意义：

平均数是一组资料的代表值，表示资料中观察值的中心位置与集中趋势。

种类：

主要有 算术平均数（Average）、

中位数（Median）、

众 数（Mode）

几何平均数（Geometric mean）.

算术平均数($\bar{\chi}$) 的性质

（1） $\bar{\chi}$ 的大小受样本内每个值的影响

（2）若每个X_i都乘以相同的数K，则 $\bar{\chi}$ 也应乘以K

（3）若每个Xi都加上相同的数A，则 $\bar{\chi}$ 也应加上A。

（4）离均差之和等于零。

（5）离均差之平方和最小。

本次内容

- 标准差
- 变异系数
- 第二章 概率分布

度量数据的离散程度：

极差

平均离差

标准差

方差

变异系数

常用的几种标志变异指标

	概　念	计　算	特　点
1．极差（R）	数列中最大值与最小值之差	R=最大值-最小值	优点：容易理解，计算方便 缺点：不能反映全部数据分布状况
2．平均离差（A、D）	各观察值与均值离差绝对值的算术平均	简单：$A\cdot D=\frac{\sum\lvert X_i-\bar{X}\rvert}{n}$ 加权：$A\cdot D=\frac{\sum\lvert X_i-\bar{X}\rvert f_i}{\sum f_i}$	优点：反映全部数据分布状况 缺点：取绝对值，数字上不尽合理

	概　念	计　算	特　点
3.方差（S^2, σ²）和 标准差（S, σ）	各标志值与均值离差平方的平均。 方差的平方根（取正根）	简单：$\sigma^2=\frac{\sum(X_i-\bar{X})^2}{n}$ 加权：$\sigma^2=\frac{\sum(X_i-\bar{X})^2 f_i}{\sum f_i}$ $\sigma=\sqrt{\sigma^2}$	优点：反映全部数据分布状况. 缺点：受计量单位和平均水平影响，不便于比较
4. 变异系数（CV）	标准差与均值之商，可以用来比较不同资料的相对变异程度	$CV=\frac{S}{\bar{X}}$	优点：适宜不同数据的比较,是无量纲的系数 缺点：对数据结构变化反应不灵敏

方差（σ²）和标准差（σ）是应用最广的标志变异指标

标准差 (Standard deviation)

- 意义:表示每个观察值与平均值的平均变异程度，反映观察值分布的离散趋势。
- 公式

（一）直接法

$$S=\sqrt{\frac{\sum\left(x_i-\bar{x}\right)^2}{n-1}}$$

$$s=\sqrt{\frac{\sum x^2-(\sum x)^2/n}{n-1}}$$

（二）频数资料的计算

$$s=\sqrt{\frac{\sum f_i(x_i-\bar{x})^2}{\sum f_i-1}}=\sqrt{\frac{\sum f_i x_i^2-(\sum f_i x_i)^2/\sum f_i}{\sum f_i-1}}$$

例　计算以下数据的标准差：**26 25 28 24 23 25 27 27 30 21**。

解 最好列成以下表格的形式计算

	x	x^2
	26	676
	25	625
	28	784
	24	576
	23	529
	25	625
	27	729
	27	729
	30	900
	21	441
$\sum x$	256	6614

$$s=\sqrt{\frac{\sum x_i^{\,2}-\frac{(\sum x_i)^2}{n}}{n-1}}$$

$$=2.59$$

某纯系蛋鸡200枚蛋重资料次数分布及标准差计算

组别	组中值（x）	次数（f）	fx	fx^2
44.15—	45.0	3	135.0	6075.0
45.85—	46.7	6	280.2	13085.34
47.55—	48.4	16	774.4	37480.96
49.25—	50.1	22	1102.2	55220.22
50.95—	51.8	30	1554.0	80497.20
52.65—	53.5	44	2354.0	125939.00
54.35—	55.2	28	1545.0	85317.12
56.05—	56.9	30	1707.0	97128.30
57.75—	58.6	12	703.2	41207.52
59.45—	60.3	5	301.5	18180.45
61.15—	62.0	4	248.0	15376.00
合计		$\Sigma f=200$	107051	575507.11

利用某纯系蛋鸡200枚蛋重资料的次数分布表（见表）计算标准差。

将表中的Σ*f*、Σ*fx*、$\sum fx^2$ 代入上式得：

$$S=\sqrt{\frac{\sum fx^2-(\sum fx)^2/\sum f}{\sum f-1}}=\sqrt{\frac{57550.7\ 11-107051^2/200}{200-1}}=3.55(g)$$

即某纯系蛋鸡200枚蛋重的标准差为3.55*g*。

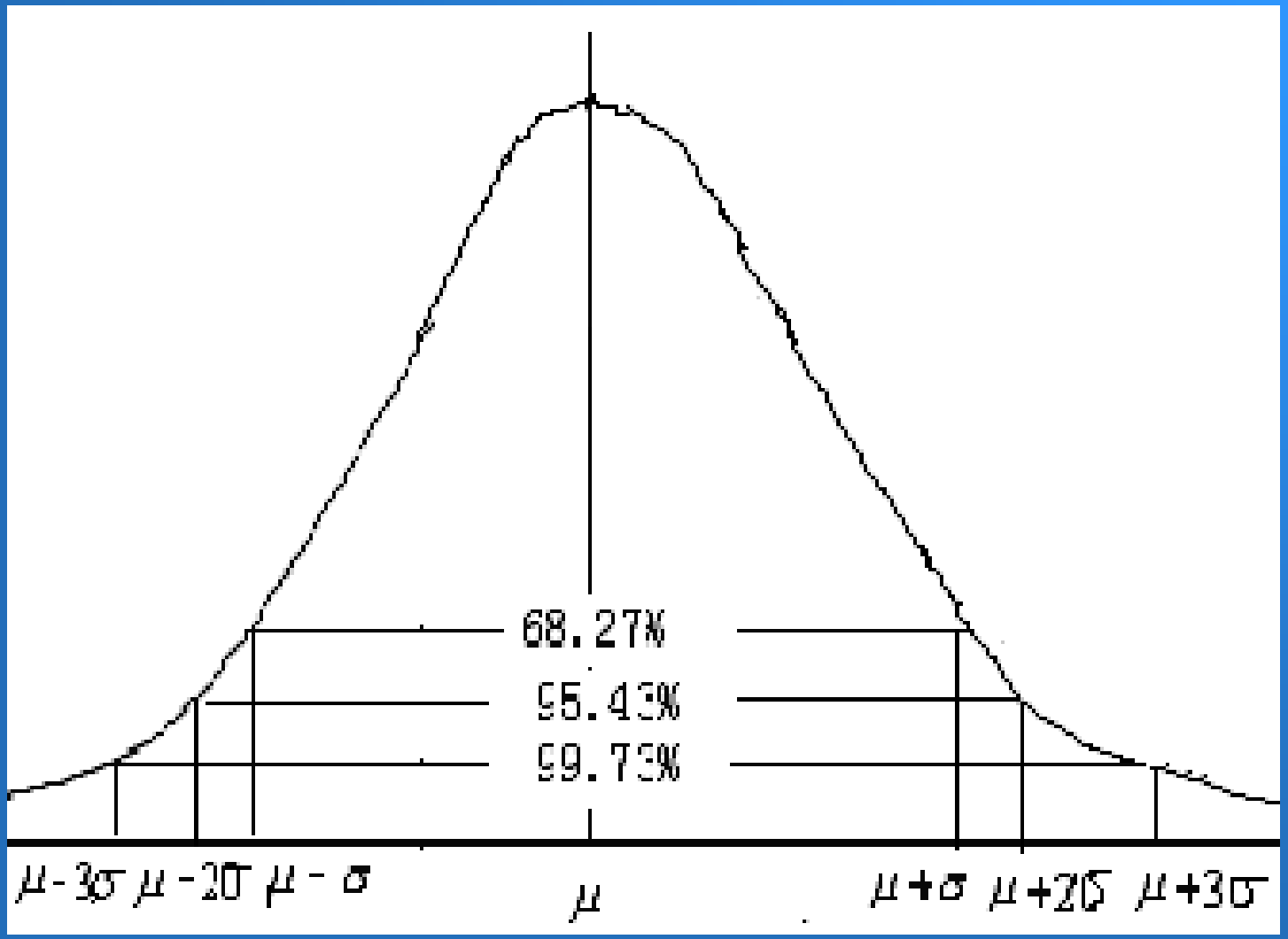
标准差的性质：

（1）标准差大小受观察值影响，其单位与观察值相同。

（2）在计算时，对各观测值加上或减去一个常数，其数值不变。

（3）当每个观测值乘以或除以一个常数a，则所得的标准差是原来标准差的a倍或1/a倍。

（4）在资料服从正态分布的条件下，约有68.26%的观测值在平均数左右一倍标准差（$\bar{X}$ ±S）范围内；约有95.43%在平均数左右两倍标准差（$\bar{X}$ ±2S）范围内；约有99.73%在平均数左右三倍标准差（$\bar{X}$ ±3S） 范 围内。


68.27%
95.43%
99.73%
μ-3σ
μ-2σ
μ-σ
μ
μ+σ
μ+2σ
μ+3σ

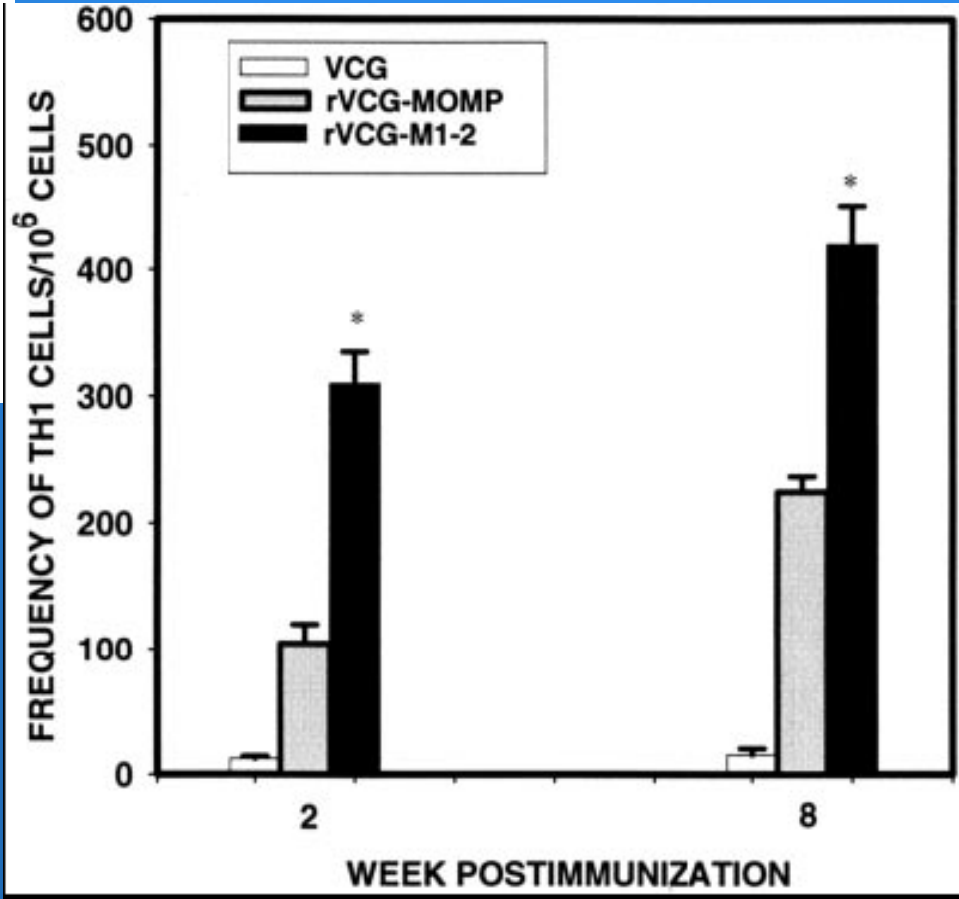
实验数据的科学表示法：

$\bar{x} \pm s$ 或 $\bar{x} \pm SD$

Table 1
Fresh ovarian weight and preovulatory follicles from infected and uninfected rabbits

Groups	Left ovary (g)	Right ovary (g)	Preovulatory follicles
Controls	0.32 ± 0.04	0.33 ± 0.04	6.5 ± 1.7
MV-ZPB Day 5	0.22 ± 0.05*	0.25 ± 0.06	5.0 ± 2.2
MV-ZPB Day 15	0.15 ± 0.05**	0.13 ± 0.06**	4.3 ± 1.9
MV-HA Day 15	0.16 ± 0.07**	0.16 ± 0.06**	4.3 ± 1.3
MV-ZPB Day 30	0.19 ± 0.05**	0.19 ± 0.07*	2.3 ± 1.5†
MV-HA Day 30	0.13 ± 0.03**	0.14 ± 0.02**	5.8 ± 1.3

Notes: the values are mean ± SD for four rabbits in each group.
* $P < 0.05$ against uninfected controls.
** $P < 0.01$ against uninfected controls.
† $P < 0.05$ against infected (MV-HA Day 30) and uninfected controls.



5．变异系数（Coefficient of variance，CV）

标准差可以反映数据的离散程度，

但对于**平均水平不同或计量单位不同的不同样本**，是不能用标准差直接比较离散程度。

$$CV = \frac{S}{\bar{X}}$$

注意，变异系数的大小，同时受平均数和标准差两个统计量的影响，因而在利用变异系数表示资料的变异程度时，最好将平均数和标准差也列出。

- 两组样本的 $\bar{X}$ 接近时可用S比较它们的变异性大小，极差较大时不可。例：10岁身高 A=120±5.0cm

  8个月身高 B=70±4.0cm

$$cv = \frac{s}{\bar{x}} \qquad cv_1 = \frac{s}{\bar{x}} = \frac{5}{120} = 0.042 \qquad cv_2 = \frac{4}{70} = 0.057$$

- 当两个样本的观察值单位不同时，也不能直接比较。

  即两种性状，如身高、体重

  例：大学男生身高：165±12cm CV_1=0.073

  体重：58.3±4.8kg CV_2=0.082

恭喜大家！

- 你们现在已基本入门。下一章的难度要大些，

一定要加油喔！

在前面的章节中，我们介绍了搜集、整理和描述统计数据的一些基本方法，通过对统计数据的整理和描述可以使我们对客观事物的概貌有一个了解。统计数据中蕴含着非常丰富的重要信息，要有效的充分利用统计数据，需要运用统计推断的方法。

统计推断就是在搜集、整理观测样本数据的基础上，对有关总体做出推断，其特点是根据随机性的观测样本数据以及问题的条件和假定，对未知事物作出的以概率形式表述的推断。

第二章 概率分布

本章介绍**随机变量**和**概率分布**两个重要概念

第一节 随机事件

2.1.1 事 件(random event)

- 必然现象与随机现象

必然现象： 指在某些条件下，一定会发生的现象。

必然事件: 必然现象的每一个结果称为必然事件。

随机现象：

在相同的条件下重复进行试验，具有多种可能发生的结果，每次究竟发生哪一种结果事先不能肯定，这类在个别试验中其结果呈现偶然性、不确定性现象，称为**随机现象**或**不确定性现象**。

随机事件：

随机现象的每一个可能结果称为**随机事件**，**简称事件**，通常用字母A、B、C……等表示。

如，新生儿的性别

随机试验 (random trial)

随机试验： 对随机现象的每一次测定称为**随机试验**。

随机试验的三个特征

（1）试验可以在相同条件下多次重复进行；

（2）每次试验的可能结果不止一个 ，并且事先知道会有哪些可能的结果；

（3）每次试验总是恰好出现这些可能结果中的一个 ，但在一次试验之前却不能肯定这次试验会出现哪一个结果。

2.1.2. 统计规律——频率的稳定性

- 在相同条件下进行大量重复试验时，其试验结果呈现出某种固有的特定的规律性—**频率的稳定性，通常称之为随机现象的统计规律性。**

统计规律性表现为单独一次结果的不确定性和累积结果的有规律性。 如天气预报、虫害发生预报。

频率：对于事件A，若在K次试验中出现了L次，则称K/L 为随机事件A在K次试验中出现的频率。

例如抛硬币试验，可能结果：
事件A:国徽面，事件B:币值面

次数	币值面次数	频率
1	0	0
	1	1
3	0	0
	1	0.33
	2	0.67
	3	1

将抛钱币次数增大到100、1000、10000次以上，则事件B出现的频率将接近于0.5

实验者	抛硬币次数	出现正面次数	频率
蒲丰	4040	2048	0.5069
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

通过大量重复试验，我们得出了抛硬币试验出现正反面的机会相等这一规律。

2. 2 概率的定义

2.2.1. 概率的统计定义

根据频率的稳定性，在大量随机试验中，事件A出现的频率稳定地接近于某一定值。据此提出概率的统计定义：

对于随机事件A，当试验次数无限增多时，频率 $f(A)=\frac{l}{k}$ 稳定地接近于一个定值 p，那么我们就将这个 p 值定义为事件A发生的概率，记作 $P(A)=p$

概率的性质：

1.任何事件（A）的概率均满足 $0 \leq P(A) \leq 1$

2.必然事件（W）的概率为1，即$P(W)=1$

3.不可能事件（V）的概率为0，即 $P(V)=0$

概率的统计定义是在大量的试验中以频率的稳定性为基础提出来的。

它适用于一般的随机现象和随机事件。

一般的随机事件是指随着试验次数的改变，事件A出现的次数也是变化的，如抛硬币实验。

频率和概率是不相同的，只有当试验次数无限增大时，任一事件的频率趋于稳定，这时频率又称统计概率．这时的频率和概率才是一样的．

调查株数（n）	5	25	50	100	200	500	1000	1500	2000
受害株数（a）	2	12	15	33	72	177	351	525	704
植株受害频率（a/n）	0.40	0.48	0.30	0.33	0.36	0.354	0.351	0.350	0.352

在一般情况下，随机事件的概率p是不可能准确得到的。通常以试验次数n充分大时随机事件A的频率作为该随机事件概率的近似值。

2.2.2 小概率事件与小概率原理

当事件A的概率与0非常接近时，称此事件为**小概率事件**。小概率事件虽然不是不可能事件，但通常认为在一次试验中实际上是不可能发生的，称之为**“小概率事件实际不可能性原理”**。**这是统计假设检验的基础。**

总体与样本

总体：随机变量可能取值的全体。

样本：随机变量的n次独立观察值。以小写字母 x_i 表示第i次观察值。

观察值：随机变量所取得值称为观察值，用 X_i 表示

一个随机事件可用随机变量的关系式表示。

例：某品种成年兔体重X在2公斤到3公斤之间

这一随机事件可表示为2∠X∠3。

课前提问

- 标准差的含义及计算式是什么？
- 变异系数的应用条件、计算式
- 对随即现象的不确定性怎样研究？
- 什么是随机变量，试举例说明？
- 概率的统计定义
- 为什么要研究概率分布？

对于随机事件A，当试验次数无限增多时，

频率 $f(A)=\frac{l}{k}$稳定地接近于一个定值 p，那么

我们就将这个 p 值定义为事件A发生的概率，记作P(A)=p

本次内容

要点

- 离散型与连续型随机变量的特点
- 分别采用那些函数描述它们的概率分布
- 计算连续型随机变量在任意区间的概率
- 数学期望

2. 3 概率分布(probability distribution)

事件的概率表示了一次试验某一个结果发生的可能性大小。若要全面了解试验，则必须知道试验的全部可能结果及各种可能结果发生的概率，即必须知道随机试验的概率分布。为了深入研究随机试验，我们先引入随机变量(random variable)的概念。

2.3.1随机变量

一、随机变量（random variables）

（一）定义：

在随机试验中，被测定的量可取不同的值，且它究竟取何值具有偶然性，这样的变量称为随机变量。用大写字母X、Y、U、V等表示。

随机变量不同于一般数学上的变量，它取各种不同数值的规律是概率性的，从而表现出随机因素的影响。

（二）随机变量的分类

离散型随机变量

随机变量的可能值为有限个或可数无穷个孤立的数值，并可一一列举，则称为离散型随机变量。

特点：可一一列举

连续型随机变量

如果随机变量可取**某一区间**（有限或无限）内的任何数值，则称为连续型随机变量。

特点：随机变量的可能结果不能一一列举。

身高、体重、面积、体积等。如津丰小麦250株的高度是从55厘米到70厘米之间的任何值。

二、离散型随机变量的概率分布

1. 目的： 要了解离散型随机变量 X 的统计规律，就必须知道它的一切可能值及取每种可能值的概率.

研究概率分布的目的就是为了全面刻化一个随机变量。

2. 定义：

将X的一切可能值 x_1, x_2, … , x_n…，以及取得这些值的概率$p(x_1)$，$p(x_2)$…，$p(x_n)$，…排列起来，即**构成离散型随机变量的概率分布**。

3. 表示法：

可用概率分布表(列)和概率分布图

分布表（列）：表示离散型随机变量的概率分布：

变量x	x_1	x_2	…	x_n	…
概率P	p_1	p_2	…	p_n	…

概率分布图

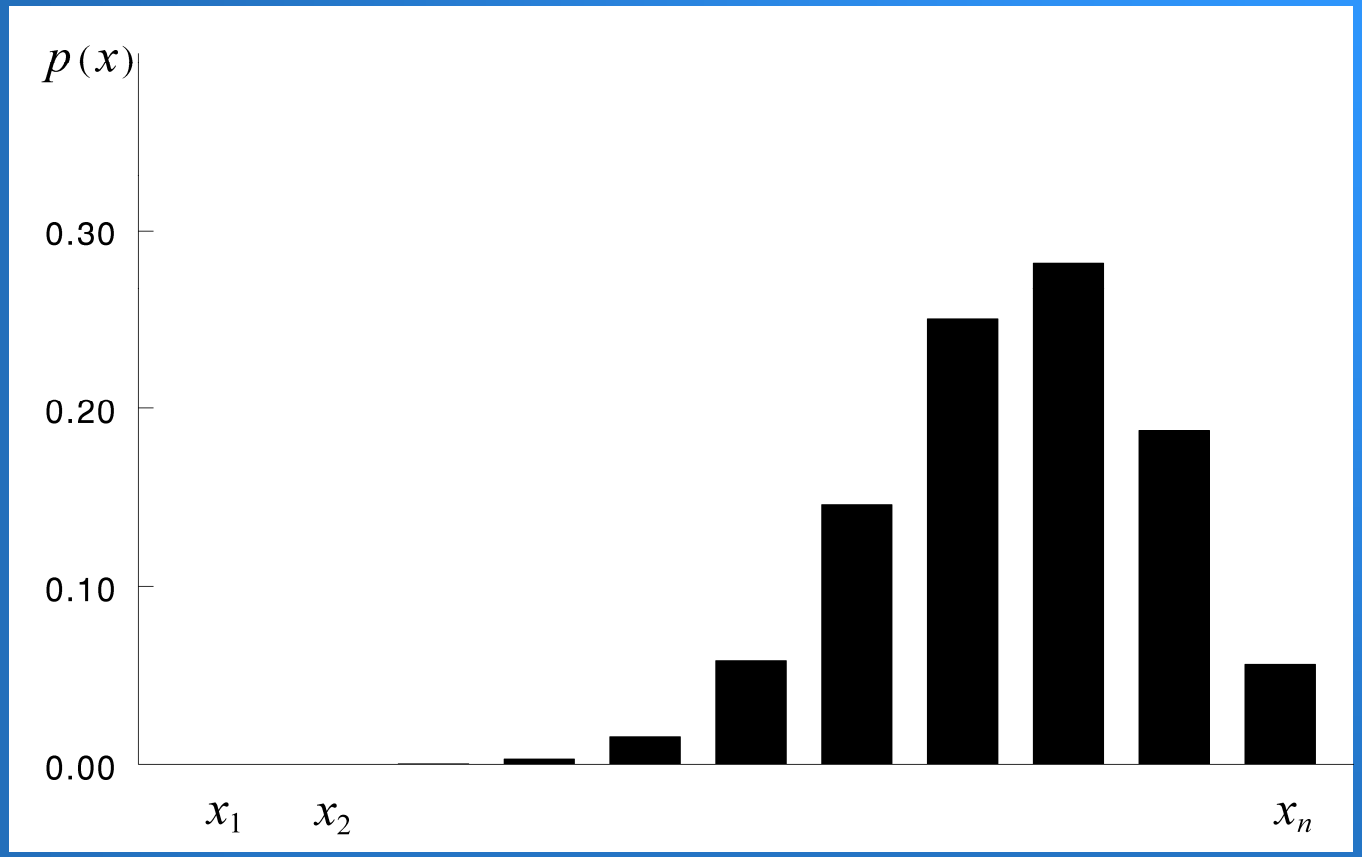


图2—1 离散型随机变量概率分布图

4. 概率函数：

将随机变量 X 所取得值 x_i 的概率 $P(X=x_i)$ 写成 X 的函数 $p(x)$，这样的函数称为随机变量 X 的概率函数

$$p(x)=P(X=x_i)$$

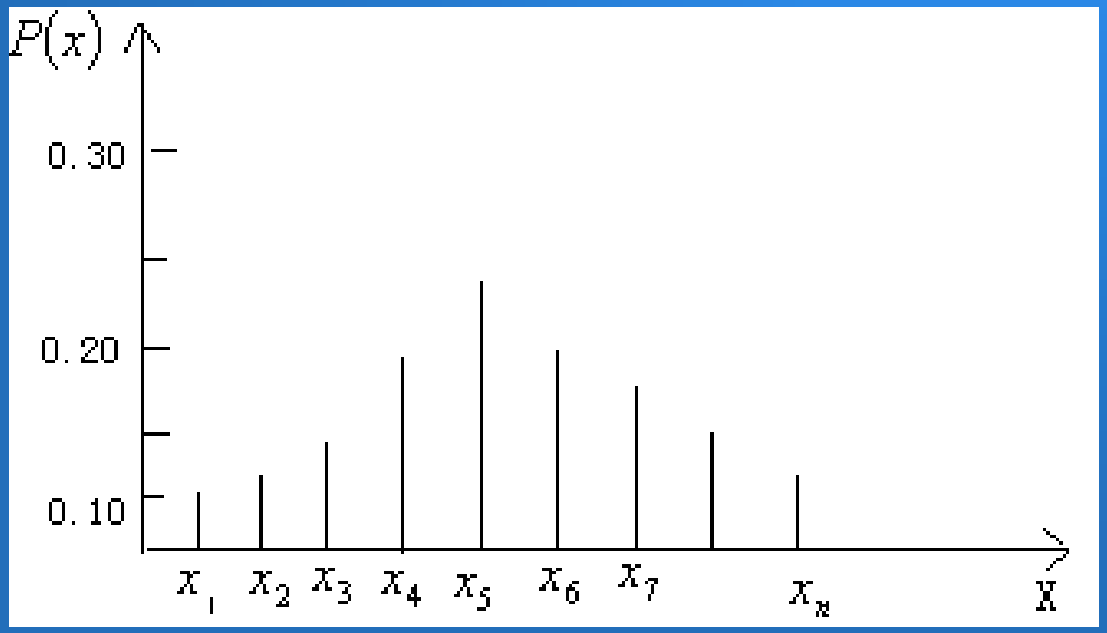
称为当 $X=x_i$ 时，随机变量X的概率函数，

概率函数应该满足：

$$\begin{cases} p(x)\geq 0 \\ \sum p(x)=1 \end{cases}$$

例：对某种鸟每窝产卵数分布的描述

X的取值	1	2	3	4	5	6	7	8	9	10	11	12
对应概率	0.01	0.03	0.08	0.15	0.30	0.25	0.09	0.05	0.02	0.01	0.01	0



小 结

- 研究概率分布的目的就是为了全面刻化一个随机变量。

- 离散型与连续型随机变量的特点分别是：

  可以一列举，不可以一列举

- 用概率函数描述离散型随机变量的概率分布。

  也就是说离散型随机变量的概率由概率函数确定。

频率分布与概率分布

频率分布称为统计分布或经验分布，用于描述样本.

概率分布称为理论分布或总体分布，用于描述总体.

频数分布图是实际观测的结果，有多种类型。对于各种不同的频数分布，都有其相应的理论分布来近似，即概率分布，这些理论分布是随机变量变化规律的理想化的数学模型，如：二项分布，正态分布等。

三、连续型随机变量的概率分布

连续型随机变量的概率分布不能用分布列来表示，因为其可能取的值是不可数的。

连续型随机变量可以是某一区间内的任何值，由于测量技术上的限制，在测量时，不可能无限提高精确度。

- 连续型随机变量取某一值的概率等于零。

- 只能用随机变量x在某个区间内取值的概率 $P(a \leqslant x < b)$来表示。

- 连续型随机变量的概率由概率分布密度函数确定

$$P(a \le x < b) = \int_a^b f(x)\,dx$$

（一）概率密度函数和概率密度分布曲线

连续型随机变量的实际观察值只能是落在某一小区间内的值，比如区间 $(x, x+\Delta x)$ 内，落在此区间内的概率值为 $P(x < X < x+\Delta x)$ ，Δx是区间长度。

当$\Delta x \to 0$时，$\dfrac{P(x < X < x+\Delta x)}{\Delta x}$的极限表示随机变量

X在点 X 处的概率密度函数，用 $f(x)$ 表示.

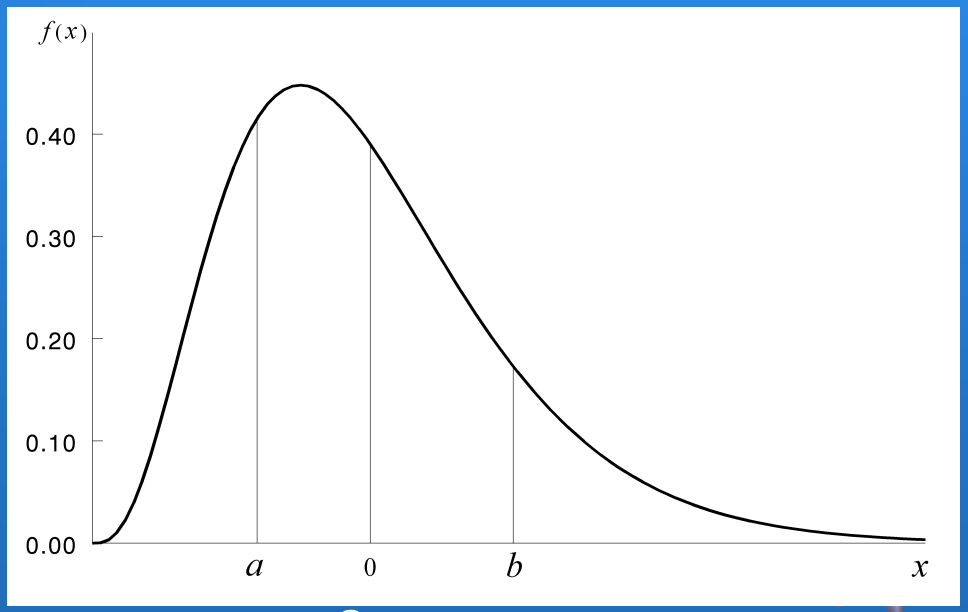
$$f(x) = \lim \frac{p(x \le X \le x+\Delta x)}{\Delta x}$$

- 概率密度函数，用 $f(x)$ 表示：

$$f(x)=\lim\frac{p(x\le X\le x+\Delta x)}{\Delta x}$$

连续型随机变量X落在区间（a，b）内的概率p(a<X<b) 也就是区间（a，b）所夹的曲线下的面积，

$$p(a\langle x\langle b)=\int_a^b f(x)dx$$



概率分布密度曲线

概率密度函数对x作图，得到概率密度的分布曲线。

$$f(x)=\frac{8}{3}(x+2)^3e^{-2x-4}$$

（二）连续型随机变量概率分布的性质

❶分布密度函数大于或等于0，即 $f(x) \geq 0$

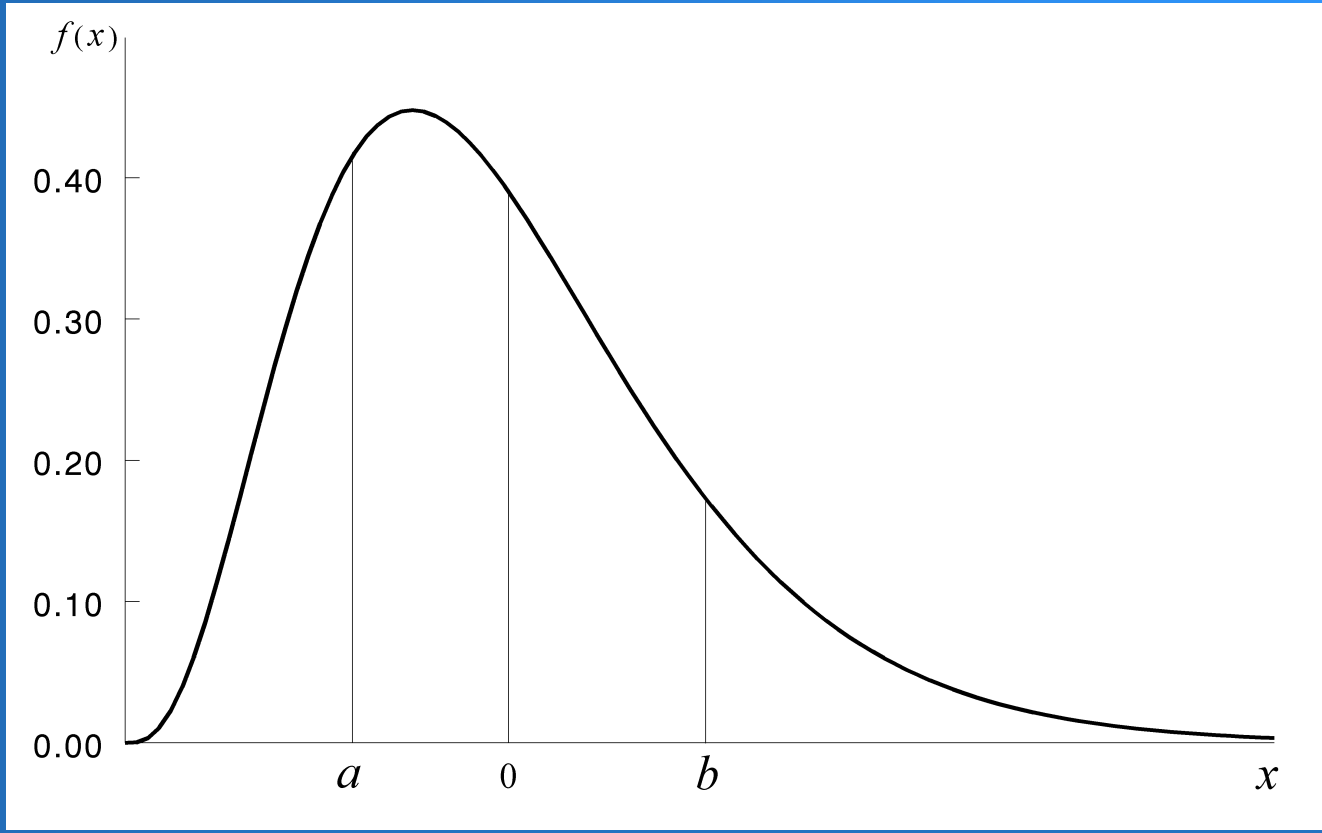
❷当随机变量x取某一特定值时，其概率为0，即

$$p(x=c)=\int_c^c f(x)dx=0$$

c为任意实数

❸在一次试验中x取值必在 $-\infty < x < +\infty$ 范围内，为一必然事件。因此

$$p(-\infty < x < +\infty)=\int_{-\infty}^{+\infty} f(x)dx=1$$



连续型随机变量的概率密度分布曲线

四、累积函数（分布函数）

分布函数：随机变量取得小于 X_0 的值的概率，记为F（x_0）

设X是一个随机变量, 称函数 $F(x_0)=p(x\le x_0),(-\infty<x<+\infty)$

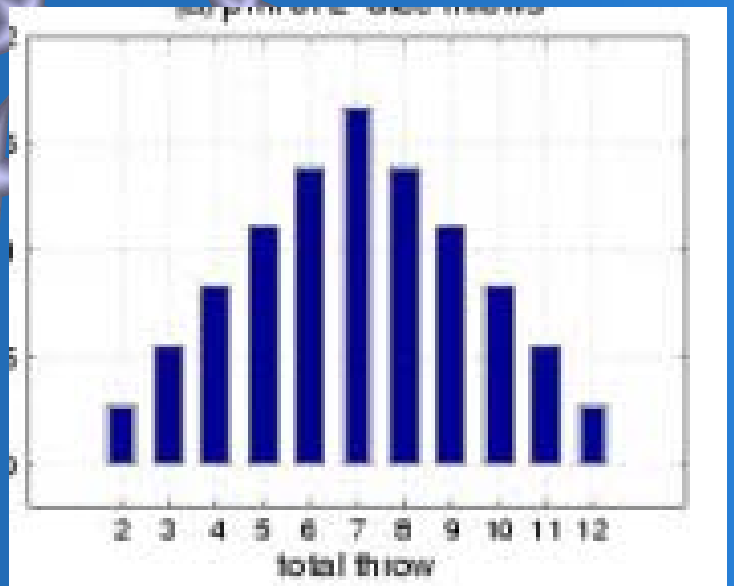
$$F(x_0)=\sum_{x_i\le x_0}p(x_0)=p(x\le x_0)$$ 为X的分布函数。

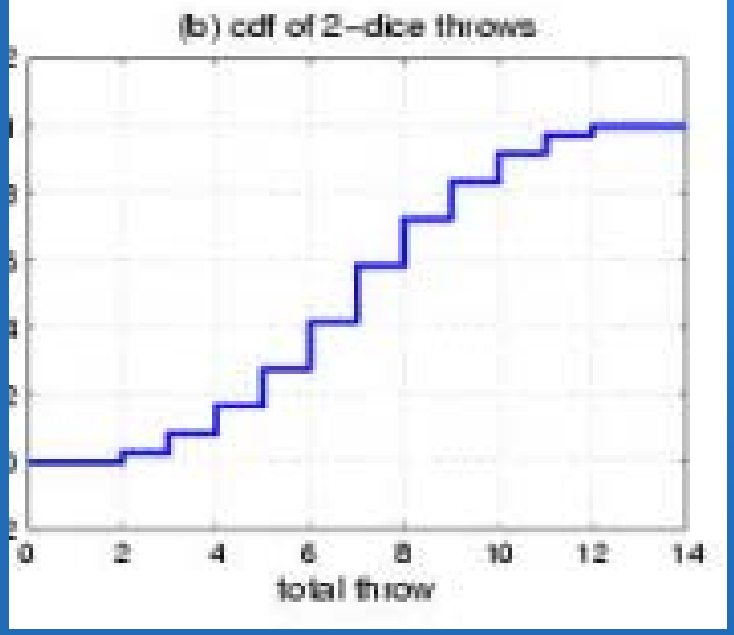
（1）若X为离散型随机变量，设X的概率函数是 $p(x_i)$，

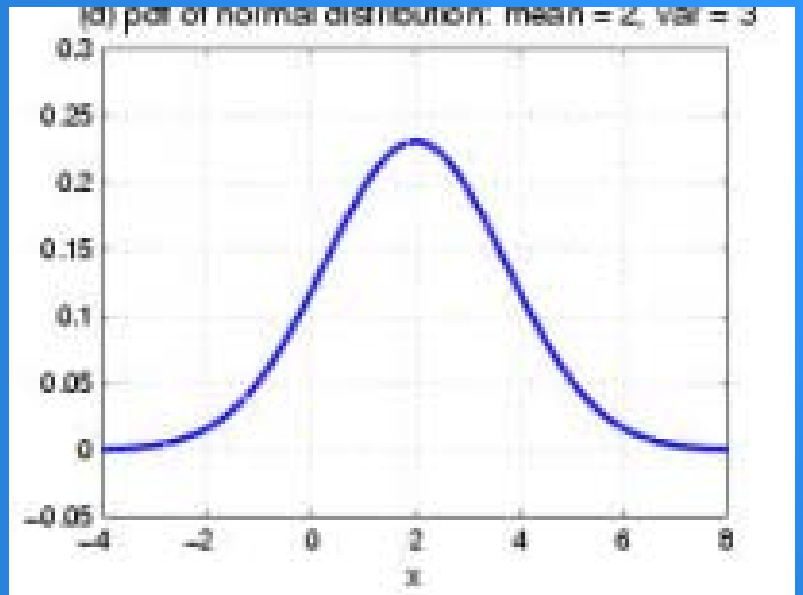
i=1,2,……则其分布函数为：$F(x_0)=p(x\le x_0)=\sum_{x_i\le x_0}p(x_i)$

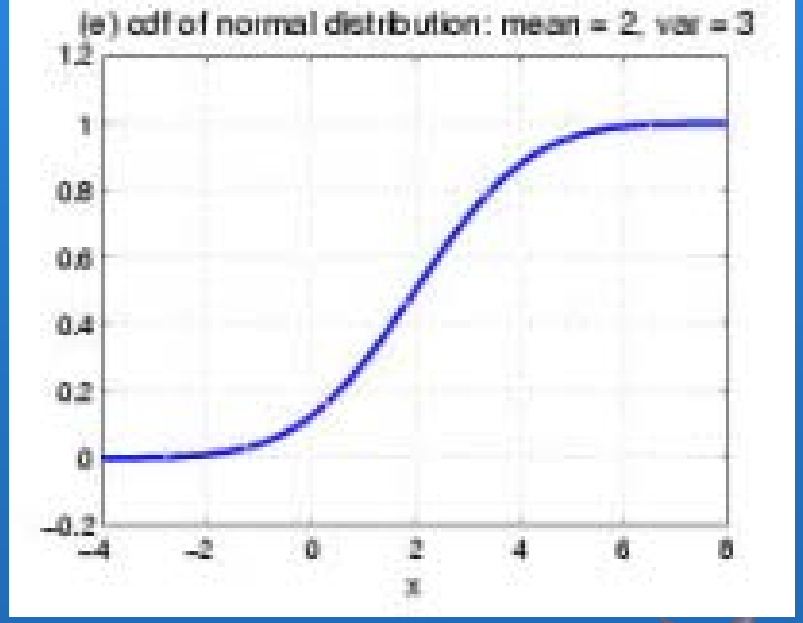
即X等于或小于某一可能值（x_0）的概率。

例 F(3)=p(1)+ p(2) + p(3)


2 3 4 5 6 7 8 9 10 11 12
total throw


(b) cdf of 2−dice throws
0 2 4 6 8 10 12 14
total throw


(d) pdf of normal distribution: mean = 2, var = 3
0.3
0.25
0.2
0.15
0.1
0.05
0
−0.05
−4 −2 0 2 4 6 8
x

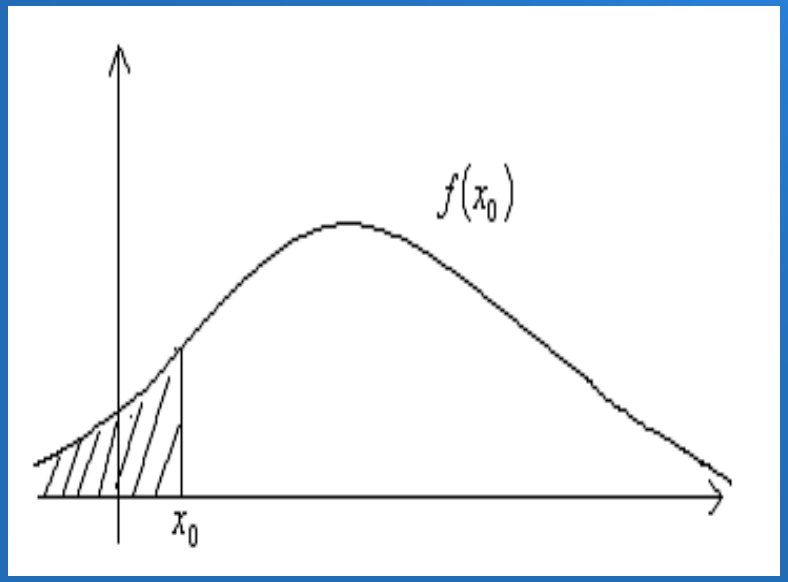

(e) cdf of normal distribution: mean = 2, var = 3
1.2
1
0.8
0.6
0.4
0.2
0
−0.2
−4 −2 0 2 4 6 8
x

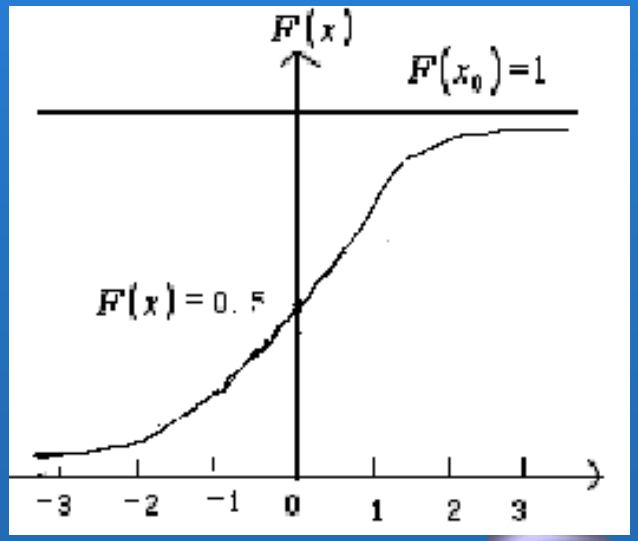
2）若X为连续型随机变量，设X的密度函数为f(x)，则其分布函数为：

$$F(x_0) = p(x \le x_0) = \int_{-\infty}^{x_0} f(x)dx$$

且满足F(-∞)=0, F(+∞)=1

$F(x_0)$ 的数值等于分布曲线中点 x_0 左边的曲线下的面积，则任一区间内的概率就可以用分布函数来表示。





$$p(x<a)+p(a<x<b)=p(x<b)$$

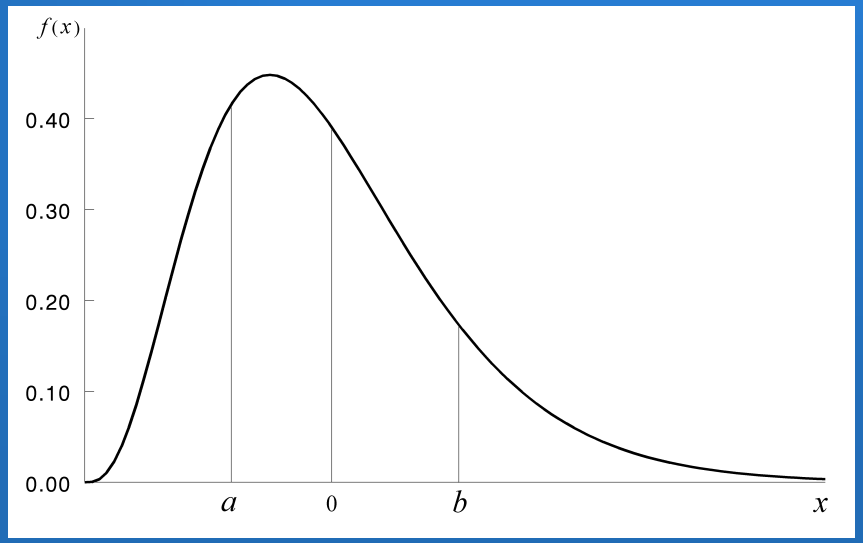
即：$F(a)=p(X<a), F(b)=p(X<b)$

$$\therefore\ p(a<x<b)=F(b)-F(a)$$

$$p(a\le x<b)=F(b)-F(a)$$

$$p(a<x\le b)=F(b)-F(a)$$

$$p(a\le x\le b)=F(b)-F(a)$$



小 结

- 用概率密度函数描述连续型随机变量的概率分布。

  也就是说连续型随机变量的概率由概率密度函数确定。

- 计算连续型随机变量在任意区间的概率可以用

概率密度函数：$p(a\langle x\langle b)=\int_a^b f(x)dx$

累积函数：$p(a<x<b)=F(b)-F(a)$

2.4　随机变量的数学期望和方差

平均值是衡量集中程度的量，标准差是衡量离散程度的量，它们能够表明随机变量的主要性质，作为随机变量的数字特征或特征数。

- 样本特征数：

  样本平均数$\overline{X}$，样本标准差 S。

- 总体特征数：

  总体的平均数 μ，标准差σ，或方差σ^2 。

2.4.1 随机变量的数学期望（总体平均数）

例：测定一批种子发芽所需的平均天数，用100粒种子进行发芽试验，其中发芽的有98粒，按发芽天数列成下表。

发芽天数	1	2	3	4	5	6	7
发芽种子数	19	34	21	11	9	3	1

每粒种子发芽所需的平均天数为：

（1×19+2×34+3×21+4×11+5×9+6×3+7×1）/ 98=2.7（天）

$$\frac{1}{98}(1\times 19+2\times 34+3\times 21+4\times 11+5\times 9+6\times 3+7\times 1)$$

$$=1\times\frac{19}{98}+2\times\frac{34}{98.}+3\times\frac{21}{98}+4\times\frac{11}{98}+5\times\frac{9}{98}+6\times\frac{3}{98}+7\times\frac{1}{98}$$

$$=0.194+0.694+0.643+0.449+0.459+0.184+0.071$$

$$=2.69$$

从上述发芽试验：

$$\bar{x}=\frac{\sum_{i=1}^{k} f_i x_i}{n}=\sum_{i=1}^{k} x_i\left(\frac{f_i}{n}\right)\xrightarrow{n\to\infty}\sum x_i p_i$$

这个值就是随机变量的总体平均值，也叫做随机变量的数学期望，即理论平均数（含义）。

记作 $E(X)=\sum x_i p(x_i)=\mu$

同样上述发芽试验，每种发芽天数与平均发芽天数之间有变异，其平均的变异程度可用S^2表示：

$$S^2=\sum_{i=1}^{k}\frac{f_i}{n-1}\left(x_i-\bar{x}\right)^2 \xrightarrow{n-1\to\infty} \sum p\left(x_i\right)\left(x_i-\mu\right)^2$$

记作 $\mathrm{var}\left(x\right)=\sigma^2$ 叫做总体方差。

X或X的函数的数学期望可用通式表示：

$$E[g(x)] = \sum_{x} p(x)g(x)$$

当g(x)为x时，表示总体平均数。

g(x)为（x−μ）^{2}时，表示总体方差。

连续型随机变量的数学期望定义为：

$$\mu = E(X) = \int xf(x)dx$$

连续型随机变量的方差定义为：

$$\sigma^2 = E(X-\mu)^2 = \int_{-} (x-\mu)^2 f(x)dx$$

例：用一种复合饲料饲养动物，每天增重斤数及其相应概率如下：

每天增重x_i（斤）	0.5	1.0	1.5	2.0
Pi	0.10	0.20	0.50	0.20

问：每天增重的数学期望是多少？平均增重多少？

解：

$$E(x)=\sum_{x}p(x)x=\sum_{i=1}^{4}p(x_i)x_i=0.10\times0.5+0.20\times1.0+0.50\times1.5+0.20\times2.0=1.4\text{斤}$$

例：A、B两块棉田，种植的棉花株数相同，每株棉花上棉铃虫的头数分别用X、Y表示。经过一段时间的考察，得到X、Y的概率分布如下：

X	0	1	2	3	Y	0	1	2	3
P_x	0.7	0.1	0.1	0.1	p_y	0.5	0.3	0.2	0.0

试问哪一块田的棉铃虫多？

解：用X、Y的数学期望来比较：

$$E(x)=0\times0.7+1\times0.1+2\times0.1+3\times0.1=0.6$$
$$E(y)=0\times0.5+1\times0.3+2\times0.2+3\times0.0=0.7$$
$$\therefore E(y)>E(x)$$

说明棉田B平均每株上的虫数较高，而两块棉田的株数相同，故棉田B的棉铃虫多些.

2.4.2 特征数的运算

1.数学期望

（1）E（C）=C 常量的均值是其自身。

（2）E（CX）=CE（X）常量c可以从括号里提出来。

（3）E（X+C）=E（X）+C

（4）E（CX+A）=CE（X）+A

（5）E（X_1+X_2）=E（X_1）+E（X_2）

（6） E（X_1X_2）=E（X_1）E（X_2）（ X_1、X_2 相互独立）

2.方差

赋予总体方差σ^2=var(X) （variance）

$$\sigma^2 = \sum p(x)(x-\mu)^2 = E(x-\mu)^2$$

方差具有下列性质：

（1）var（c）=0

(2)var(c+X)=var(X)

(3)var(cX)=c^2var(X)

（4）var(cX+A)=c^2var(X)

小测验 1

1.生物统计学的任务是什么？

2.如何对随机现象的不确定性进行研究？

3.频数资料的算术平均数如何计算？

4.概率的统计定义？

5.为什么要研究概率分布?离散型与连续型随机变量的特点？分别采用那些函数描述它们的概率分布？

6. 标准差的含义及计算式是什么？

7.变异系数的应用条件及其计算式是什么？

8. 数学期望的含义是什么？如何推导出E(X)和Var(X) ？

9 用哪两种方法可以计算连续型随机变量在任意区间的概率？

课前提问

- 算术平均数和加权平均数的计算
- 标准差的意义、性质和计算公式
- 实验数据的科学表示法
- 变异系数的意义、应用条件
- 对随即现象的不确定性怎样研究？
- 概率的统计定义
- 什么是随机变量，试举例说明？