1. 中心极限定理.

(1) 样本期望符合正态分布

$\{x_1, x_2, \cdots, x_n\}$ iid. 未必是正态分布
但要已知 $E[X_i]=\mu$
$var(X_i)=\sigma^2$.

$$\frac{\sum_{i=1}^{n} x_i}{n} \sim N(?, ?)$$

(2) 期望: $E\left[\frac{\sum_{i=1}^{n} x_i}{n}\right]=\mu$

方差: $var\left(\frac{\sum x_i}{n}\right)=\frac{1}{n^2}var(\sum x_i)$
$=\frac{1}{n^2}\sum var(x_i)$
$=\frac{\sigma^2}{n}$

因此, $\frac{\sum x_i}{n} \sim N(\mu, \frac{\sigma^2}{n})$

(3) 化简

$$\frac{\frac{\sum x_i}{n}-\mu}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0, 1)$$

$$\frac{\sum x_i - n\mu}{\sigma\sqrt{n}} \sim N(0, 1)$$

2. 大数定理: 可用样本均值近似总体均值.

例1. 100题 判断 70

扔100次硬币 70个head 概率

$\mu = \cancel{50\%}\ 0.5$

$\sigma = \sqrt{0.5\cdot 0.5} = 0.5$

$$\frac{X_1+\dots+X_{100}-50}{0.5\cdot 10} \sim N(0,1)$$

$S = X_1+\dots+X_{100} > 70$

$$= \frac{S-50}{5} \sim N(0,1)$$

$P(S>70)$, $\frac{S-50}{5} \sim N(0,1)$

$$P\left(\frac{S-50}{5} > \frac{70-50}{5}\right)$$

$$= P(Z>4) \approx 0$$

例2. 1000个乘客 2条铁路 S个座位 n个乘客

Want: $P(n > S) \le 1\%$

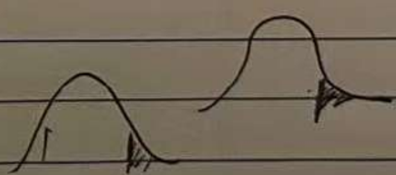
$\{X_1, \dots, X_{1000}\}$

$S = X_1+\dots+X_{1000}$ 上第一辆车的总人数

$= n$

$\mu = 0.5$

$\sigma = 0.5$



$$\frac{S-500}{0.5\sqrt{1000}} \sim N(0,1)$$

$P(n>S) = \cancel{P\left(\frac{n-500}{5\sqrt{10}} > z\right)} \le 1\%$

$= P\left(\frac{n-500}{5\sqrt{10}} > \frac{S-500}{5\sqrt{10}}\right) = \cancel{N(500-n)} = P\left(Z > \frac{S-500}{5\sqrt{10}}\right) \le 1\%$

$N\left(\frac{500-S}{5\sqrt{10}}\right) \le 1\%$

$\frac{500-S}{5\sqrt{10}} \le -2.32$

$S = 500 + 5\sqrt{10}\cdot 2.32$

$= 537$

3. 简单：第一层 x_i 分布　　一次实验中掷 n 次个硬币

↓

第二层 $\frac{\sum x_i}{n}$ 分布　　多次实验，虽然并不直的需要

复杂：第一层 x_i 分布　　掷硬币 n 次实验

m 次硬币 → 中位数

↓

⇒ n 个中位数

第二层 x_i 分布　　统计值的分布

↓

第三层 $\frac{\sum x_i}{n}$ 分布　　$N(?\ ,\ ?)$

↑

LLN

$\mu = \frac{1}{\lambda}$

$\sigma = \frac{1}{\lambda}$