1. 中心极限定理

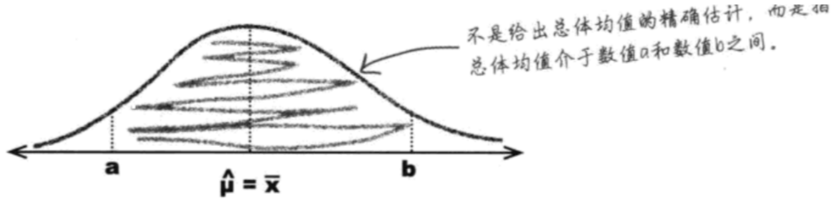
样本的平均值约等于总体的平均值。不管总体是什么分布，任意一个总体的样本平均值都会围绕在总体的整体平均值周围，并且呈正态分布。
演示小程序地址：https://link.zhihu.com/?target=http%3A//onlinestatbook.com/stat_sim/sampling_dist/index.html

2. 样本均值的抽样分布

- 总体参数已知，需要某个特定样本的概率。因为假设样本均值的抽样分布也是服从正态分布的，实际上就是求标准差。
  样本抽样分布的标准差 = 总体标准差 / n^0.5
  样本标准差是一个点估计量和正态分布里面的标准差不是同一个东西
- 伯努利分布均值和方差
  均值 = p，方差 = n(1-p)

3. 置信区间

样本有时候无法给出足够正确的结果。置信区间是考虑了不确定性的方法。



求解置信区间的四个步骤：

- 选择总体统计量
- 求出其抽样分布

  $E(\bar{X}) = \mu \qquad Var(\bar{X}) = \frac{\sigma^2}{n}$

  而 σ^2 是未知的，这时候可以使用样本的方差来计算。 $Var(\bar{X}) = \frac{s^2}{n}$
- 决定置信水平(百分多少)
- 求出置信上下限

置信区间的简易算法

总体统计量	总体分布	条件	置信区间
μ	正态	σ^2已知 n 可大可小 $\bar{x}$为样本均值	$\left(\bar{x}-c\frac{\sigma}{\sqrt{n}}, \bar{x}+c\frac{\sigma}{\sqrt{n}}\right)$
μ	非正态	σ^2已知 n 很大（至少30） $\bar{x}$为样本均值	$\left(\bar{x}-c\frac{\sigma}{\sqrt{n}}, \bar{x}+c\frac{\sigma}{\sqrt{n}}\right)$
μ	正态或非正态	σ^2未知 n 很大（至少30） $\bar{x}$为样本均值 s^2为样本方差	$\left(\bar{x}-c\frac{s}{\sqrt{n}}, \bar{x}+c\frac{s}{\sqrt{n}}\right)$
p	二项	n 很大 p_s 为样本比例 q_s 等于$1-p_s$	$\left(p_s-c\sqrt{\frac{p_s q_s}{n}}, p_s+c\sqrt{\frac{p_s q_s}{n}}\right)$