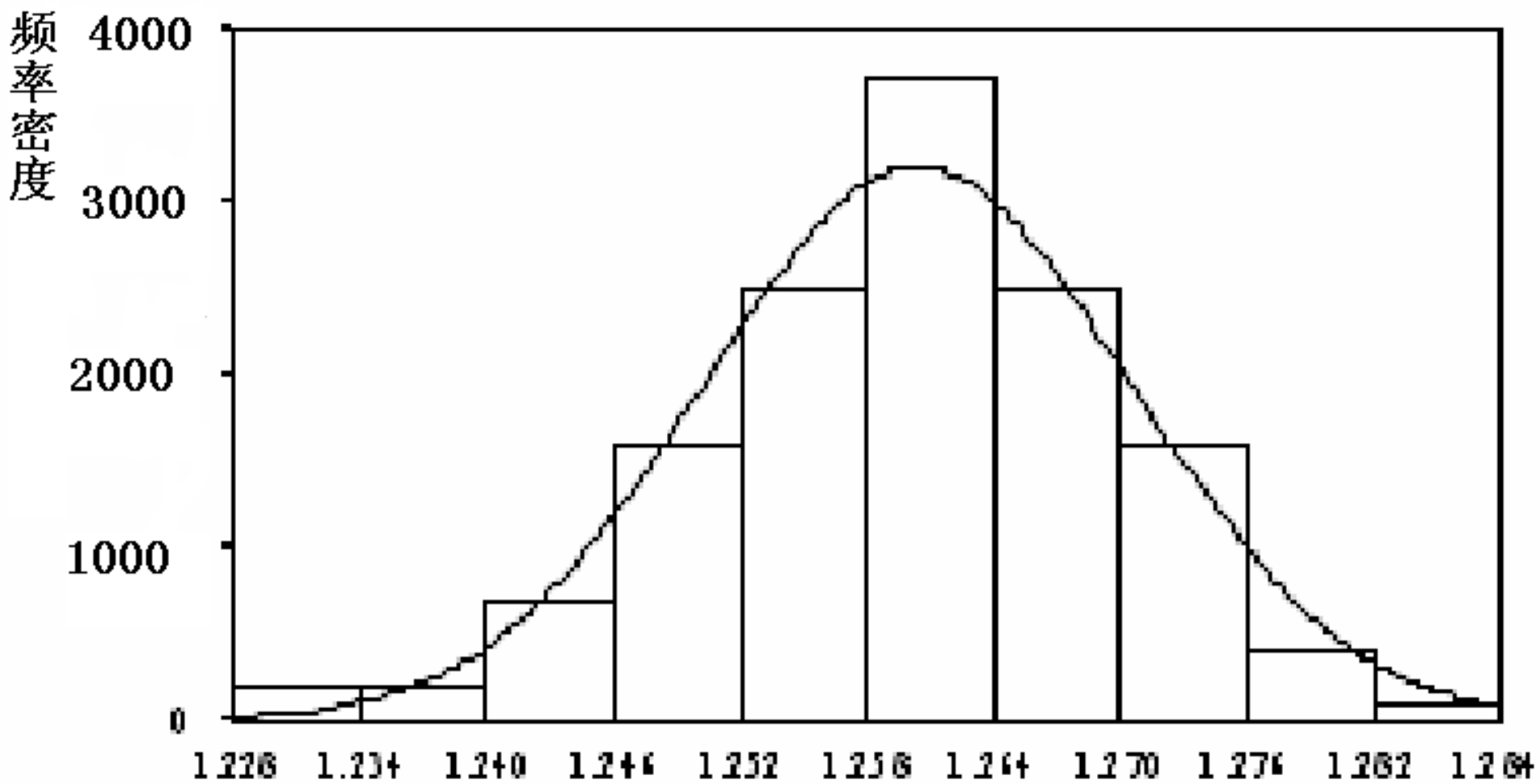


第四章 常用概率分布

第一节 正态分布

一、正态分布的概念

正态分布是自然界最常见的一种分布，若指标 X 的频率分布曲线对应于数学上的正态分布曲线，则称该指标服从正态分布。



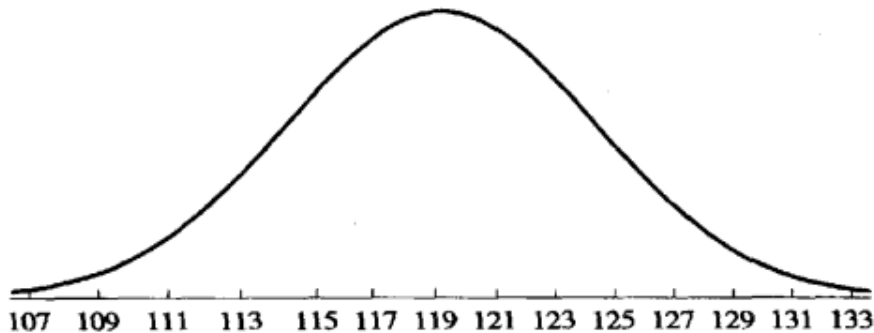
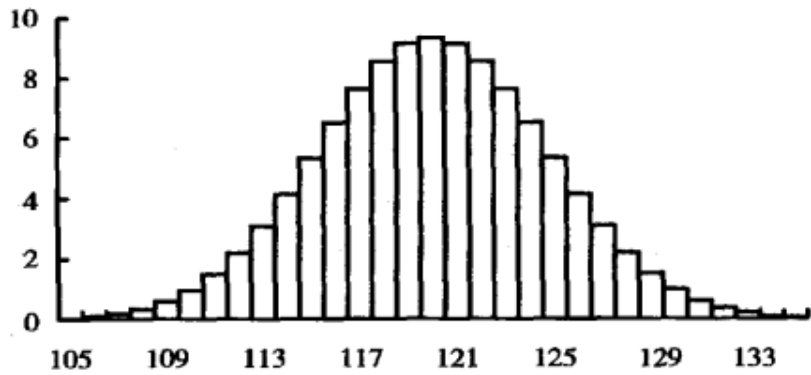
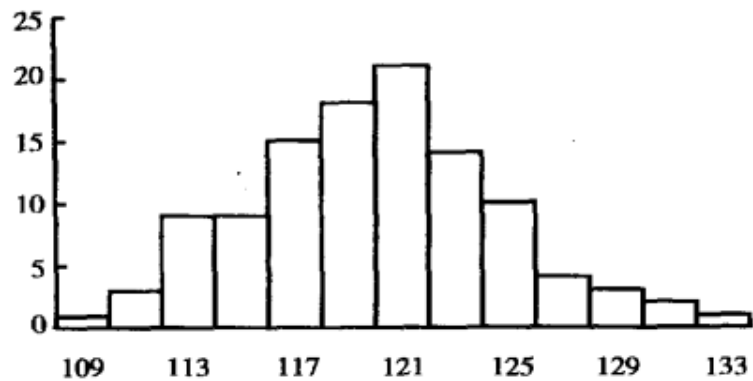
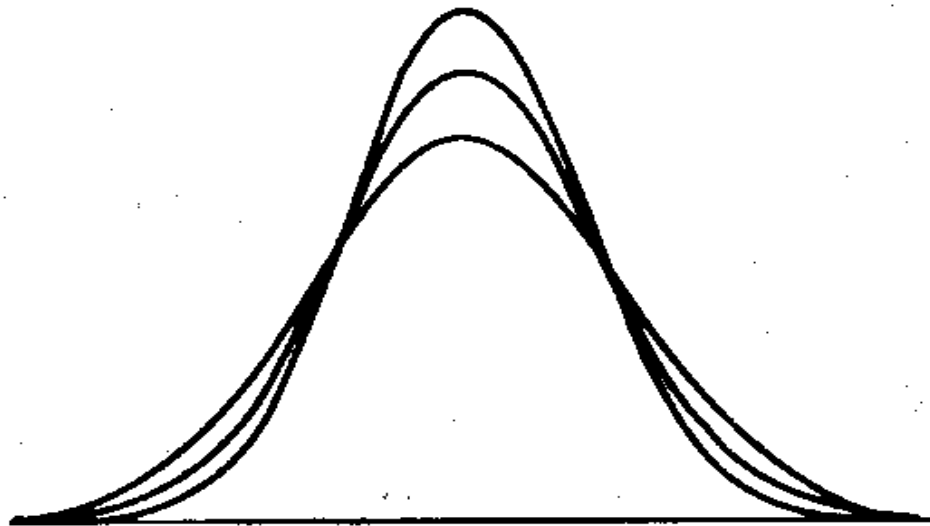


图 3.1 频数分布逐渐接近正态分布示意图

正态分布的概率密度函数（即纵轴的高度）

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad -\infty < X < +\infty$$



b. 均数相同、标准差不同($\sigma_1 < \sigma_2 < \sigma_3$)的三条正态曲线

图 4-5 正态曲线位置、形状与 μ 、 σ 关系示意图

均数为0，标准差为1的正态分布，这种正态分布称为**标准正态分布**。

对于任意一个服从正态分布 $N(\mu, \sigma^2)$ 的随机变量，可作如下的**标准化变换**，也称**Z变换**，

$$Z = \frac{X - \mu}{\sigma}$$

标准正态分布的密度函数：

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad -\infty < z < +\infty$$

正态分布的特征

1. 关于 $x = \mu$ 对称。即正态分布以均数为中心，左右对称。
2. 在 $x = \mu$ 处取得概率密度函数的最大值，在 $x = \mu \pm \sigma$ 处有拐点，表现为钟形曲线。即正态曲线在横轴上方均数处最高。

3. 正态分布有两个参数，即均数 μ 和标准差 σ 。
 μ 是位置参数， σ 是变异度参数(形状参数)。常用
 $N(\mu, \sigma^2)$ 表示均数为 μ ，标准差为 σ 的正态分布；用
 $N(0, 1)$ 表示标准正态分布。

4. 正态曲线下面积分布有一定规律。横轴上
正态曲线下的面积等于1（也常写作100%）。

二、正态曲线下面积的分布规律

正态方程的积分式(分布函数):

$$F(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^X e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$F(X)$ 为正态变量 X 的累计分布函数，反映正态曲线下，横轴尺度自 $-\infty$ 到 X 的面积，即下侧累计面积。

标准正态分布方程积分式(分布函数):

$$\Phi(Z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^Z e^{-\frac{z^2}{2}} dz$$

$\Phi(Z)$ 为标准正态变量 Z 的累计分布函数，反映标准正态曲线下，横轴尺度自 $-\infty$ 到 Z 的面积，即下侧累计面积。

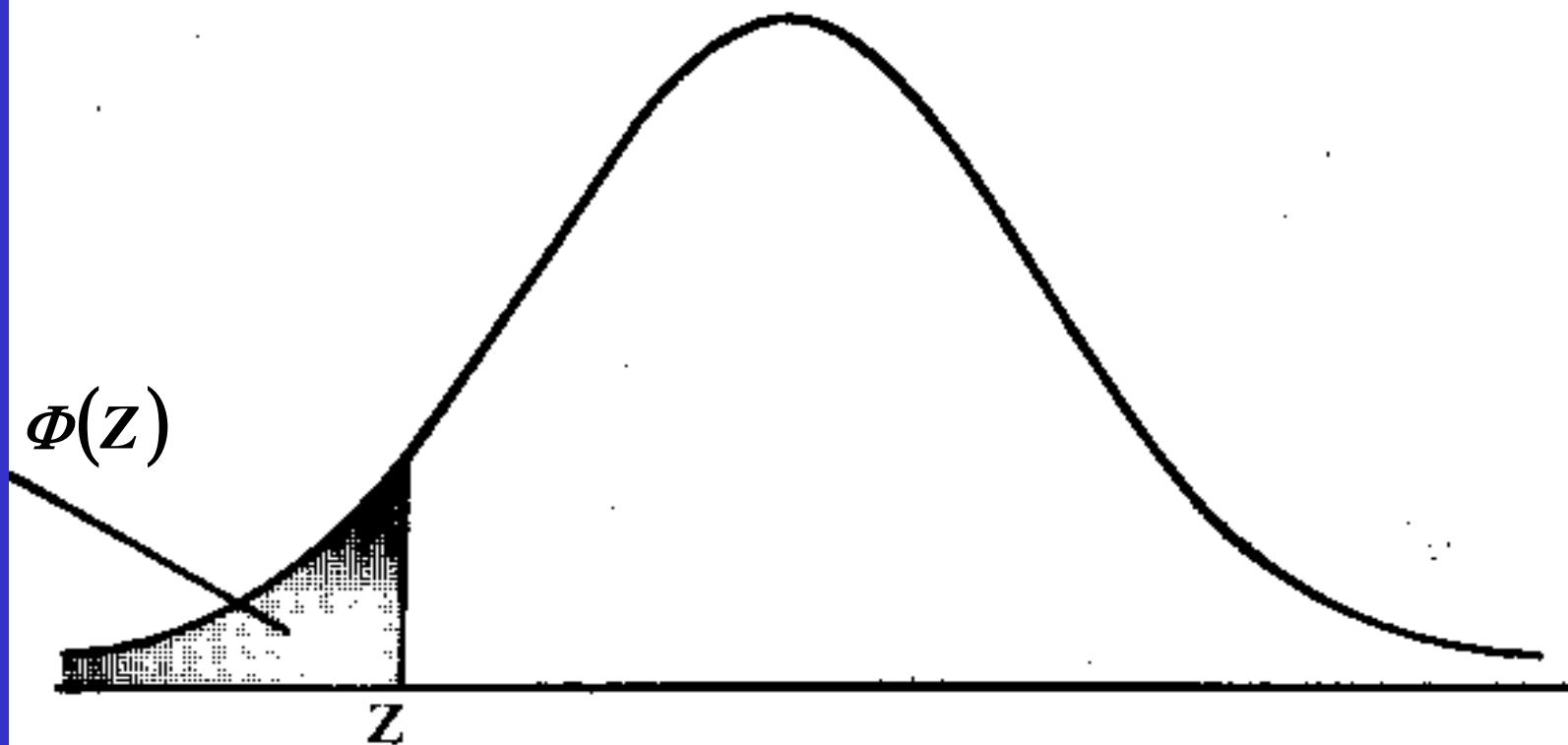


图 4-7 标准正态分布的
分布函数示意图

三、标准正态分布表

用查表代替计算必须注意：

1) 表中曲线下面积为 $-\infty$ 到Z的面积。

2) 当 μ, σ 和X已知时，先求出Z值，
$$Z = \frac{(X - \mu)}{\sigma}$$

再用Z值查表，得所求区间占总面积的比例。

当 μ 和 σ 未知时，要用样本均数和样本标准差S来估计Z值。

$$Z = \frac{(X - \bar{X})}{S}$$

3) 曲线下对称于0的区间，面积相等。

4) 曲线下横轴上的面积为1（即100%）。

正态分布是一种对称分布，其对称轴为直线 $X=\mu$ ，即均数位置，理论上：

$\mu \pm 1\sigma$ 范围内曲线下的面积占总面积的68.27%

$\mu \pm 1.96\sigma$ 范围内曲线下的面积占总面积的95%

$\mu \pm 2.58\sigma$ 范围内曲线下的面积占总面积的99%

实际应用中：

$\bar{X} \pm 1 S$ 范围内曲线下的面积占总面积的68.27%

$\bar{X} \pm 1.96 S$ 范围内曲线下的面积占总面积的95%

$\bar{X} \pm 2.58 S$ 范围内曲线下的面积占总面积的99%

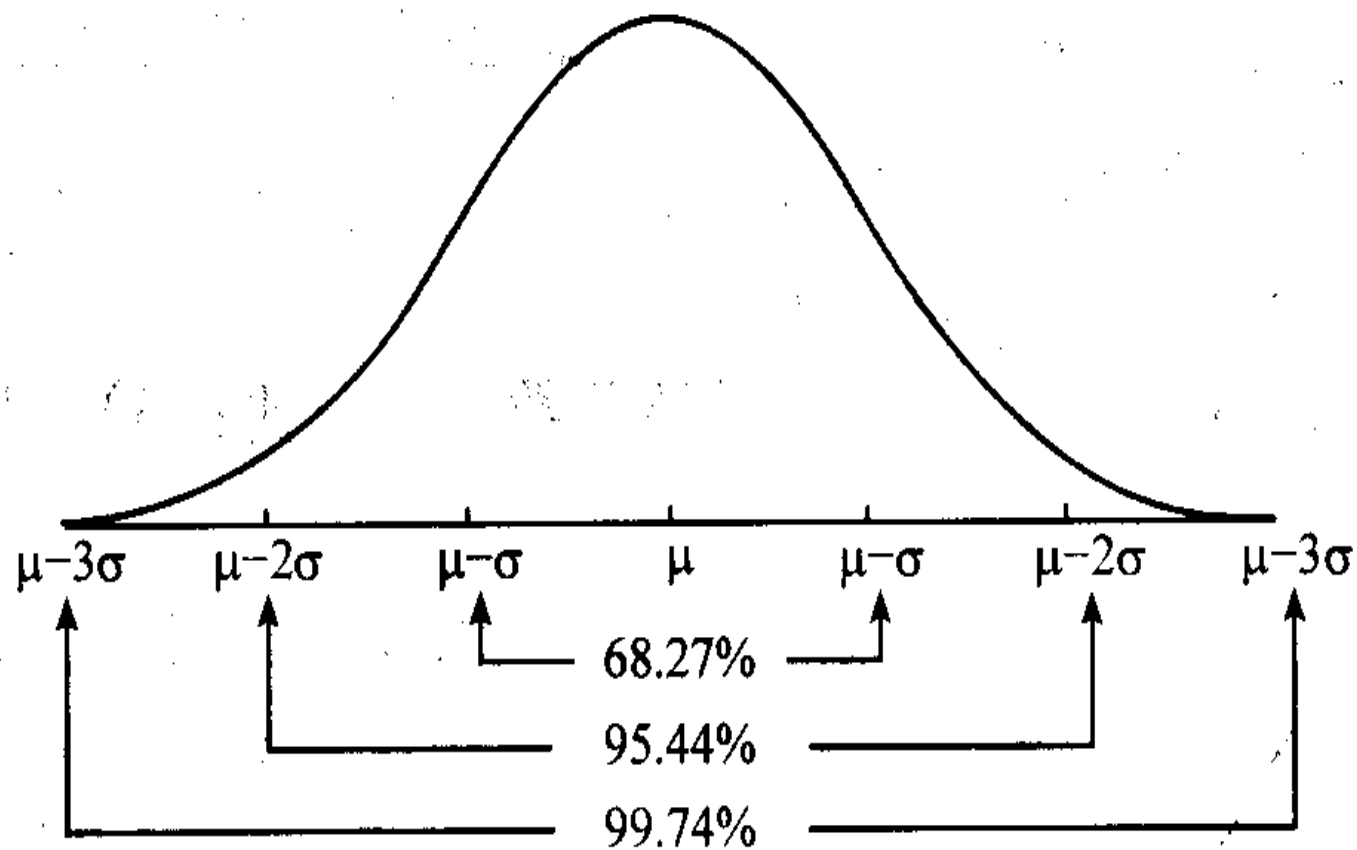


图 4-6 正态分布曲线下的面积分布

标准正态分布的 $\mu=0$, $\sigma=1$, 则

$\mu \pm \sigma$ 相当于区间 $(-1, 1)$,

$\mu \pm 1.96\sigma$ 相当于区间 $(-1.96, 1.96)$,

$\mu \pm 2.58\sigma$ 的区间相当于区间 $(-2.58, 2.58)$ 。

区间 $(-1,1)$ 的面积: $1-2\Phi(-1)=1-2 \times 0.1587=0.6826=68.26\%$

区间 $(-1.96,1.96)$ 的面积: $1-2\Phi(-1.96)=1-2 \times 0.0250=0.9500=95.00\%$

区间 $(-2.58,2.58)$ 的面积: $1-2\Phi(-2.58)=1-2 \times 0.0049=0.9902=99.02\%$

例 4-11 X服从均数为 μ ，标准差为 σ 的正态分布，试估计（1）X取值在区间 $\mu \pm 1.96\sigma$ 上的概率；
 （2）X取值在区间 $\mu \pm 2.58\sigma$ 的概率；

先做标准化变换：



正态曲线下面积对称，则区间 $(1.96, \infty)$ 的面积也是0.025。Z取值于 $(-1.96, 1.96)$ 的概率为 $1 - 2 \times 0.025 = 0.95$ ，即X取值在区间 $\mu \pm 1.96\sigma$ 上的概率为95%。

例 4-12 已知某地1986年120名8岁男童身高均数 $\bar{X}=120.2$ ， $S=4.79$ cm，估计(1)该地8岁男孩身高在130 cm以上者占该地8岁男孩总数的百分比；(2)身高界于120cm~128cm者占该地8岁男孩总数的比例；(3)该地80%男孩身高集中在哪个范围？

先做标准化变化：

$$Z = \frac{X - \bar{X}}{S} = \frac{X - 120.2}{4.79}$$

$$(1) \rightarrow (14.0)'$$

理论上该地8岁男孩身高在130 cm以上者占该地8岁男孩总数的7.21%。

(2)

$$z_1 = \frac{1202032}{479} = 06$$

~~(1) (10) (3)~~

$$z_2 = \frac{1282032}{479} = 104$$

~~(2) (10) (3)~~

~~(1) (1) (3) (3)~~

(3)

查附表1，标准正态分布曲线下左侧面积为0.10所对应的Z值为-1.28，所以80%的8岁男孩身高值集中在 $\bar{X} \pm 1.28S$ 区间内，即116.9cm~129.2cm

四、正态分布的应用

(一) 制定医学参考值范围

❖ 参考值范围：指特定的“正常”人群的解剖、生理、生化、免疫等各种数据的波动范围。

❖ 制定参考值范围的步骤：

1. 选择足够数量的正常人作为调查对象。
2. 样本含量足够大。
3. 确定取单侧还是取双侧正常值范围。
4. 选择适当的百分界限。
5. 选择适当的计算方法。

❖ 估计医学参考值范围的方法:

1. **正态近似法**: 适用于正态分布或近似正态分布的资料。
2. **百分位数法**: 适用于偏态分布资料。

%	正态分布法			百分位数法		
	双侧	单侧 只有下限	单侧 只有上限	双侧	单侧 只有下限	单侧 只有上限
90	$\bar{X} \pm 1.64S$	$\bar{X} - 1.28S$	$\bar{X} + 1.28S$	$P_5 \sim P_{95}$	P_{10}	P_{90}
95	$\bar{X} \pm 1.96S$	$\bar{X} - 1.64S$	$\bar{X} + 1.64S$	$P_{2.5} \sim P_{97.5}$	P_5	P_{95}
99	$\bar{X} \pm 2.58S$	$\bar{X} - 2.33S$	$\bar{X} + 2.33S$	$P_{0.5} \sim P_{99.5}$	P_1	P_{99}

制定

过低异常 过高异常

过低异常 过高异常

例4-13 某地调查120名健康女性血红蛋白，直方图显示，其分布近似于正态分布，得均数为117.4g/L，标准差为10.2g/L，试估计该地正常女性血红蛋白的95%医学参考值范围。

分析：正常人的血红蛋白过高过低均为异常，要制定双侧正常值范围。

~~$117.4 \pm 1.96 \times 10.2$~~

该指标的95%医学参考值范围为 $97.1133(g/L)$

例4.14 某地调查110名正常成年男子的第一秒肺通气量，得均数为4.2 L，标准差为0.7 L，试估计该地正常成年男子第一秒肺通气量的95%参考值范围。

分析：正常人的第一秒肺通气量近似正态分布，且只以过低为异常，要制定单侧下限。

$$\bar{x} - 1.645s$$

该地正常成年男子第一秒肺通气量的95%参考值范围为：不低于3.052L。

例 4.15 某年某市调查了 200 例正常成人血铅含量 ($\mu\text{g}/100\text{g}$) 如下, 试估计该市成人血铅含量的 95% 医学参考值范围。

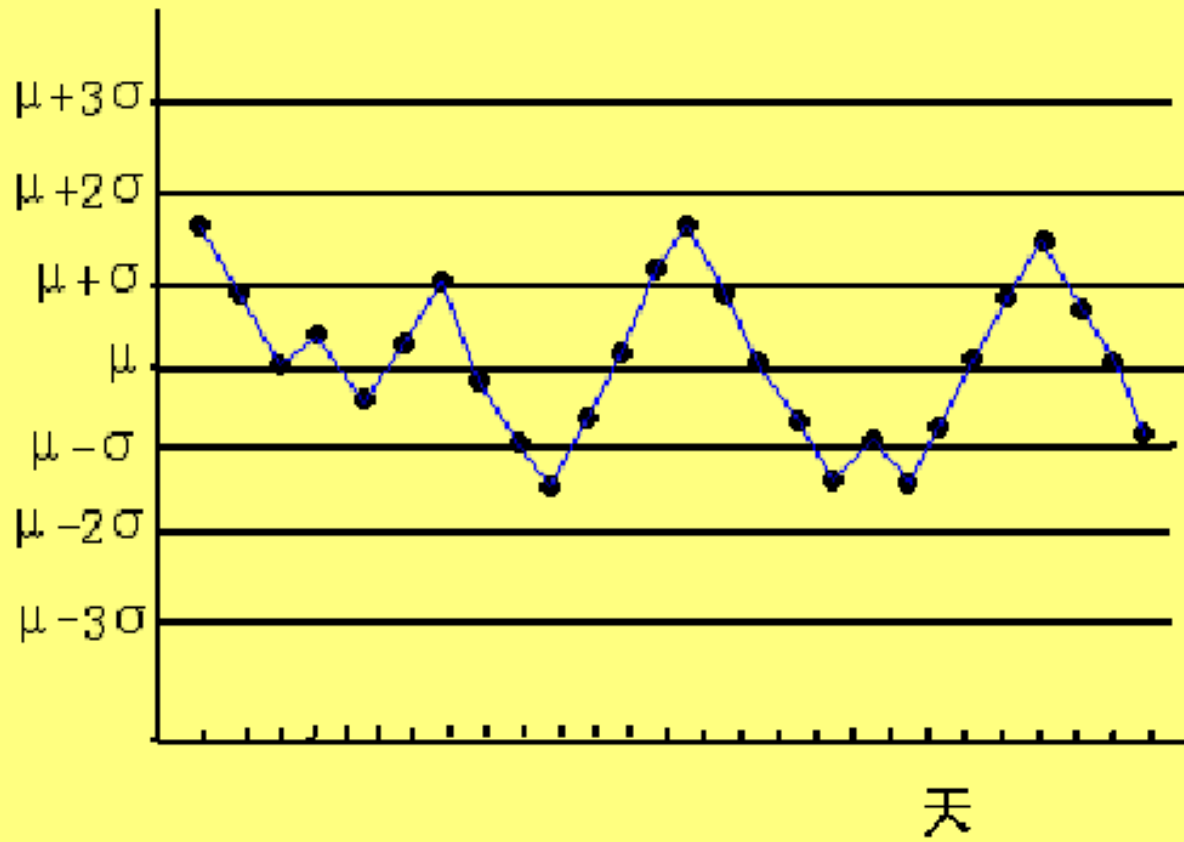
表 7-5 200 名血铅频数表及 P_{95} 计算表

组段	频数 f	累计频数 Σf	累计频率 (%)
3~	36	36	18.0
8~	39	75	37.5
12~	47	122	61.0
18~	20	152	76.0
23~	18	170	85.0
28~	16	186	93.0
33~	3	189	94.5
38~	7	196	98.0
43~	1	197	98.5
48~	1	198	99.0
53~	1	199	99.5
58~62	1	200	100.0

分析：血铅的分布为偏峰分布，且血铅含量只以过高为异常，要用百分位数法制定单侧上限。



二、质量控制



质量控制图

第三节 正态分布及其应用

判断异常的8种情况是：

- ❖ 有一个点距中心线的距离超过3个标准差（控制限以外）
- ❖ 在中心线的一侧连续有9个点
- ❖ 连续6个点稳定地增加或减少
- ❖ 连续14个点交替上下
- ❖ 连续3个点中有两个点距中心线距离超过2个标准差（警戒限以外）

第三节 正态分布及其应用

- ❖ 连续5个点中有4个点距中心线距离超过1个标准差
- ❖ 中心线一侧或两侧连续15个点距中心线距离都在1个标准差以内
- ❖ 中心线一侧或两侧连续8个点距中心线距离都超出1个标准差范围。

三、统计处理方法的理论基础

如 统计描述中计算算术平均数、标准差、
统计推断中进行总体均数置信区间估计、
t 检验、F 检验、相关与回归等分析

第二节 二项分布

一、二项分布的概念与特征

(一) 成败型实验 (Bernoulli 实验)

在医学卫生领域的许多实验或观察中，人们感兴趣的是某事件是否发生。如用白鼠做某药物的毒性实验，关心的是白鼠是否死亡；某种新疗法临床实验观察患者是否治愈；观察某指标的化验结果是否呈阳性等。将我们关心的事件A出现称为成功，不出现称为失败，这类试验就称为成-败型实验。指定性资料中的二项分类实验。

成-败型 (Bernoulli) 实验序列:

满足以下三个条件的 n 次实验构成的序列称为成-败型实验序列。

1) 每次实验结果, 只能是两个互斥的结果之一 (A或非A)。

2) 相同的实验条件下, 每次实验中事件A的发生具有相同的概率 π 。(非A的概率为 $1-\pi$)。

实际工作中要求 π 是从大量观察中获得的较稳定的数值。

3) 各次实验独立。各次的实验结果互不影响。

(二) 二项分布的概率函数

二项分布是指在只能产生两种可能结果（如“阳性”或“阴性”）之一的 n 次独立重复实验中，当每次试验的“阳性”概率保持不变时，出现“阳性”的次数 $X=0, 1, 2, \dots, n$ 的一种概率分布。

若从阳性率为 π 的总体中随机抽取大小为 n 的样本，则出现“阳性”数为 X 的概率分布即呈现二项分布，记作 $B(n, \pi)$ 。

举例 设实验白鼠共3只，要求它们同种属、同性别、体重相近，且他们有相同的死亡概率，即事件“白鼠用药后死亡”为A，相应死亡概率为 π 。记事件“白鼠用药后不死亡” \bar{A} 为，相应不死亡概率为 $1-\pi$ 。设实验后3只白鼠中死亡的白鼠数为X，则X的可能取值为0，1，2和3，则死亡鼠数为X的概率分布即表现为二项分布。

表 7.1 3 只白鼠各种试验结果及其发生概率

死亡数 X	未死亡数 $3-X$	试验结果		
		甲	乙	丙
0	3	生	生	生
1	2	死	生	生
		生	死	生
		生	生	死
2	1	死	死	生
		死	生	死
		生	死	死
3	0	死	死	死

互不相容事件的
加法定理

独立事件的
乘法定理

$$\begin{aligned}(A + B)^3 &= A^3 + 3A^2B + 3AB^2 + B^3 \\ &= C_3^0 A^{3-0} B^0 + C_3^1 A^{3-1} B^1 + C_3^2 A^{3-2} B^2 + C_3^3 A^{3-3} B^3\end{aligned}$$

$$(A + B)^n = C_n^0 A^{n-0} B^0 + C_n^1 A^{n-1} B^1 + C_n^2 A^{n-2} B^2 + \cdots + C_n^n A^{n-n} B^n$$

构成成-败型实验序列的n次实验中，事件A出现的次数X的概率分布为：

$$P(X) = C_n^X \pi^X (1 - \pi)^{n-X}$$

其中 $X=0, 1, 2, \dots, n$ 。

n, π 是二项分布的两个参数。

$$C_n^X = \frac{n!}{X!(n-X)!}$$

对于任何二项分布，总有 $\sum_{x=0}^n P(X) = 1$

例4-2 临床上用针灸治疗某型头疼，有效的概率为60%，现以该疗法治疗3例，其中2例有效的概率是多大？

分析：治疗结果为有效和无效两类，每个患者是否有效不受其他病例的影响，有效概率均为0.6，符合二项分布的条件。

（有效）²（无效）¹



2例有效的概率是0.432

一例以上有效的概率为:

$$\begin{aligned} P(X \geq 1) &= P(1) + P(2) + P(3) \\ &= \frac{3}{1(3-1)!} 0.6(1-0.6)^{3-1} + \frac{3}{2(3-2)!} 0.6^2(1-0.6)^{3-2} \\ &\quad + \frac{3}{3(3-3)!} 0.6^3(1-0.6)^{3-3} = 0.288 + 0.432 + 0.216 \\ &= 0.936 \end{aligned}$$

或

$$\begin{aligned} P(X \geq 1) &= 1 - P(0) \\ &= 1 - 0.064 \end{aligned}$$

(三) 二项分布的特征

1. 二项分布的图形特征

n , π 是二项分布的两个参数, 所以二项分布的形状取决于 n , π 。可以看出, 当 $\pi = 0.5$ 时分布对称, 近似对称分布。当 $\pi \neq 0.5$ 时, 分布呈偏态, 特别是 n 较小时, π 偏离 0.5 越远, 分布的对称性越差, 但只要不接近 1 和 0 时, 随着 n 的增大, 分布逐渐逼近正态。因此, π 或 $1 - \pi$ 不太小, 而 n 足够大, 我们常用正态近似的原理来处理二项分布的问题。

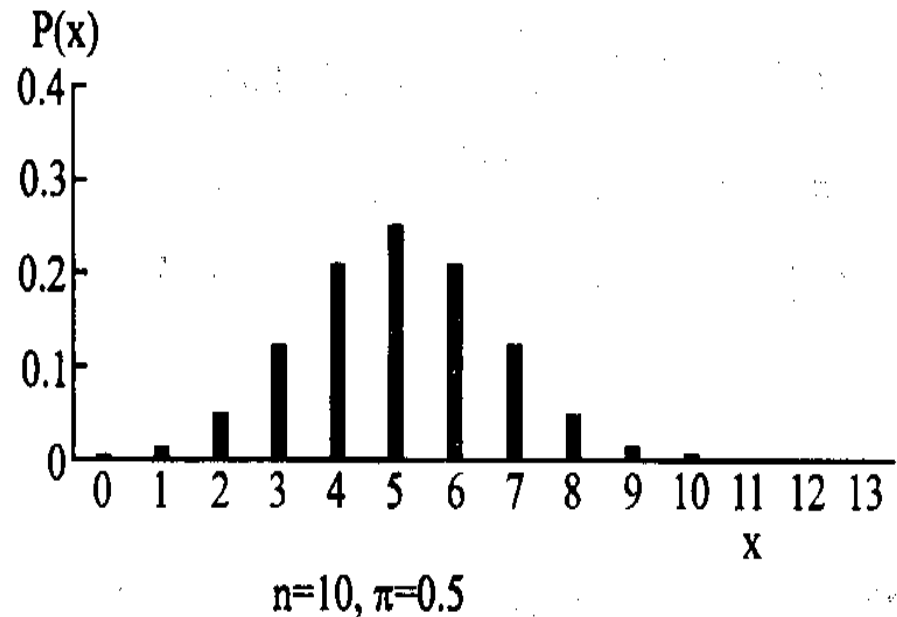
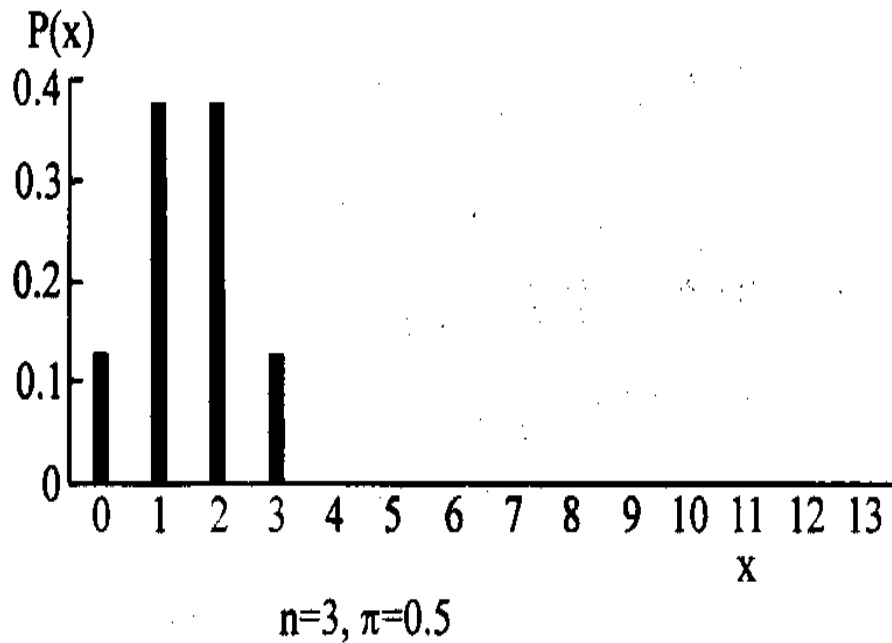


图 4-1 $\pi=0.5$ 时,不同 n 值对应的二项分布

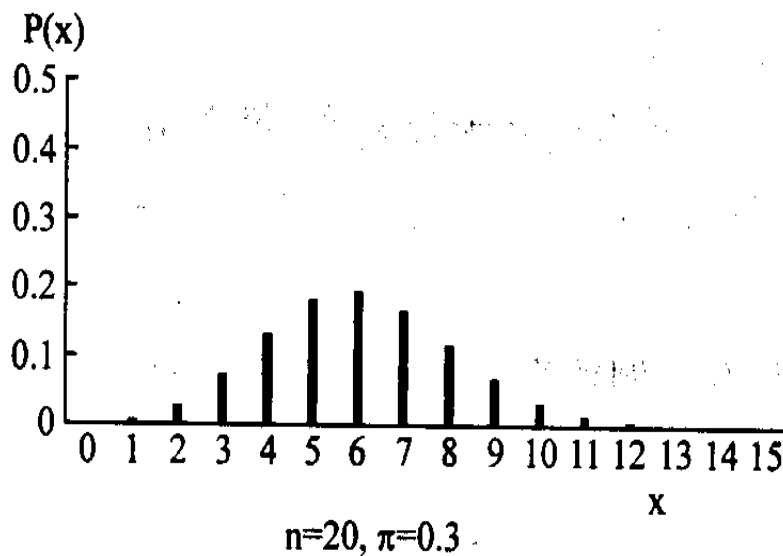
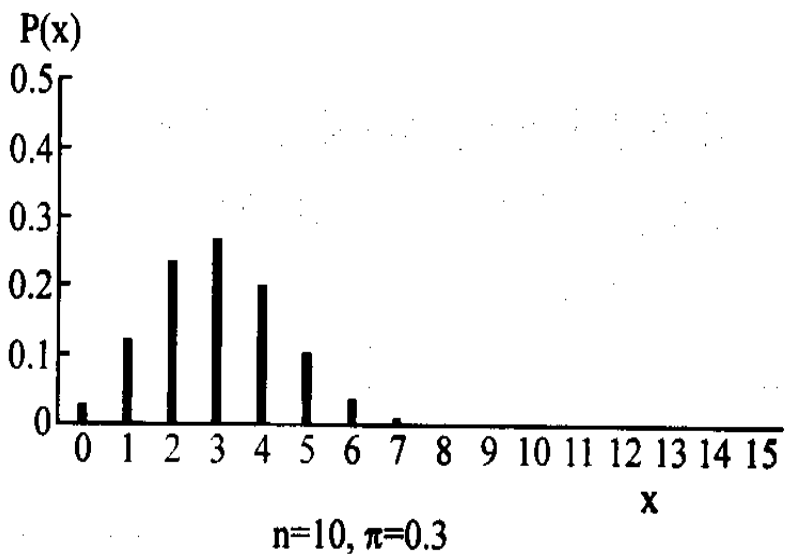
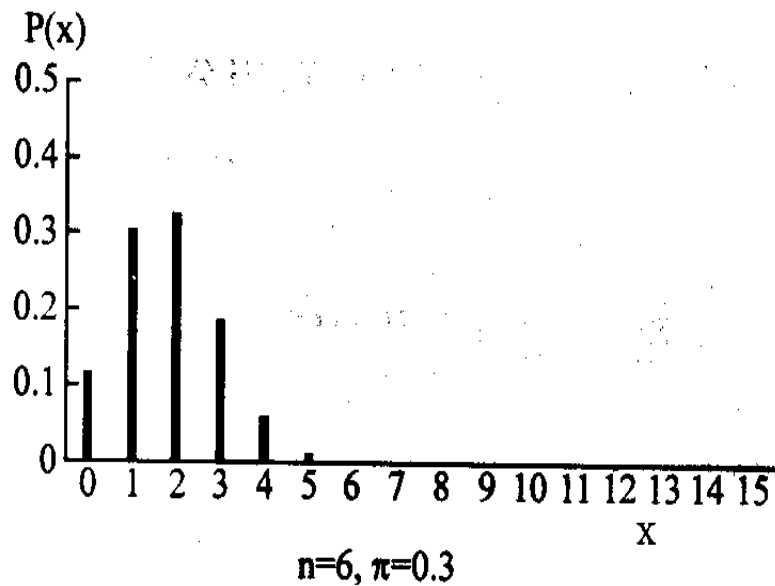
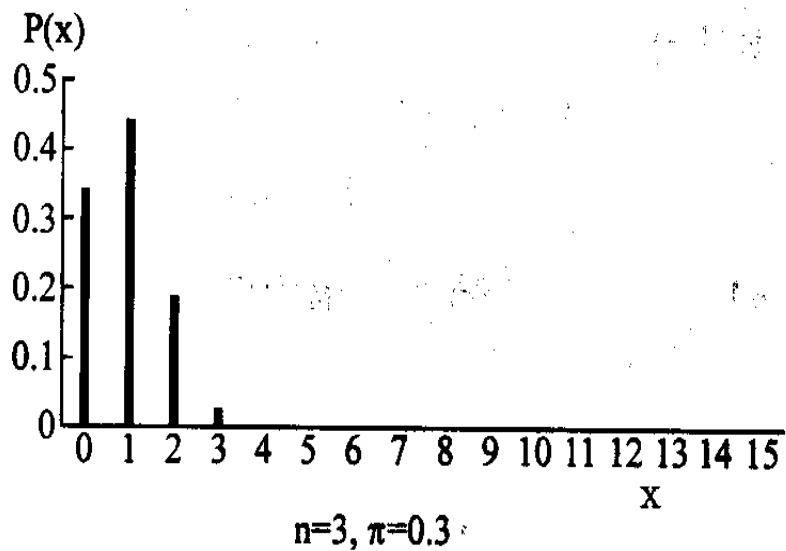


图 4-2 $\pi = 0.3$ 时, 不同 n 值对应的二项分布

2. 二项分布的均数和标准差

对于任何一个二项分布 $B(n, \pi)$ ，如果每次试验出现“阳性”结果的概率均为 π ，则在 n 次独立重复实验中，出现阳性次数

X的总体均数为 $\mu = n\pi$

方差为 $\sigma^2 = n\pi(1 - \pi)$

标准差为 $\sigma = \sqrt{n\pi(1 - \pi)}$

例 实验白鼠3只，白鼠用药后死亡的死亡概率 $\pi=0.6$ ，则3只白鼠中死亡鼠数X的总体均数

$$\mu = n\pi$$

$$= 3 \times 0.6 = 1.8 \text{ (只)}$$

方差为

$$\sigma^2 = n\pi(1-\pi)$$

$$= 3 \times 0.6 \times 0.4 = 0.72$$

标准差为

$$\sigma = \sqrt{n\pi(1-\pi)}$$

$$= \sqrt{0.72} = 0.8485$$

如果以率表示，将阳性结果的频率记为 $p = \frac{X}{n}$ ，则 p 的总体均数 $\mu_p = \pi$

总体方差为 $\sigma_p^2 = \frac{\pi(1-\pi)}{n}$

总体标准差为 $\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$

式中 σ_p 是频率 p 的标准误，反映阳性频率的抽样误差的大小。

例4-4 如果某地钩虫感染率为6.7%,随机观察当地150人,样本钩虫感染率为 p ,求 p 的抽样误差 σ_p 。

$$n=150, \theta=0.067$$

$$\sigma_p = \sqrt{\frac{0.067(1-0.067)}{150}} = 0.021$$

二、二项分布的应用

(一) 概率估计

例4-5 如果某地钩虫感染率为13%，随机观察当地150人，其中有10人感染钩虫的概率有多大？

$$P(X=10) = \binom{150}{10} (0.13)^{10} (0.87)^{140}$$

$$P(X=10) = 0.0055$$

$$= 0.0055$$

(二)单侧累计概率计算

二项分布出现阳性次数至少为 k 次的概率为

$$P(X \geq K) = \sum_{x=k}^n P(X) = \sum_{x=k}^n \frac{n!}{X!(n-X)!} \pi^X (1-\pi)^{n-X}$$

阳性次数至多为 k 次的概率为

$$P(X \leq K) = \sum_{x=0}^k P(X) = \sum_{x=0}^k \frac{n!}{X!(n-X)!} \pi^X (1-\pi)^{n-X}$$

例4-6 如果某地钩虫感染率为13%，随机观察当地150人，其中至多有2人感染钩虫的概率有多大？至少有2人感染钩虫的概率有多大？至少有20人感染钩虫的概率有多大？

至多有2名感染的概率为：

~~$$P(X \leq 2) = P(X=0) + P(X=1) + P(X=2)$$

$$= \binom{150}{0} (0.13)^0 (0.87)^{150} + \binom{150}{1} (0.13)^1 (0.87)^{149} + \binom{150}{2} (0.13)^2 (0.87)^{148}$$~~

$$P(0) + P(1) + P(2)$$

$$= \frac{150}{0! 50} (0.13)^0 (0.87)^{50} + \frac{150}{1! 49} (0.13)^1 (0.87)^{49}$$

$$+ \frac{150}{2! 48} (0.13)^2 (0.87)^{48}$$

$$= 23.10$$

至少有2名感染的概率为:

$$\begin{aligned}
 P(X \geq 2) &= \sum_{x=2}^n P(X=x) = 1 - \sum_{x=0}^1 P(X=x) \\
 &= 1 - [P(0) + P(1)] = 1
 \end{aligned}$$

至少有20名感染的概率为:

$$\begin{aligned}
 P(X \geq 20) &= \sum_{x=20}^n P(X=x) = 1 - \sum_{x=0}^{19} P(X=x) \\
 &= 1 - [P(0) + P(1) + P(2) + \dots + P(19)] \\
 &= 0.4879
 \end{aligned}$$

第二节 Poisson分布的概念与特征

一、Poisson分布的概念

Poisson分布也是一种离散型分布，用以描述**罕见事件**发生次数的概率分布。Poisson分布也可用于研究单位时间内(或单位空间、容积内)某罕见事件发生次数的分布，如分析在单位面积或容积内细菌数的分布，在单位空间中某种昆虫或野生动物数的分布，粉尘在观察容积内的分布，放射性物质在单位时间内放射出质点数的分布等。Poisson分布一般记作 $\Pi(\lambda)$

● Poisson分布作为二项分布的一种极限情况

Poisson分布可以看作是发生的概率 π 很小，而观察例数很大时的二项分布。除要符合二项分布的三个基本条件外，**Poisson分布还要求 π 或 $1-\pi$ 接近于0和1**。有些情况 π 和 n 都难以确定，只能以观察单位(时间、空间、容积、面积)内某种稀有事件的发生数 X 等来表示，如每毫升水中大肠杆菌数，每个观察单位中粉尘的计数，单位时间内放射性质点数等，只要细菌、粉尘、放射性脉冲在观察时间内满足以上条件，就可以近似看为Poisson分布。

二、Poisson分布的特征

1. Poisson分布的概率函数为：

$$P(X) = e^{-\lambda} \frac{\lambda^X}{X!}$$

式中 $\lambda = n\pi$ 为Poisson分布的总体均数，X为观察单位时间内某稀有事件的发生次数；e为自然对数的底，为常数，约等于2.71828。

如某地20年间共出生短肢畸形儿10名，平均每年0.5名。就可用 $\lambda=0.5$ 代入Poisson分布的概率函数来估计该地每年出生此类短肢畸形儿的人数为0, 1, 2...的概率P(X)。

$$P(X) = e^{-\lambda} \frac{\lambda^x}{x!}$$

$$P(X=0) = e^{-0.5} \frac{0.5^0}{0!} = 0.6065$$

$$P(X=1) = e^{-0.5} \frac{0.5^1}{1!} = 0.3033$$

$$P(X=2) = e^{-0.5} \frac{0.5^2}{2!} = 0.0758$$

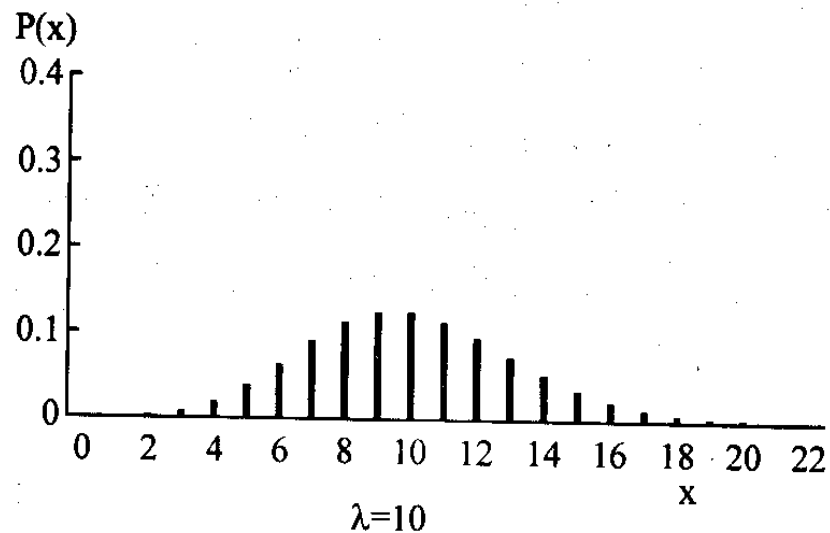
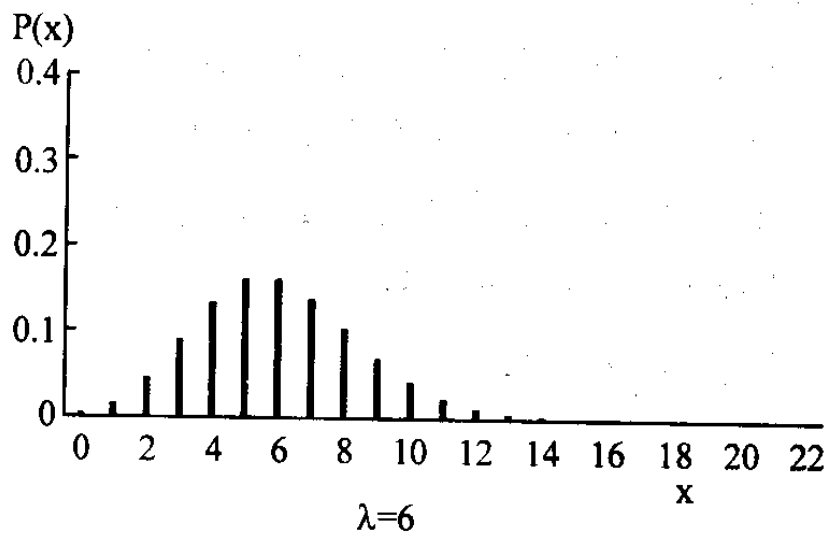
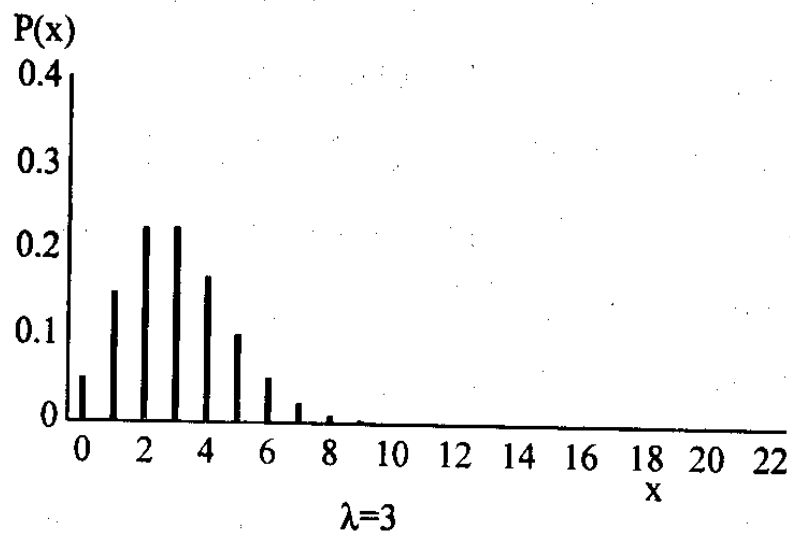
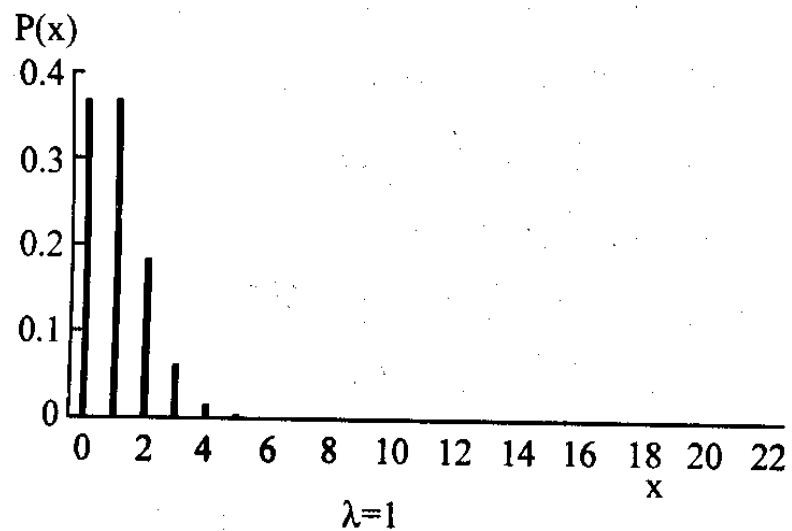


图 4-3 λ 取不同值时的 Poisson 分布图

2.Poisson分布的特性:

- (1) Poisson分布的的**总体均数**与**总体方差**相等, 均为 λ 。
- (2) Poisson分布的观察结果有**可加性**。即对于服从Poisson分布的 **m 个互相独立的随机变量 $X_1, X_2 \dots X_m$** , 它们之和也服从Poisson分布, 其**均数为这 **m 个随机变量的均数之和****。

从总体均数为 λ_1 的服从Poisson分布总体中随机抽出一份样本，其中稀有事件的发生次数为 X_1 ，再独立地从总体均数为 λ_2 的Poisson分布总体中随机抽出另一份样本，其中稀有事件的发生次数为 X_2 ，则他们的合计发生数 $T=X_1+X_2$ 也服从Poisson分布，总体均数为 $\lambda_1+\lambda_2$ 。

Poisson分布的这些性质还可以推广到多个Poisson分布的情形。例如，从同一水源独立地取水样5次，进行细菌培养，每次水样中的菌落数分别为 $X_i, i=1, 2, \dots, 5$ ，均服从Poisson分布，分别记为 $P(\lambda_i), i=1, 2, \dots, 5$ ，把5份水样混合，其合计菌落数 $\sum X_i$ 也服从Poisson分布，记为 $P(\lambda_1 + \lambda_2 + \dots + \lambda_5)$ 其均数为 $(\lambda_1 + \lambda_2 + \dots + \lambda_5)$ 。

医学研究中常利用Poisson分布的可加性，将小的观察单位合并以增大发生次数 X ，以使用正态近似法进行统计推断。

(二)单侧累计概率计算

Poisson分布出现阳性次数至多为K次的概率为

$$P(X \leq K) = \sum_{x=0}^k P(X) = \sum_{x=0}^k e^{-\lambda} \frac{\lambda^x}{x!}$$

阳性次数至少为K次的概率为

$$P(X \geq K) = 1 - P(X \leq k - 1)$$

例4-7 (2) 如果某地新生儿先天性心脏病的发病率为8‰，那么该地120名新生儿中至多有4人患先天性心脏病的概率有多大？至少有5人患先天性心脏病的概率有多大？

至多有4人患先天性心脏病的概率：

$$P(X \leq 4) = \sum_{k=0}^4 \binom{120}{k} (0.008)^k (0.992)^{120-k}$$

$$= \binom{120}{0} (0.008)^0 (0.992)^{120} + \binom{120}{1} (0.008)^1 (0.992)^{119} + \binom{120}{2} (0.008)^2 (0.992)^{118} + \binom{120}{3} (0.008)^3 (0.992)^{117} + \binom{120}{4} (0.008)^4 (0.992)^{116}$$

至少有5人患先天性心脏病的概率

$$P(X \geq 5) = 1 - P(X \leq 4)$$

例4-7 (3) 实验显示某100cm²培养皿平均菌落数为6个，试估计该培养皿菌落数小于3个的概率，大于1个的概率。

该培养皿菌落数小于3个的概率

$$\begin{aligned} P(X < 3) &= \sum_{x=0}^2 P(X=x) = \sum_{x=0}^2 e^{-6} \frac{6^x}{x!} \\ &= P(0) + P(1) + P(2) = 0.062 \end{aligned}$$

该培养皿菌落数大于1个的概率

~~$$P(X > 1) = 1 - P(X=0) - P(X=1) = 1 - e^{-6} - 6e^{-6} = 0.938$$~~

三、二项分布、Poisson分布的正态近似

当 n 很大 π 很小

二项分布

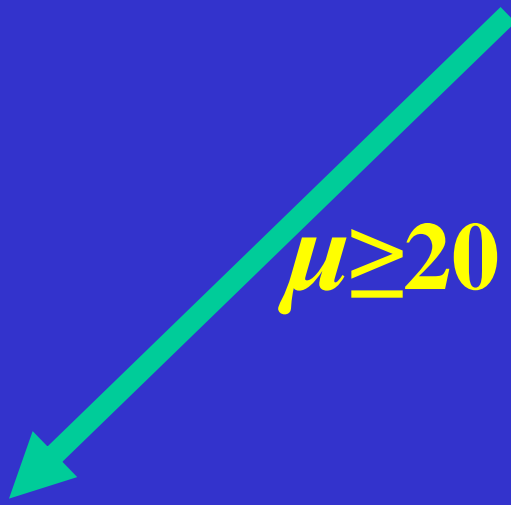
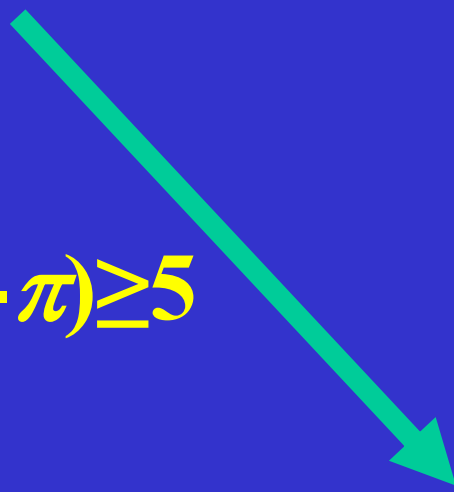
Poisson分布



$n\pi \geq 5,$
且 $n(1-\pi) \geq 5$

$\mu \geq 20$

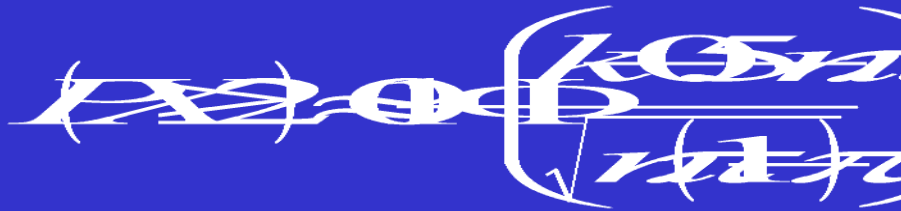
正态分布



例4-14 如果某地钩虫感染率为13%，随机观察当地150人，其中至少有20人感染钩虫的概率有多大？

$$n\pi = 150 \times 0.13 = 19.5$$

$$n(1-\pi) = 150 \times (1-0.13) = 130.5$$



至少有20人感染钩虫的概率为50%。

小 结

1. 正态分布和标准正态分布
2. 实际资料的正态性判断
3. 医学参考值范围
4. 二项分布的应用条件及特征参数
5. Poisson分布的应用条件及特征参数