概率论与数理统计
第三章

随机变量的数字特征

大纲

1. 数学期望（均值）的定义
2. 数学期望（均值）的性质
3. 条件数学期望（条件均值）
4. 中位数

大纲

1. 数学期望（均值）的定义
2. 数学期望（均值）的性质
3. 条件数学期望（条件均值）
4. 中位数

引例

例 3.1.1. 一甲乙两人赌技相同，各出赌金100元，约定先胜三局者为胜，取得全部200元。现在甲胜2局乙胜1局的情况下中止，问赌本该如何分？

定义（离散型随机变量）

定义 3.1.1. 设X为一离散型随机变量，其分布律为

$$P(X = x_i) = p_i, \quad i = 1, 2, \cdots$$

如果$\sum_{i=1}^{\infty} |x_i| p_i < +\infty$，则称

$$\sum_{i=1}^{\infty} x_i p_i$$

为随机变量X的数学期望(均值)，用符号EX表示. 若$\sum_{i=1}^{\infty} |x_i| p_i = +\infty$，则称$X$的数学期望(均值)不存在.

定义（连续型随机变量）

定义 3.1.2. 如果连续型随机变量X具有密度函数$f(x)$, 则当

$$\int_{-\infty}^{\infty}|x|f(x)dx<\infty$$

时, 我们将积分

$$\int_{-\infty}^{\infty}xf(x)dx$$

的值称为X的数学期望, 记作EX. 如果

$$\int_{-\infty}^{\infty}|x|f(x)dx=\infty,$$

则称X的数学期望不存在.

常见分布的期望

- 二项分布 $X \sim B(n, p)$
- Poisson分布 $X \sim P(\lambda)$
- 正态分布 $X \sim N(\mu, \sigma^2)$
- 均匀分布 $X \sim U(a, b)$
- 指数分布 $X \sim Exp(\lambda)$

例子

例 3.1.3. (Cauchy分布)设

$$p(x) = \frac{1}{\pi(1+x^2)}, \quad x \in \mathcal{R},$$

则: 该分布的期望不存在.

大纲

1. 数学期望（均值）的定义
2. 数学期望（均值）的性质
3. 条件数学期望（条件均值）
4. 中位数

性质一

- 若干个随机变量的和的期望等于各变量的期望之和，即：

$$E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$$

这里我们假设各随机变量X_i的期望都存在。

- 例：
  - 假设随机变量$X \sim B(n,p)$，求 $E(X)$。

性质二

- 若干个独立随机变量的积的期望等于各变量的期望之积，即：

$$E(X_1X_2\cdots X_n)=E(X_1)E(X_2)\cdots E(X_n)$$

这里假设各随机变量X_i的期望都存在。

性质三：随机变量函数的期望

- 设随机变量X为离散型，有分布$P(X=a_i)=p_i, i=1,2,\ldots,$或者为连续型，有概率密度函数$f(x)$。则

$$
Eg(X)=\begin{cases}\sum_i g(a_i)p_i\,, & \sum_i |g(a_i)|p_i<\infty\ \ ; \\ \int_{-\infty}^{+\infty} g(x)f(x)dx, & \int_{-\infty}^{+\infty} |g(x)|f(x)dx\ <\infty. \end{cases}
$$

- 例：
  - 假设c为常数，则$E(cX)=c\,E(X)$。
  - 假设随机变量$X\sim N(0,1)$，求 $Y=X^2+1$的数学期望。

育天下英才 創寰宇學府
嚴濟慈 一九八八年五月 題

例子

- 机场大巴上有20位乘客，离开机场后总共有10各车站可以下车，若某个车站没有人下车则该车站不停车。设乘客在每个车站下车的可能性相等，以X表示停车的次数，求$E(X)$。

大纲

1. 数学期望（均值）的定义
2. 数学期望（均值）的性质
3. 条件数学期望（条件均值）
4. 中位数

条件数学期望

- 设X和Y为随机变量，若(X,Y)为离散型，且在给定$X=x$下，Y有分布$P(Y=a_i|X=x)=p_i, i=1,2,\dots,$ 或者(X,Y)为连续型，且在给定$X=x$下，Y的条件密度函数为$f(y|x)$。则

$$E(Y|X=x)=\begin{cases}\sum_i a_i p_i\,, & (X,Y)\text{为离散型；}\\ \int_{-\infty}^{+\infty} yf(y|x)dy, & (X,Y)\text{为连续型．}\end{cases}$$

- 例：
  - 设$(X,Y)\sim N(a,b,\sigma_1^2,\sigma_2^2,\rho)$，试计算$E(Y|X=x)$。

中国科学技术大学
University of Science and Technology of China

全期望公式

- 设X, Y为两个随机变量，则有

$$EX = E[E(X|Y)]$$

$$Eg(X) = E[E(g(X)|Y)]$$

- 例：
  - 设$(X, Y) \sim N(a, b, \sigma_1^2, \sigma_2^2, \rho)$，试计算$E(XY)$。

例子

- 一窃贼被关在有3个门的地牢里，其中第一个门通向自由，出这个门走3小时便可以回到地面。第二个门通向另一个地道，走5个小时将返回地牢。第三个门通向更长的地道，走7个小时也回到地牢。若窃贼每次选择3个门的可能性总是相同，求他为获得自由而奔走的平均时间。

大纲

1. 数学期望（均值）的定义
2. 数学期望（均值）的性质
3. 条件数学期望（条件均值）
4. 中位数

中位数

- 假设X为连续型随机变量，如果

$$P(X \le m) = \frac{1}{2}, P(X \ge m) = \frac{1}{2},$$

- 则称m为随机变量X的中位数。

- 假设X为离散型随机变量，如果

$$P(X \le m) \ge \frac{1}{2}, P(X \ge m) \ge \frac{1}{2},$$

- 则称m为随机变量X的中位数。

大纲

1. 方差和标准差
2. 矩
3. 协方差
4. 相关系数

大纲

1. 方差和标准差
2. 矩
3. 协方差
4. 相关系数

方差

- 设X为随机变量，分布为F，则称

$$Var(X) = E(X - EX)^2$$

- 为X（或分布F）的方差，其平方根$\sqrt{Var(X)} = \sigma$（取正值）成为X（或分布F）的标准差。

- 显然有

$$Var(X) = EX^2 - (EX)^2$$

性质

1. 常数的方差为0。
2. 若c为常数，则$Var(X+c)=Var(X)$。
3. 若c为常数，则$Var(cX)=c^2Var(X)$。
4. 如果随机变量X和Y相互独立，a和b为常数，则
$$Var(aX+bY)=a^2Var(X)+b^2Var(Y).$$

常见分布的方差

- 二项分布 $X \sim B(n, p)$
- Poisson分布 $X \sim P(\lambda)$
- 正态分布 $X \sim N(\mu, \sigma^2)$
- 均匀分布 $X \sim U(a, b)$
- 指数分布 $X \sim Exp(\lambda)$

标准化

- 我们称

$$Y = \frac{X - EX}{\sqrt{Var(X)}}$$

- 为随机变量X的标准化。

- $E(Y) = 0$
- $Var(Y) = 1$

矩

- 设X为随机变量，c为常数，k为正整数，则
$$E[(X-c)^k]$$
- 称为X关于c点的k阶矩。

- 两个重要的情况：
  - $c=0$。这时 $a_k=E\left(X^k\right)$ 称为X的k阶原点矩。
  - $c=E(X)$。这时 $\mu_k=E\left((X-EX)^k\right)$ 称为X的k阶中心矩。

育天下英才 創寰宇學府
嚴濟慈 一九八八年五月 題

偏度和峰度

- 偏度系数：

$$\beta_1 = \mu_3/\mu_2^{3/2}$$

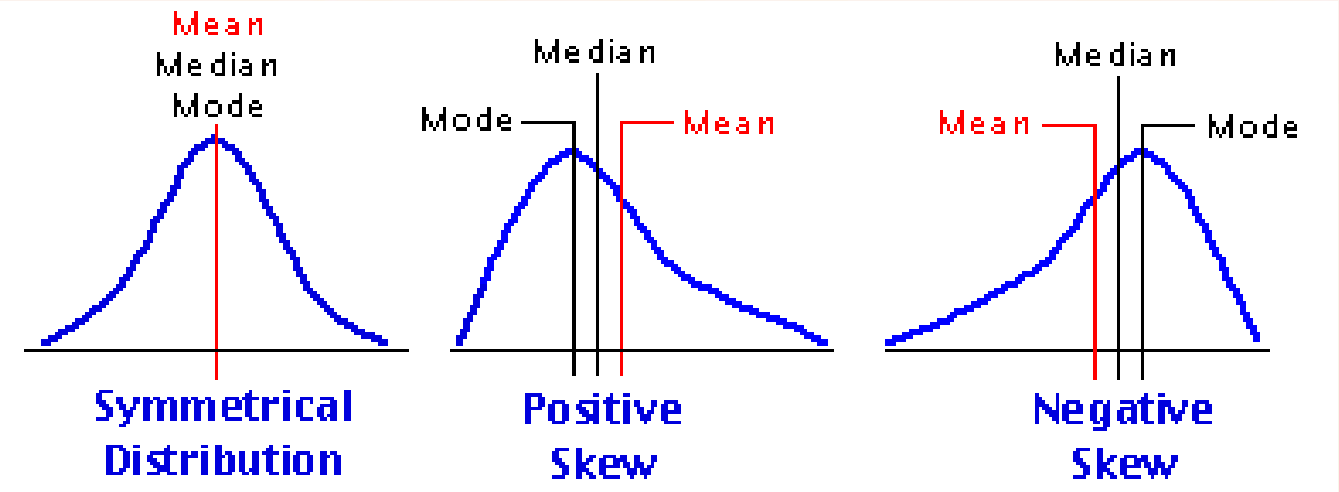
- 峰度系数：

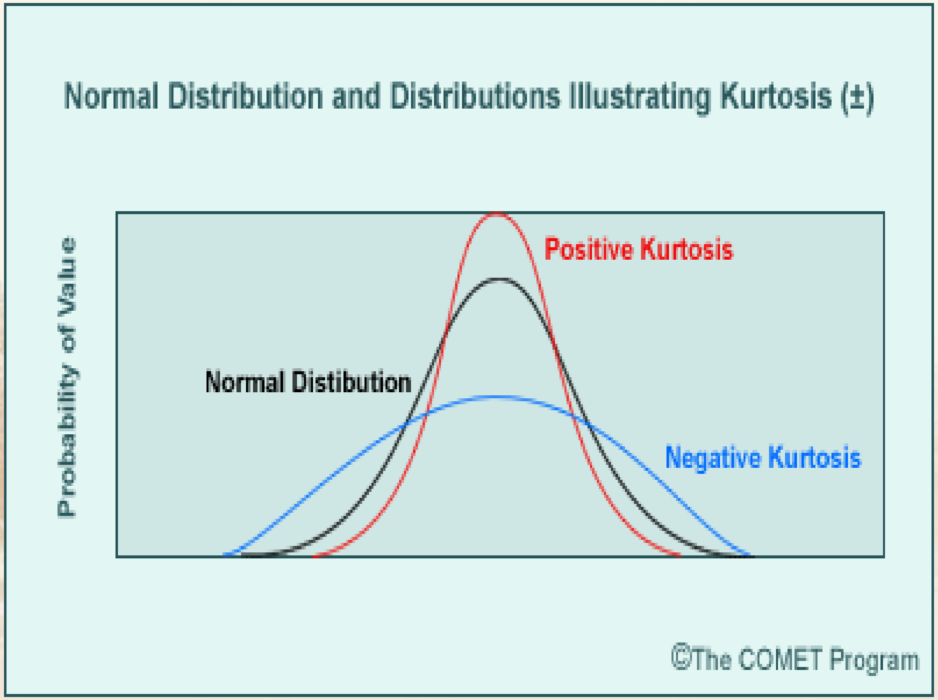
$$\beta_2 = \frac{\mu_4}{\mu_2^2} - 3$$

例：

正态分布的偏度和峰度系数？

创寰宇学府 育天下英才 严济慈 一九八八年五月 题


Mean
Median
Mode
Symmetrical Distribution
Median
Mode
Mean
Positive Skew
Median
Mean
Mode
Negative Skew


Normal Distribution and Distributions Illustrating Kurtosis (±)
Probability of Value
Positive Kurtosis
Normal Distibution
Negative Kurtosis
©The COMET Program

大纲

1. 方差和标准差
2. 矩
3. 协方差
4. 相关系数

协方差

- 设X与Y为随机变量，我们称

$$Cov(X,Y)=E(X-EX)(Y-EY)$$

- 为X与Y的协方差。

- 例：
  - 设$(X,Y)\sim N(a,b,\sigma_1^2,\sigma_2^2,\rho)$，设求$Cov(X,Y)$

性质

1. $Cov(X,Y) = Cov(Y,X), Cov(X,X) = Var(X)$。
2. $Cov(X,Y) = EXY - EXEY$。
3. 若X与Y相互独立，则$Cov(X,Y) = 0$。
4. $Cov(X_1 + X_2, Y) = Cov(X_1, Y) + Cov(X_2, Y)$。
5. $Cov(aX, bY) = abCov(X,Y)$。
6. $Var(X + \mathrm{Y}) = Var(X) + 2cov(X,Y) + Var(Y)$

大纲

1. 方差和标准差
2. 矩
3. 协方差
4. 相关系数

相关系数

- 设X与Y为随机变量，我们称

$$Corr(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

- 为X与Y的相关系数。当$Corr(X,Y)=0$，则称X与Y不相关。
- 标准尺度下的协方差

- 例：
  - 设$X\sim U\left(-\frac{1}{2},\frac{1}{2}\right)$, $\mathrm{Y}=\cos(X)$，求$Corr(X,Y)$。

性质

1. 若X与Y相互独立，则$Corr(X,Y)=0$。
2. $|Corr(X,Y)|\le 1$，等号成立当且仅当X与Y之间存在严格的线性关系，即存在实数a和b使得Y $= aX+b$。

$Corr(X,Y)$只能刻画X与Y之间的线性相依程度。

- $|Corr(X,Y)|$越接近1，就表示X与Y之间的线性相关程度越高。
- $Corr(X,Y)=0$只能表明X与Y之间不存在线性关系，但可以存在非线性的函数关系。

创寰宇学府 育天下英才
严济慈 一九八八年五月题

独立性和不相关性

- 对于随机变量X与Y：
  - 如果X与Y相互独立，那么它们一定不相关；
  - 但如果X与Y不相关，它们未必相互独立。

- 例：
  - 若(X,Y)服从单位圆内的均匀分布，则X与Y不相关但不独立。
  - 若$(X,Y)\sim N(a,b,\sigma_1^2,\sigma_2^2,\rho)$，则X与Y独立等价于$X$与$Y$不相关。

大纲

1. 大数定律
2. 中心极限定理

大纲

1. 大数定律
2. 中心极限定理

大数定律

- 设$\{X_n\}$是一列独立同分布的随机变量序列，具有公共的数学期望μ和方差σ^2，则对任意给定的$\epsilon > 0$，有
$$\lim_{n\to\infty} P(|\bar{X}_n - \mu| \geq \epsilon) = 0,$$
- 其中$\bar{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$。我们称$\bar{X}_n$服从（弱）大数定律。

- $\bar{X}_n$依概率收敛于μ，记为$\bar{X}_n \xrightarrow{p} \mu$.

证明

- 马尔可夫不等式:
  - 若Y为只取非负数的随机变量，则对任意的常数$\epsilon > 0$，都有
$$P(Y \geq \epsilon) \leq \frac{E(Y)}{\epsilon}.$$
- 切比雪夫不等式:
  - 若随机变量X的方差存在，则对任意的常数$\epsilon > 0$，都有
$$P(|X - EX| \geq \epsilon) \leq \frac{Var(X)}{\epsilon^2}.$$

例子：伯努力大数定律

- 若以Y_n表示n重伯努力实验中的成功次数，则对任意的常数$\epsilon > 0$，都有

$$\lim_{n\to\infty} P(|p_n - p| \geq \epsilon) = 0$$

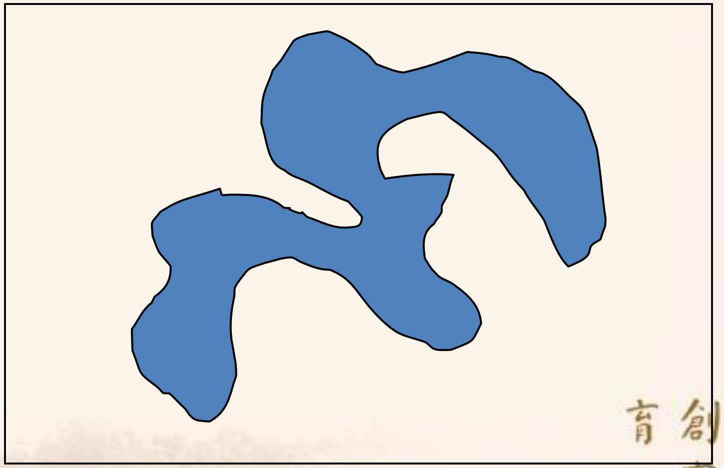
其中$\mathrm{p_n} = \frac{Y_n}{n}$。

- 频率收敛于概率

例3：求不规则区域的面积

- 随机地把 N 个点投入方形区域（面积=1），落入不规则区域的个数为 n，则不规则区域面积 S 可以用比率 n/N 逼近 （N非常大）

$$n/N \to S$$

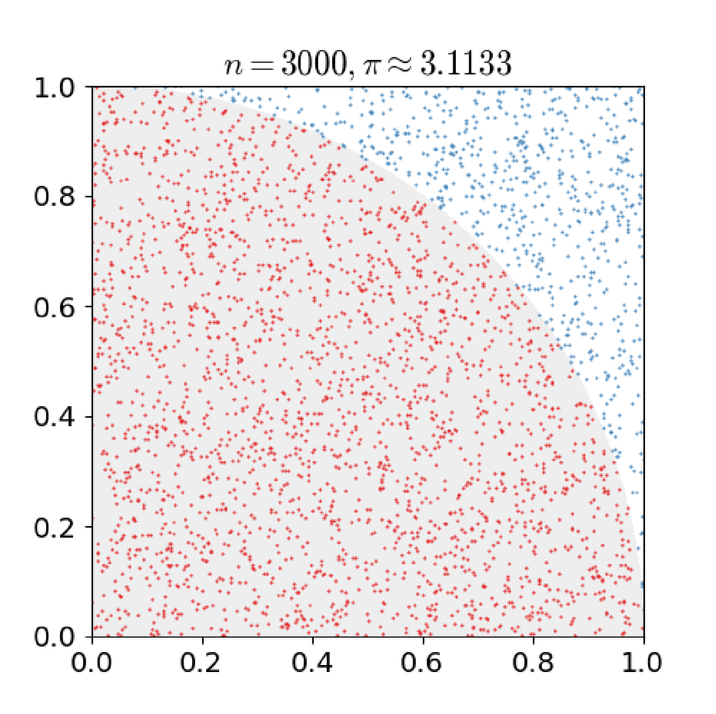


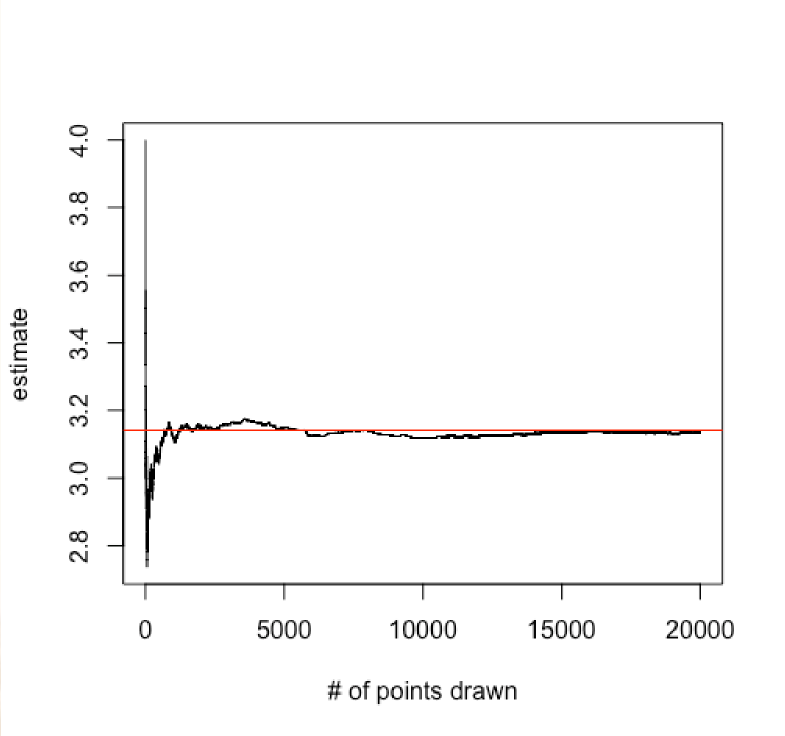
大数律

- 应用： 用蒙特卡洛方法估计圆周率

应用：估计圆周率

- 用蒙特卡洛方法估计圆周率





https://en.wikipedia.org/wiki/Monte_Carlo_method

大纲

1. 大数定律
2. 中心极限定理

中心极限定理

- 设$\{X_n\}$是一列独立同分布的随机变量序列，具有公共的的数学期望μ和方差σ^2，则$X_1+\cdots+X_n$的标准化形式$\frac{1}{\sqrt{n}\sigma}(X_1+\cdots+X_n-n\mu)$满足中心极限定理。即对任意实数$x$，有

$$\lim_{n\to\infty} P\left(\frac{1}{\sqrt{n}\sigma}(X_1+\cdots+X_n-n\mu)\le x\right)=\Phi(x),$$

其中$\Phi(x)$为标准正态分布$N(0,1)$的分布函数。记为

$$\frac{1}{\sqrt{n}\sigma}(X_1+\cdots+X_n-n\mu)\xrightarrow{d} N(0,1).$$

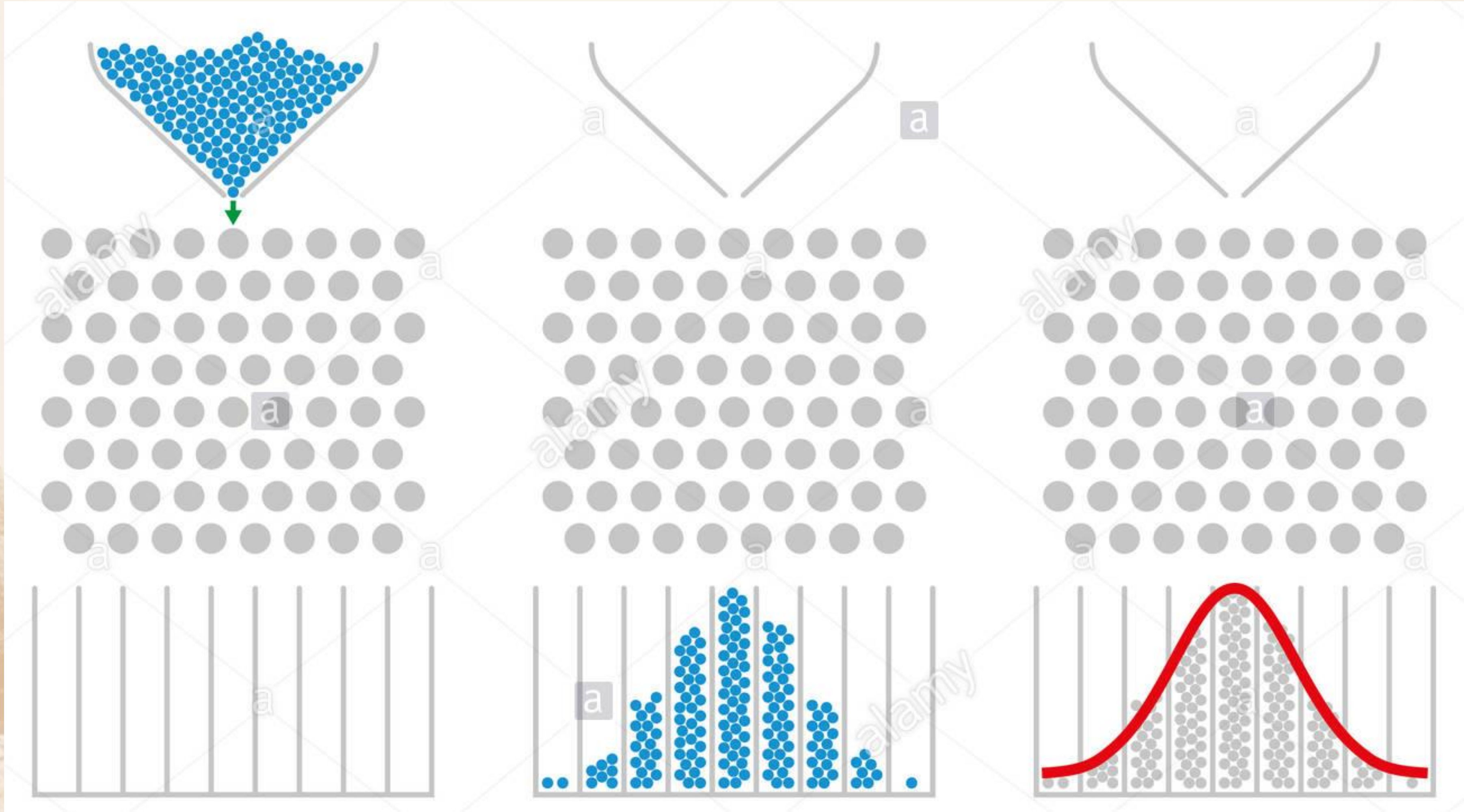
棣莫弗-拉普拉斯定理

- 设$X_1, \dots, X_n$相互独立且具有相同的分布
  $P(X_i = 1) = p, P(X_i = 0) = 1 - p, 0 < p < 1$
- 则对于任意实数x，有

$$\lim_{n\to\infty} P\left(\frac{1}{\sqrt{np(1-p)}}(X_1 + \cdots + X_n - np)\right.$$

中国科学技术大学
University of Science and Technology of China

高尔顿板



https://en.wikipedia.org/wiki/Bean_machine

例子

- 设一考生参加100道题的英语标准化考试（每道题均有连个备选答案的选择题，有且只有一个答案是正确的），每道题他都随机地选择一个答案，假设评分标准为：选对得一份，选错或者不选不得分。是给出该考生最终得分大于等于50的概率。

Thank You!

創寰宇學府
育天下英才
嚴濟慈
一九八八年五月題