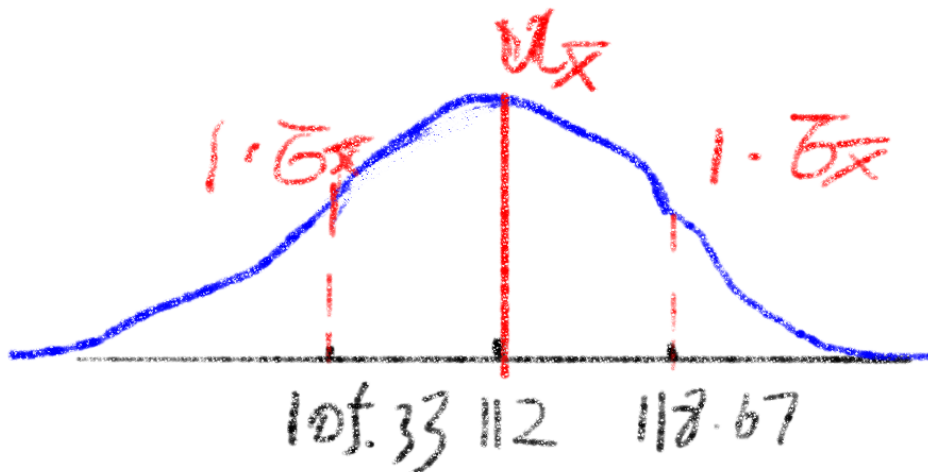
首先，我们求抽样分布的均值。抽样分布的均值等于总体均值，它的标准差等于：

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

我们不知道总体的标准差，只能用最好的估计值来代替总体标准差，这个最好的估计值可以采用样本的均值，也就是40克。因此可以求得抽样分布的标准差为：

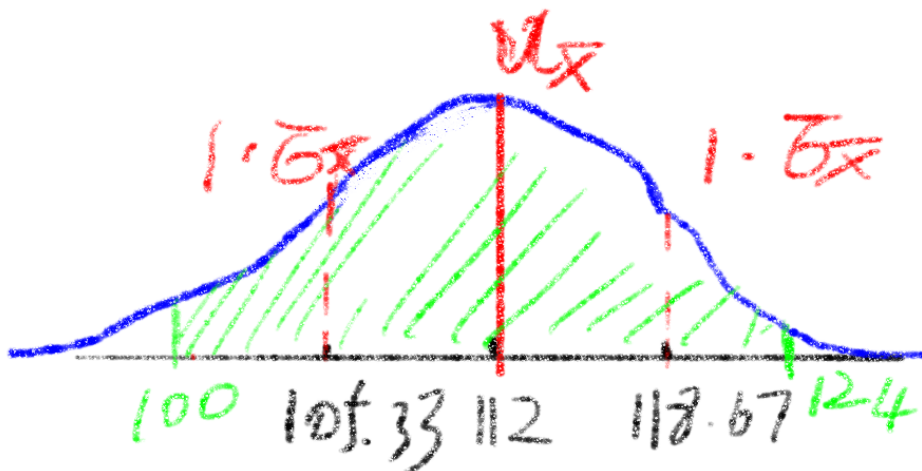
$$\sigma_{\bar{x}} = \frac{40}{\sqrt{36}} = 6.67$$

因此，我们知道抽样分布的情况：



(2) 求区间的概率

求重量在100到124克之间的概率，也就是求标准差为12克之内的苹果的概率。即下图绿色区域的面积：



那么12对应多少个标准差呢？也就是z分数：

$$\frac{12}{6.67} = 1.8$$

因为对称性，我们可以先求112到124之间的概率。通过查z分数表，我们知道小于124（也就是1.8个z分数）的概率为0.9641，我们需要减去112克（z分数为0）以下的概率：0.5也就是求得112到124之间的概率为0.4641

最后，根据对称性，求得20万个苹果的重量均值处在100到124克之间的概率为：

$$0.4641 \times 2 = 0.9282 = 92.82\%$$