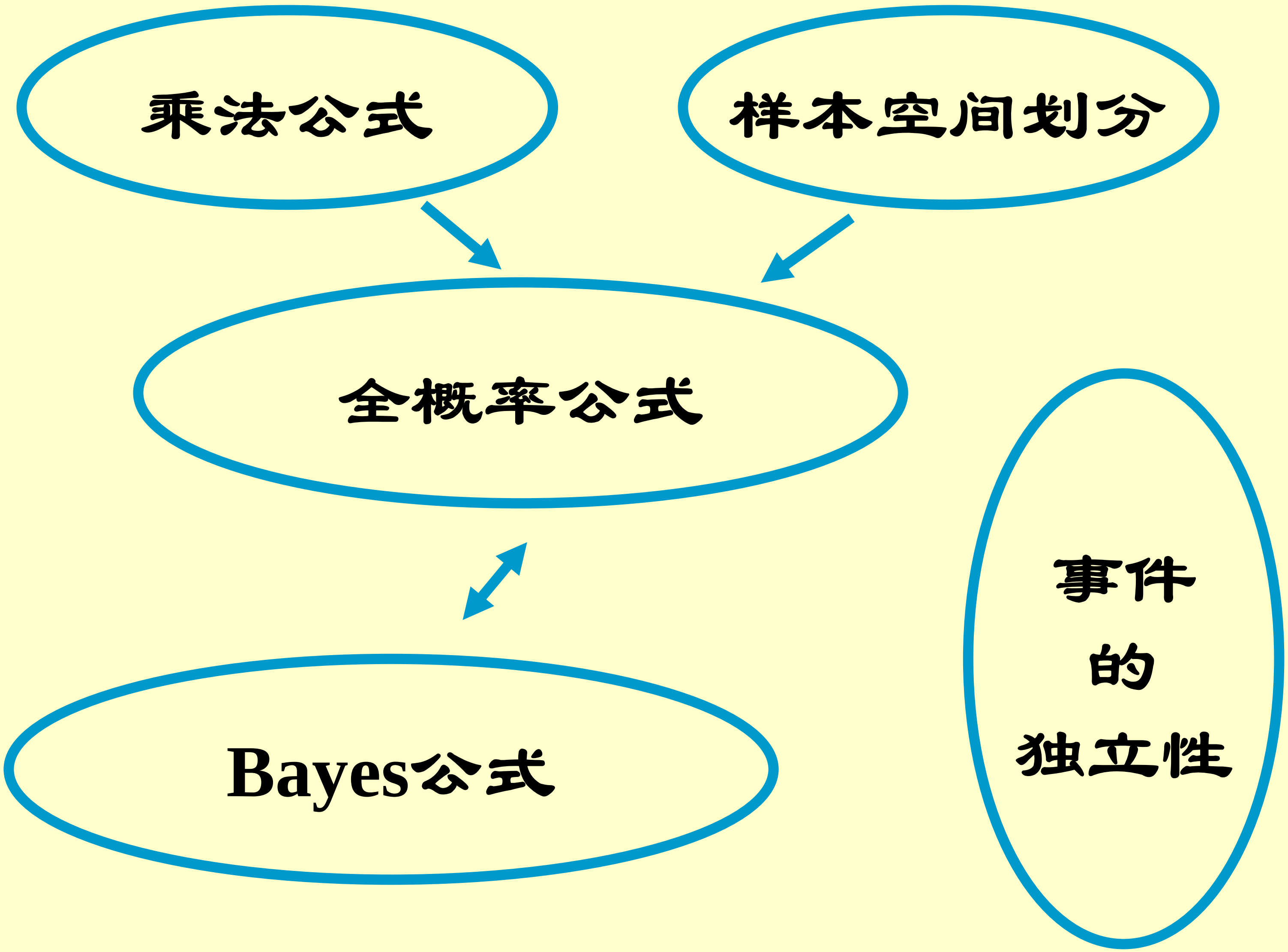
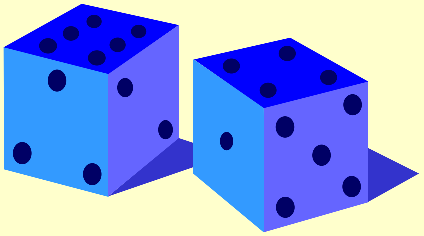

乘法公式
样本空间划分
全概率公式
Bayes公式
事件
的
独立性

- 第一章 小结
- 六个概念（随机试验、事件、概率、条件概率、独立性）
- 四个公式（加法公式、乘法公式、全概率公式、贝叶斯公式）
- 二个概型（古典概型、 几何概型）

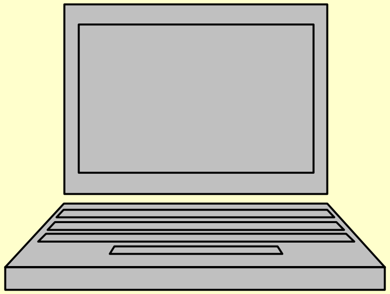
第二章
随机变量及其分布

- 抽样数据的描述统计
- 随机变量及其概率分布
- 随机变量的数学期望
- 随机变量的方差
- 常用随机变量的分布
- 应用案例



2.1 抽样数据的描述统计

建立频率分布表，作频率直方图



具体步骤：

确定分组数k；

确定组距d；

确定组限，即各小区间的上、下限；

计算样本落入各小区间的频数、频率；

得频率分布表，作出频率直方图。

上述过程可通过Excel软件来实现

Excel软件的使用

加载“分析工具库”：

工具 / 加载宏 / 分析工具库

频率计算与直方图显示：

工具 / 数据分析 / 直方图

计算样本统计量：

工具 / 数据分析 / 描述统计

样本数据的中心描述

样本均值：$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

样 本 中 位 数 Me ： 将 原 始 数 据 按 大 小 排 序， 记 为：

$x_{(1)} \le x_{(2)} \le \cdots x_{(n)}$. 处于中间位置的数称为中位数。

$$Me = \begin{cases} x_{(\frac{n+1}{2})}, & n\text{为奇数} \\ \frac{1}{2}[x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}], & n\text{为偶数。} \end{cases}$$

众 数 Mod ： 它 是 一 组 数 据 中 出 现 机 会 最 大 的 数 。

样本数据的离散程度描述

样本方差 S_{n-1}^2：

$$S_{n-1}^2 = \sigma_{n-1}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n-1}[\sum_{i=1}^{n} x_i^2 - n\bar{x}^2].$$

标准偏差 σ_{n-1}：

$$\sigma_{n-1} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

注：有的书上用

$$S_n^2 = \sigma_n^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2, \sigma_n = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}.$$

极差 R： $R = x_{\max} - x_{\min} = x_{(n)} - x_{(1)}$；

2.2 随机变量及其概率分布

随机变量的定义：

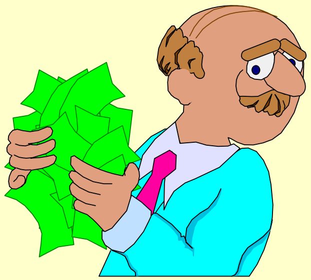
在$(\Omega, \mathcal{F}, P)$概率空间框架下，如果对一切$x \in R$，有事件$\mathbf{A} = \{\xi(\omega) \le x\} \in \mathcal{F}$，则实值函数$\xi = \xi(\omega)$（$\omega \in \Omega$）称为随机度量，有时为强调$\xi$与$\mathcal{F}$的关系，$\xi$又称为$\mathcal{F}$可测的随机变量（简记为$r.v.$），记为$\xi \in \mathcal{F}$。

随机变量与普通函数的区别

1） 随机变量随着试验结果而取不同的值，在试验前只知道它可能取值的范围，而不能预知它取什么值。

2) 随机变量取各个值有一定的概率。

3) 普通函数定义在实数轴上（一元），而随机变量是定义在样本空间上的，样本空间的元素不一定是实数。



分布函数的定义：

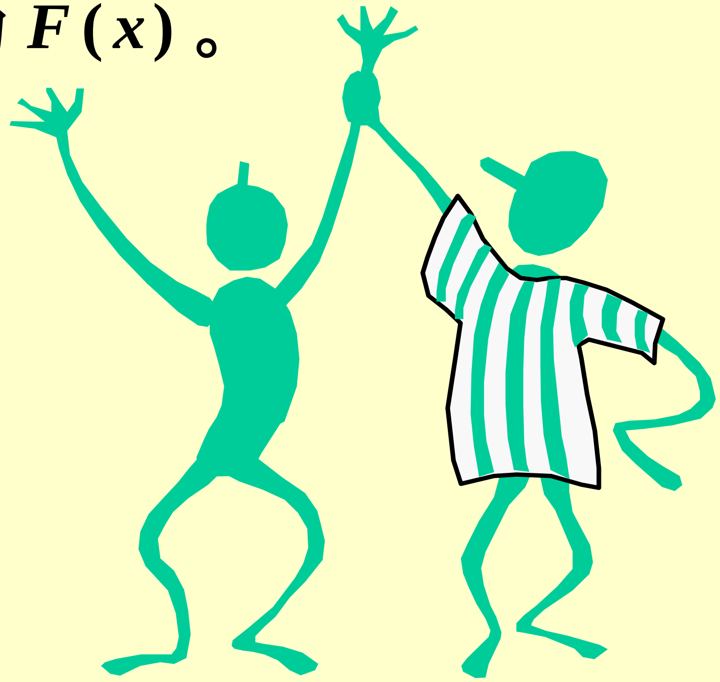
设ξ是定义在$(\Omega, \mathcal{F}, P)$上的随机度量，对于任意的实数$x \in R$，

$$F_{\xi}(x) = P\{\omega \mid \xi(\omega) \le x\} = P\{\xi \le x\},$$

称为ξ的分布函数，记为$F(x)$。

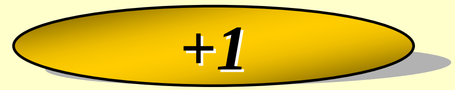
注意：

任意一个随机变量都存在唯一的分布函数，而同样的分布函数，又是可以对应不同的随机变量.



例 1. 掷一枚硬币，把出现正面与反面分别记为 ＋1 与－1，这样可以把掷硬币的结果用随机变量ξ来表示，显然对均匀硬币 $P(\xi=\pm 1)=\dfrac{1}{2}$，

（1）试求ξ的分布函数$F_1(x)$； (2) 若令$\eta=-\xi$，再求η的分布函数$F_2(x)$。



解：

$$F_1(x)=\begin{cases} 0, & x\in(-\infty,-1) \\ \dfrac{1}{2}, & x\in[-1,1) \\ 1, & x\in[+1,+\infty) \end{cases}$$

注意到 $P(\eta=\mp 1)=P(\xi=\pm 1)=\dfrac{1}{2}$

$$\therefore F_\eta(x)=F_2(x)=F_1(x)。$$

性质

（1）非降性：$F(x)$是单调非降函数，即对任意的

$x_1 < x_2$，有 $F(x_1) \le F(x_2)$；

(2) 0－1 性：对任意 $x \in R$,有 $0 \le F(x) \le 1$，且

$$F(-\infty) = \lim_{x \leftarrow \infty} F(x) = \mathbf{0}, \quad F(+\infty) = \lim_{x \to \infty} F(x) = \mathbf{1};$$

(3) 右连左极性：$F(x)$是 X 的右连左极函数，

即对任意的 $x_0 \in R$,有

$$\lim_{x \to x_0^+} F(x) = F(x_0)，即 F(x_0 + 0) = F(x_0)$$

$$\lim_{x \to x_0^-} F(x) = F(x_0 -) \le F(x_0)。$$

利用分布函数计算概率：

对任意的实数$a<b$，必有

$$P(a<\xi\le b)=F(b)-F(a)$$

$$P(\xi>b)=F(+\infty)-F(b)=1-F(b).$$

$$P(\xi<a)=\lim_{n\to\infty}P(\xi\le a-\frac{1}{n})$$

$$=\lim_{n\to\infty}F(a-\frac{1}{n})=F(a-0)$$

$$P(\xi=a)=P(\xi\le a)-P(\xi<a)$$

$$=F(a)-F(a-0)$$

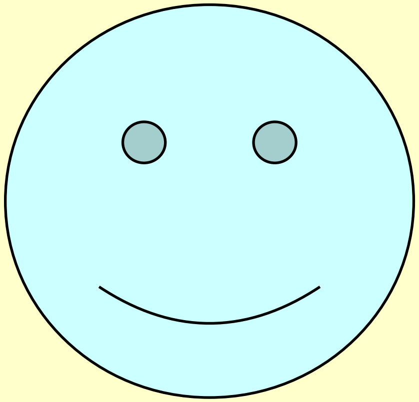
$$P(a\le\xi<b)=F(b-0)-F(a-0)$$

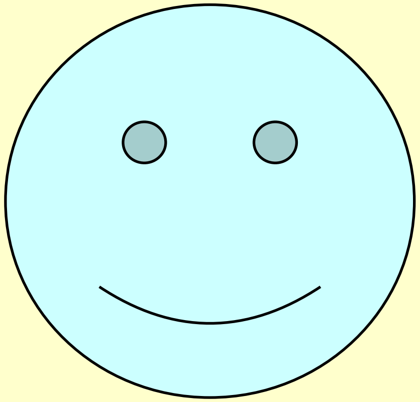
★随机变量的优点：

可以用数学分析（微积分）的方法来研究随机试验。

★随机变量的分类：（按随机变量可能取值范围）

随机变量
- 离散型随机变量
- 非离散型
  - 连续型
  - 奇异型（混合型）





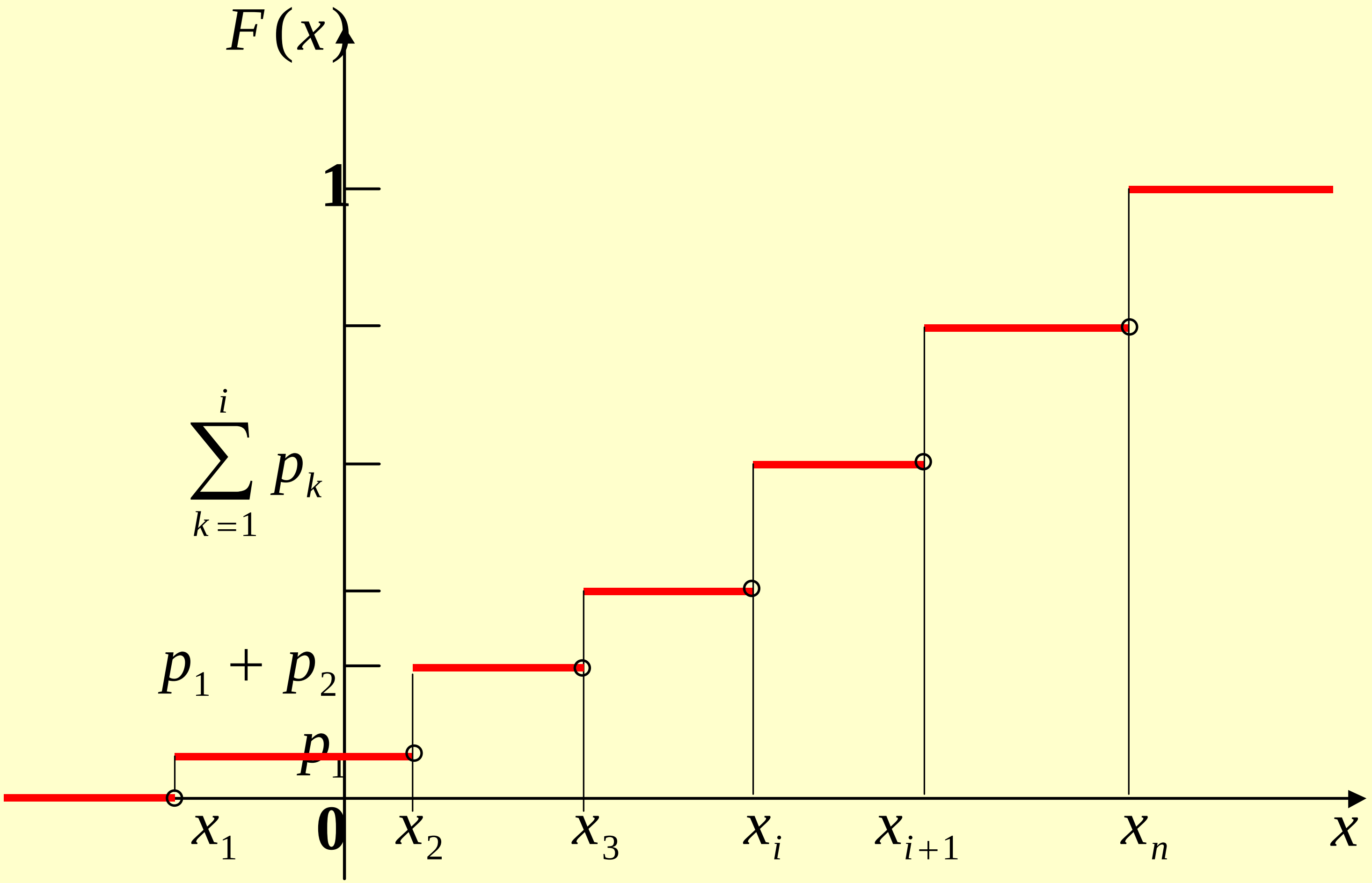
2.3 离散型随机变量及其分布列

1. 离散型随机变量 X 的概率分布

定义：设 R 上的函数 $F(x)$ 满足分布函数三个充要条件，而且是阶梯型函数，则对应的随机变量称为离散型随机变量。

复习

图形特点：右连续，台阶形



另一个定义：

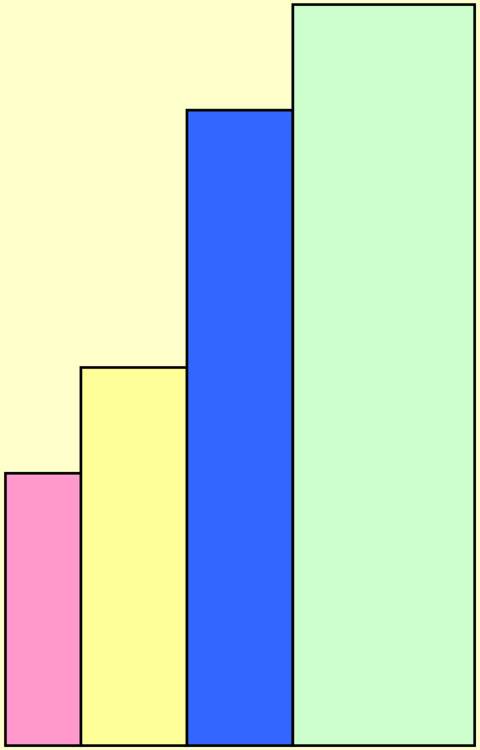
随机变量 X 对应于每一取值的概率为：

$$P(X=x_k)=p_k\,,\quad k=1,2,\cdots$$

称为离散型随机变量 X 的概率分布或分布律

分布律的表格形式：

X	x_1	x_2	$\cdots$	x_n	$\cdots$
p_k	p_1	p_2	$\cdots$	p_n	$\cdots$



性质：

①非负性： $p_k\geq 0,\quad k=1,2,\cdots.$

②规范性： $\sum_k p_k=1.$

★对离散型随机变量，若已知分布律，就可求出它的分布函数。

$$F(x) = P\{X \le x\} = \sum_{x_i \le x} P\{X = x_i\}$$

例如：$P\{X = x_i\} = p_i, i = 1,2,\cdots,n$

$$F(x) = \begin{cases} 0, & x < x_1 \\ p_1, & x_1 \le x < x_2 \\ p_1 + p_2, & x_2 \le x < x_3 \\ p_1 + p_2 + p_3, & x_3 \le x < x_4 \\ \vdots & \vdots \\ p_1 + p_2 + \cdots + p_n = 1, & x \ge x_n \end{cases}$$

例 **1.**随机变量 X 的分布律为

X	**-1**	**2**	**3**
p_k	**0.25**	**0.5**	**0.25**

求 X 的分布函数，并求 $P\{X \le \frac{1}{2}\}$， $P\{\frac{3}{2} < X \le \frac{5}{2}\}$，

$P\{2 \le X \le 3\}$。

解： X 的分布函数为

$$F(x) = \begin{cases} 0, & x < -1 \\ 0.25, & -1 \le x < 2 \\ 0.75, & 2 \le x < 3 \\ 1, & x \ge 3 \end{cases}$$

$$P\{X \le \frac{1}{2}\} = P\{X = -1\} = 0.25,$$

$$P\{\frac{3}{2} < X \le \frac{5}{2}\} = P\{X = 2\} = 0.5,$$

$$P\{2 \le X \le 3\} = P\{X = 2\} + P\{X = 3\}$$

$$= 0.5 + 0.25 = 0.75$$

解二：利用$F(x)$求。

$$P\{X \le \frac{1}{2}\} = F(\frac{1}{2}) = 0.25,$$

$$P\{\frac{3}{2} < X \le \frac{5}{2}\} = F(\frac{5}{2}) - F(\frac{3}{2}) = 0.75 - 0.25 = 0.5$$

$$P\{2 \le X \le 3\} = P\{2 < X \le 3\} + P\{X = 2\}$$

$$= F(3) - F(2) + 0.25 = 1 - 0.75 + 0.5 = 0.75$$

例 2. ξ是一个离散型随机变量，其分布列如下表所示

ξ	-1	0	1
P	0.5	$1-2q$	q^2

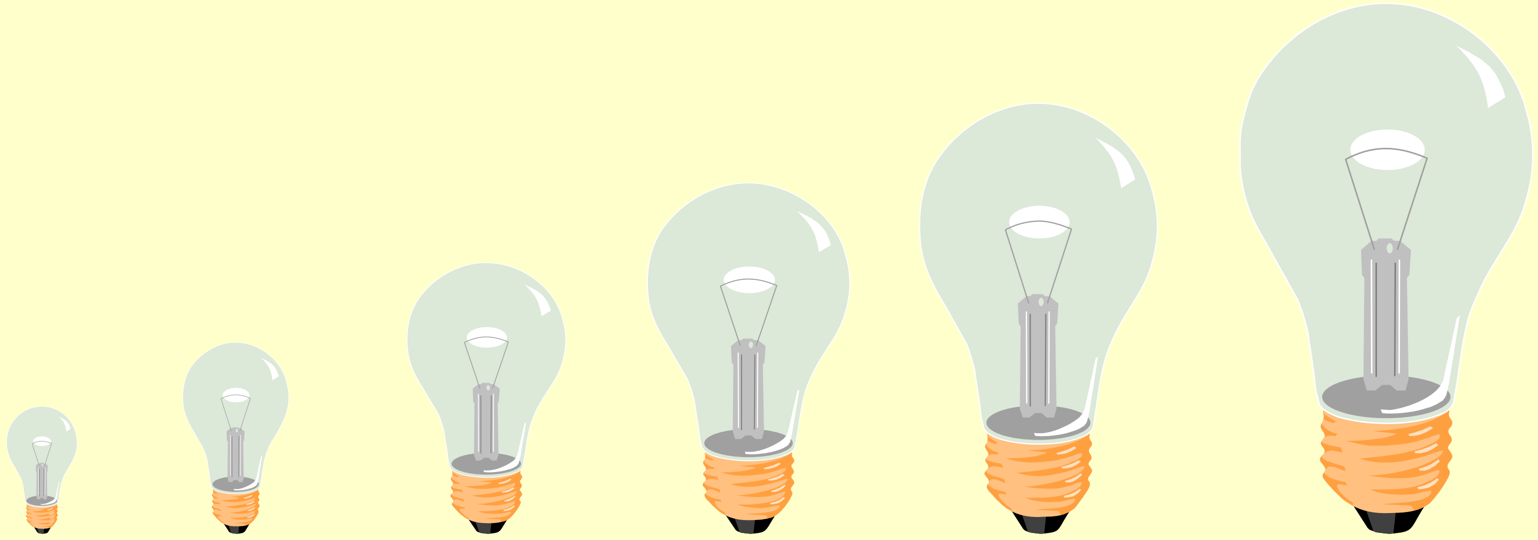
试求（1）常数q；（2）ξ的分布函数。

解：

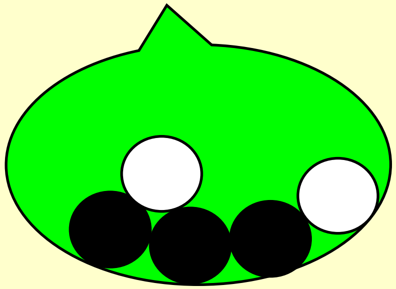
$$\begin{cases} \frac{1}{2}+1-2q+q^2=1 \\ 0\le 1-2q\le 1 \\ q^2\le 1 \end{cases}$$

解得 $q=1-\frac{\sqrt{2}}{2}$，于是有分布函数

$$F(x)=\begin{cases} 0, & x\in(-\infty,-1) \\ 0.5, & x\in[-1,0) \\ \sqrt{2}-0.5, & x\in[0,1) \\ 1, & x\in[1,+\infty) \end{cases}$$



例 3.袋中有 2 个白球，3 个黑球，每次从中任取一球，直到取到白球为止，求取球次数的概率分布。(分放回和不放回两种情形讨论.)



解：(1)放回

设 X 为取到白球时的取球次数,

则 $X=1,2,\cdots$

$$P(X=1)=\frac{2}{5},\ P(X=2)=\frac{2}{5}\cdot\frac{3}{5},$$

$$P(X=3)=\frac{2}{5}\left(\frac{3}{5}\right)^2,\cdots$$

几何级数

$$P(X=m)=\frac{2}{5}\left(\frac{3}{5}\right)^{m-1},\cdots,\quad m=1,2,\cdots$$

这种分布称为几何分布。

显然，$$\sum_{m=1}^{\infty}P(X=m)=\sum_{m=1}^{\infty}\frac{2}{5}\left(\frac{3}{5}\right)^{m-1}=\frac{2/5}{1-3/5}=1$$

(2)不放回

设 Y 为取到白球时的取球次数，$Y=1,2,3,4$，则

$$P(Y=1)=\frac{2}{5}=0.4,\qquad P(Y=2)=\frac{3}{5}\cdot\frac{2}{4}=0.3,$$

$$P(Y=3)=\frac{3}{5}\cdot\frac{2}{4}\cdot\frac{2}{3}=0.2,$$

$$P(Y=4)=\frac{3}{5}\cdot\frac{2}{4}\cdot\frac{1}{3}\cdot\frac{2}{2}=0.1,$$

Y	1	2	3	4
P	0.4	0.3	0.2	0.1

显然，$\sum_{m=1}^{4} P(Y=m)=0.4+0.3+0.2+0.1=1$