**1 基本概念**

样本均值：

总体均值：

样本方差：

总体方差：

样本标准差：

总体标准差：

方差体现数据分布的离散程度，标准差统一单位

随机变量（随机过程映射到数值的函数）：

随机变量分离散型和连续性

**2 二次分布**

重复n次的伯努利实验每次试验只有两种可能的结果

投硬币为例子：设共投n次（head : 0.5, tail : 0.5），所可能结果为：

有 0 次正面朝上概率：

有 1 次正面朝上概率：

有 2 次正面朝上概率：

. . . . . .

有 n 次正面朝上概率：

当 n 无限大时，概率密度分布趋于钟型分布，即正态分布

以投篮为例子：设共投 n 次（make : 0.3，miss : 0.7），所有可能结果未：

有 0 次命中：

有 1 次命中：

有 2 次命中：

有 3 次命中：

. . . . . .

有 n 次命中：

以下为投篮6次命中率的统计

|  |  |  |  |
| --- | --- | --- | --- |
| N | 6 |  |  |
| P(make) | 0.7 |  |  |
| P(miss) | 0.3 |  |  |
|  |  |  |  |
| K | P(make)^K\*P(miss)^(N-K) | N choose K | P(X=K) |
| 0 | 0.000729 | 1 | 0.000729 |
| 1 | 0.001701 | 6 | 0.010206 |
| 2 | 0.003969 | 15 | 0.059535 |
| 3 | 0.009261 | 20 | 0.18522 |
| 4 | 0.021609 | 15 | 0.324135 |
| 5 | 0.050421 | 6 | 0.302526 |
| 6 | 0.117649 | 1 | 0.117649 |

图 1 命中率表格

图 2 命中率柱状图

随机变量的期望值即总体均值

证明如下： 









令, ,

则, ,

则







1. **泊松分布**

以每小时通过车辆数目为例子：将时间分得越细取极限，可由二项分布推导出泊松分布



将时间 n 划分越来越细，即 n 越来越大











证明得到泊松分布实际是极限下的二项分布

**4 大数定理**

样本数量足够多的时候样本均值趋近总体均值

赌徒谬论不成立，独立事件再次发生的结果不受前面事件的影响

**5 正态分布（又名高斯分布，钟型分布）**

日常生活中最常见的分布：



3σ准则68—95—99.7可以快速估计概率

标准正态分布期望 1 ，标准差 0

Z 分数表示距离均值相差多少个标准差 =(x-u)/σ