第2章　一维随机变量及其分布

- 离散型随机变量及其分布律
- 连续型随机变量及其概率密度
- 随机变量函数的分布

1

第一节

随机变量的定义

2

随机变量

Random Variable

在前面的学习中,我们用字母A、B、C...表示事件，并视之为样本空间Ω的子集；针对等可能概型，主要研究了用排列组合手段计算事件的概率。

本章，将用随机变量表示随机事件，以便采用精确的数学方法描述、研究随机现象。

3

随机变量Random Variable

■目的

将样本空间数量化,即用数值来表示试验的结果

■ 有些随机试验的结果可直接用数值来表示.

例如: 在掷骰子试验中,结果可用**1,2,3,4,5,6**来表示

■ 有些随机试验的结果表面上与数值无关，但我们仍可将其数量化

例如: 掷硬币试验,其结果是用汉字“正面”和“反面”来表示的

可规定: 用 **1**表示 “正面朝上” 用 **0** 表示“反面朝上”

4

试验结果的数量化

例　设箱中有10个球，其中有2个红球，8个白球；从中任意抽取2个，观察抽球结果。

取球结果为：两个白球；两个红球；一红一白

如果用X表示取得的红球数，则X的取值可为0，1，2。

此时，“两只红球” = “X取到值2”，可记为 {X=2}

“一红一白”记为 {X=1}，

“两只白球”记为 {X=0}

特点：试验结果数量化了，试验结果与实数建立了对应关系

5

随机变量的定义

- 随机变量

设随机试验的样本空间为Ω，如果对于每一个样本点 $\omega\in\Omega$，均有唯一的实数 $X(\omega)$ 与之对应，称 $X=X(\omega)$ 为样本空间Ω上的随机变量。

- 随机变量的两个特征：

1) 它的取值在试验之前是不明确的，只有在做了试验之后才能确定取值

2）随机变量在某一范围内取值，表示一个随机事件

6

随机变量的实例

■ 例

➢某个灯泡的使用寿命X。

X 的可能取值为 $[0,+\infty)$

➢某电话总机在一分钟内收到的呼叫次数Y.

Y 的可能取值为 0，1，2，3，...，

➢在[0，1]区间上随机取点，该点的坐标X.

X 的可能取值为 [0，1]上的全体实数。

7

用随机变量表示事件

■ 若X是随机试验E的一个随机变量，$S\subset R$，那么

$\{X\in S\}$可表示E中的事件

■ E中的事件通常都可以用X的不同取值来表示.

如在掷骰子试验中，用X表示出现的点数, 则

“出现偶数点”可表示为：$\{X=2\}\cup\{X=4\}\cup\{X=6\}$

“出现的点数小于 4 ”可表示为：$\{X<4\}$或$\{X\le3\}$

8

第二节

随机变量的分布函数

随机变量的分布函数 Distribution Function

- 分布函数的定义

设X为一随机变量, 则对任意实数 x ,$\{X \le x\}$ 是一个随机事件，称

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

为随机变量X的分布函数。

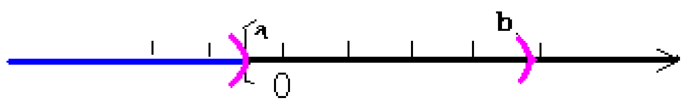
定义域为 （－∞，＋∞）； 值域为 ［０，１］。

分布函数表示事件的概率

引进分布函数F(x)后，事件的概率就可以用F(x)的函数值来表示。

- **P{X≤b}=F(b)**
- **P{X>b}=1-P{X≤b}=1-F(b)**
- **P{a<X≤b}=F(b)-F(a)**

P{a<X≤b}=P{X≤b}-P{X≤a}=F(b)-F(a)



11

例 在抛硬币的试验中，已知随机变量X的可能取值是0和1，它们的概率都是0.5，求X的分布函数

$$F(x)=P\{X\le x\}=\begin{cases}0 & x<0\\ 0.5 & 0\le x<1\\ 1 & 1\le x\end{cases}$$

12

分布函数的性质

- F(x)是单调不减函数

  若 $x_1 < x_2$ $\quad F(x_1) \le F(x_2)$

- F(x)处处右连续 $F(x^+) = F(x)$

- 0≤ F(x) ≤1, 且

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

$F(-\infty) = P\{X < -\infty\} \longrightarrow$ 不可能事件

$F(+\infty) = P\{X < +\infty\} \longrightarrow$ 必然事件

13

第三节

离散型随机变量

14

随机变量的类型

- **离散型**

  随机变量的所有取值是有限个或可数无限个

- 非离散型

  随机变量的取值有无穷多个，且不可列

其中连续型随机变量是最常见的一种重要类型

15

离散随机变量的概率分布

设离散型随机变量 X 的所有可能取值是

$x_1, x_2, \cdots, x_n, \cdots$，而取值 x_k 的概率为 p_k

即 $$\boxed{P\{X = x_k\} = p_k \quad k = 1, 2, 3, \cdots}$$

称此式为X的分布律或概率分布

16

离散随机变量分布律的表示法

- 公式法 $P\{X=x_k\}=p_k \quad k=1,2,3,\cdots$
- 概率分布表

X	x_1	x_2	$\cdots$	x_k	$\cdots$
P	p_1	p_2	$\cdots$	p_K	$\cdots$

随机变量X的概率分布全面表达了X的所有可能取值以及取各个值的概率情况

性质 1) $p_k \geq 0 \quad k=1,2,\cdots$

2) $\sum_k p_k = 1$

17

分布律确定概率

例 设X的分布律为

X	-1	1	2
P	**1/3**	**1/2**	**1/6**

求 **P{0<X≤2}**

解 **P{0<X≤2}=P{X=1}+P{X=2}**

=1/2+1/6=2/3

18

求分布律举例

例1 设有一批产品20件，其中有3件次品，从中任意抽取2件，如果用X表示取得的次品数，求随机变量X的分布律及事件“至少抽得一件次品”的概率。

解：X的可能取值为 0，1，2

$$P\{X=0\}=\frac{C_{17}^{2}}{C_{20}^{2}}=\frac{136}{190}$$

$$P\{X=1\}=\frac{C_{3}^{1}C_{17}^{1}}{C_{20}^{2}}=\frac{51}{190}$$

$$P\{X=2\}=\frac{C_{3}^{2}}{C_{20}^{2}}=\frac{3}{190}$$

19

故X的分布律为

X	0	1	2
P	$\frac{136}{190}$	$\frac{51}{190}$	$\frac{3}{190}$

而“至少抽得一件次品”$=\{X\geqslant 1\}=\{X=1\}\cup\{X=2\}$

注意：{X=1}与{X=2}是互不相容的！

故 $P\{X\geqslant 1\}=P\{X=1\}+P\{X=2\}=\frac{51}{190}+\frac{3}{190}=\frac{54}{190}=\frac{27}{95}$

思考：试求X的分布函数.

20

X	0	1	2
P	$\frac{136}{190}$	$\frac{51}{190}$	$\frac{3}{190}$

思考：试求X的分布函数.

$$F(x)=P\{X\le x\}=\begin{cases}0 & x<0\\ 136/190 & 0\le x<1\\ 187/190 & 1\le x<2\\ 1 & 2\le x\end{cases}$$

例2 从一批次品率为p的产品中，有放回抽样直到抽到次品为止。求抽到次品时，已抽取的次数X的分布律。

解 记A_i="第i次取到正品"，i=1, 2, 3, …

则 A_i，i=1, 2, 3, … 是相互独立的！

X的所有可能取值为 1，2，3，…，k, …

由于{X=k}对应着事件$A_1A_2\cdots A_{k-1}\overline{A_k}$

$$P\{X=k\}=P(A_1A_2\cdots A_{k-1}\overline{A_k})=(1-p)^{k-1}p,\ k=1,2,\ldots$$

例3 设随机变量X的分布律为

$$P\{X=k\}=b(\frac{2}{3})^k, k=1,2,3,\cdots$$

试确定常数b.

解 由分布律的性质,有

$$\sum_{k=1}^{\infty}P\{X=k\}=\sum_{k=1}^{\infty}b(\frac{2}{3})^k=\frac{b\frac{2}{3}}{1-\frac{2}{3}}$$

$$=b\frac{2/3}{1/3}=2b=1 \Longrightarrow b=\frac{1}{2}.$$

常见的离散型分布

1. 两点分布

△定义: 若随机变量X的分布律为:

X	0	1
P	1−p	p

则称**X**服从参数为***p*** 的两点分布

△由定义可知:样本空间只有两个样本点的情况都可以用两点分布来 描述。

如:上抛一枚硬币。

例 设一个袋中装有3个红球和7个白球，现在从中随机抽取一球，如果每个球抽取的机会相等，并且用数“1”代表取得红球，“0”代表取得白球，则随机抽取一球所得的值是一个离散型随机变量

$$X=\begin{cases}1 & （取得红球）\\ 0 & （取得白球）\end{cases}$$

其概率分布为 $P\{X=1\}=\dfrac{3}{10}$ $P\{X=0\}=\dfrac{7}{10}$

即X服从两点分布。

25

2. 二项分布

将试验E重复进行n次,若各次试验的结果互不影响,则称这n次试验是相互独立的.

设随机试验E只有两种可能的结果:A及$\bar{A}$，且P(A)=p,在相同的条件下重复进行n次独立试验E,则这一串试验为n重伯努利试验，简称伯努利试验(Bernolli trials).

设在一次试验中事件A发生的概率为 **p (0<p<1)**，则A在**n**次贝努里试验中恰好发生 k次的概率为

$$P_n(k)=C_n^k p^k q^{n-k} \quad （k=0，1，2，…，n）$$

26

2. 二项分布

- 随机变量X的分布律

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

$$k=0,1,2\ldots,n;$$

其中0< p <1，则称X服从参数为 n, p 的二项分布(也称 Bernolli 分布）,记为

X～B(n, p)

- 在n重贝努利试验中, 若以X表示事件A发生的次数, 则X～B(n, p)

27

俗语中的概率论

俗话说，三个臭皮匠顶个诸葛亮。意思是说，三个普通的人智慧合起来能和一个智者相当。

智者千虑，必有一失；愚者千虑，必有一得。意思是聪明的人在上千次考虑中，总会有一次失误；愚蠢的人在上千次考虑中，总会有一次收获。

设有n个普通人，每人解决问题的概率都为0.001，要想解决问题的概率超过0.999，n至少要是多少？

n≥6905

28

3. 泊松分布

■ 定义

若随机变量 X 的分布律为:

$$P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda},k=0,1,2\ldots$$

其中λ >0，则称X服从参数为λ的泊松分布

X～P(λ)

实际问题中，有些随机变量X，Y是服从或近似服从Poisson分布的

- 服务台在某时间段内接待的服务次数X；
- 交换台在某时间段内接到呼叫的次数Y；
- 矿井在某段时间发生事故的次数；
- 显微镜下相同大小的方格内微生物的数目；
- 单位体积空气中含有某种微粒的数目

体积相对小的物质在较大的空间内的稀疏分布，都可以看作泊松分布，其参数 λ 可以由观测值的平均值求出。

例 已知某电话交换台每分钟接到的呼唤次数X服从 $\lambda=4$ 的泊松分布，分别 求（1）每分钟内恰好接到3次呼唤的概率；（2）每分钟不超过4次的概率

解 $$P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda}$$

$\lambda=4, k=3$

$$P\{X=3\}=\frac{4^3}{3!}e^{-4}\approx 0.196$$

$$P\{X\le 4\}=P\{X=0\}+P\{X=1\}+P\{X=2\}+P\{X=3\}+P\{X=4\}\approx 0.629$$

31

二项分布的泊松近似

The Poisson Approximation to the Binomial Distribution

泊松定理

$$C_n^k p^k(1-p)^{n-k}\approx\frac{\lambda^k}{k!}e^{-\lambda}$$

$$\lambda=np$$

实际应用中：当n较大, p较小，np适中时，即可用泊松公式近似替换二项分布的概率公式

32

例　某人骑摩托车上街,出事故率为0.02，独立重复上街200次，求出事故至少两次的概率.

解　200次上街⇔200重Bernolli实验

记X为出事故的次数，则 $X \sim B(200, 0.02)$

$$P\{x=k\}=C_{200}^{k}(0.02)^{k}(0.98)^{200-k} \approx \frac{4^{k}}{k!}e^{-4}$$

P{X≥2}=1-P{X=0}-P{X=1}≈0.908

=1-0.98 ^{200}-200（0.02)(0.98 ^{199})

≈0.911　　泊松定理

- 结果表明，随着实验次数的增多，小概率事件总会发生的！

4. 几何分布

在独立重复试验中，设事件A在每次实验中发生的概率为p，记X为A首次发生时的实验次数，则X的分布律为

$$P\{X=k\}=(1-p)^{k-1}p \quad k=1,2,\cdots$$

这种分布称为几何分布，p为参数，记为

$$X \sim Ge(p)$$

5. 负二项分布

在独立重复试验中，设事件A在每次实验中发生的概率为p，记X为A第r次发生时的实验次数，则X的分布律为

$$P\{X=k\}=C_{k-1}^{r-1}p^{r}(1-p)^{k-r}\quad k=r,r+1,\cdots$$

这种分布称为**负二项分布**，r和p为参数，记为

$$X\sim Nb(r,p)$$

35

6. 超几何分布

设有N件产品，其中有M件不合格品，N-M件合格品。若从中不放回地随机抽取n件，记X为其中不合格品的件数，则X的分布律为

$$P\{X=k\}=\frac{C_{M}^{k}C_{N-M}^{n-k}}{C_{N}^{n}}\quad k=1,2,\cdots$$

这种分布称为**超几何分布**，n、N和M为参数，记为

$$X\sim h(n,N,M)$$

36

课后练习

P55 习题2
1--9

作业

1, 7

第四节

连续型随机变量

概率密度函数

Probability density function p.d.f.

■ 定义 设X为一随机变量，F(x)为其分布函数，若存在非负可积函数 $f(x)$，使对任意实数 x，有

$$F(x)=\int_{-\infty}^{x} f(u)du$$

则称X为连续型随机变量，$f(x)$ 称为X 的概率密度函数,简称密度函数或密度.

39

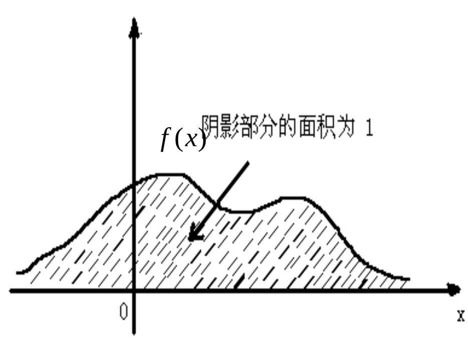
概率密度函数的性质

■ 非负性

$$f(x)\geq 0,\ \ \forall x\in(-\infty,+\infty)$$

■ 规范性

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

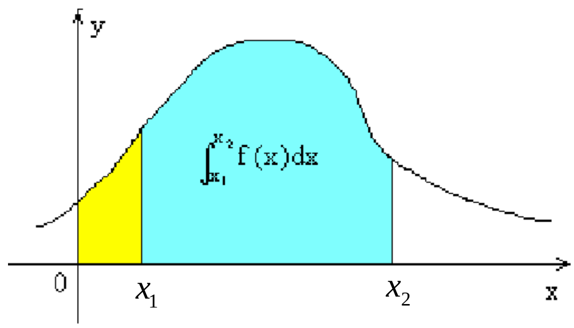


$$P\{-\infty<x<+\infty\}=1$$

40

- 密度函数在区间上的积分 =
  随机变量在区间上取值的概率

$$P\{x_1 < X \le x_2\} = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x)dx$$



41

密度函数和分布函数的关系

- 积分关系

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

$$F(x) = P\{X \le x\} = \int_{-\infty}^{x} f(t)dt$$

- 导数关系

若$f(x)$在x处连续，则$F'(x) = f(x)$

42

连续型随机变量的分布函数的性质

连续型随机变量的分布函数在实数域内处处连续

连续型随机变量取任意指定实数值a的概率为0

$$P\{X=a\}=0$$

$$P\{a \le X< b\}= P\{a<X\le b\}=P\{a \le X \le b\}=P\{a<X<b\}$$

$$=\int_a^b f(x)dx$$

X取值在某区间的概率等于密度函数在此区间上的定积分

43

例：已知密度函数求概率

随机变量 X 的概率密度为

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

$$f(x)=\begin{cases} a\cos x & |x|\le \frac{\pi}{2} \\ 0 & \text{其它} \end{cases}$$

求 $P\{0\le X\le \frac{\pi}{4}\}$

解　（先利用密度函数的性质求出 a）

$$\int_{-\infty}^{+\infty} f(x)dx=\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} a\cos x dx=1 \quad \Rightarrow \quad a=\frac{1}{2}$$

（密度函数在区间的积分得到此区间的概率）

$$P\{0\le X\le \frac{\pi}{4}\}=\int_0^{\frac{\pi}{4}} \frac{1}{2}\cos x dx=\frac{\sqrt{2}}{4}$$

44

例：已知分布函数求密度函数

随机变量 X 的分布函数为

$$F(x)=\begin{cases}0 & x\le 0\\ x^2 & 0<x\le 1\\ 1 & x>1\end{cases}$$

（1）求 $P\{0.3<X<0.7\}$

（2）X 的密度函数

解

（1）$P\{0.3<X<0.7\}=F(0.7)-F(0.3)=0.7^2-0.3^2=0.4$

（2）密度函数为

$$f(x)=F'(x)=\begin{cases}2x & 0<x\le 1\\ 0 & \text{其他}\end{cases}$$

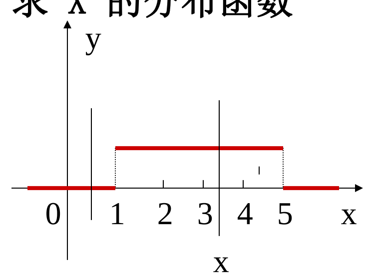
45

例：已知密度函数求分布函数

已知连续型随机变量X的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{4} & (1,5)\\ 0 & \text{其它}\end{cases}$$

求 X 的分布函数



解 当 $x<1$ 时

$$F(x)=\int_{-\infty}^{x} f(x)dx=0$$

当 $1\le x<5$ 时

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

$$=0+\int_{1}^{x}\frac{1}{4}dx=\frac{1}{4}(x-1)$$

46

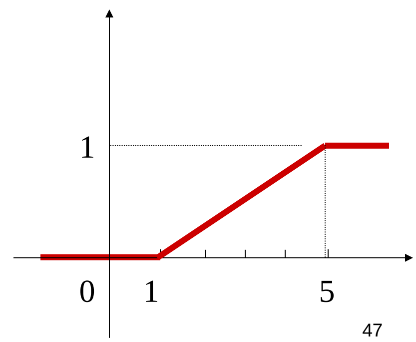
当 **x>5** 时 $F(x)=\int_{-\infty}^{x} f(x)dx=$

$$=\int_{-\infty}^{1} f(x)dx+\int_{1}^{5} f(x)dx+\int_{5}^{x} f(x)dx$$

$$=0+\int_{1}^{5}\frac{1}{4}dx+0=\frac{1}{4}(5-1)=1$$

所以

$$F(x)=\begin{cases}0 & x\le 1\\ \frac{1}{4}(x-1) & 1<x\le 5\\ 1 & x>5\end{cases}$$



练一练

已知连续型随机变量X的概率密度为

$$f(x)=Ae^{-|x|}$$

(1) 求 $P\{-1<X<1\}$

(2) 求 $P\{|X|\ge 5\}$

练一练

随机变量 X 的分布函数为

$$F(x)=\begin{cases}\frac{1}{2}e^{x} & x<0\\ \frac{1}{2} & 0\le x<1\\ 1-\frac{1}{2}e^{-(x-1)} & 1\le x\end{cases}$$

(1) 求 $P\{-1<X<2\}$　　(2) 求X的密度函数

1. 均匀分布

■ 定义　　若连续型随机变量X的密度函数为

$$f(x)=\begin{cases}\dfrac{1}{b-a} & a\le x\le b\\ 0 & \text{其他}\end{cases}$$

则称X在区间 [a，b]上服从均匀分布．记为 X ~ U[a, b]

■ 分布函数

$$F(x)=\begin{cases}0, & x<a\\ \dfrac{x-a}{b-a}, & a\le x<b\\ 1, & b\le x\end{cases}$$

■ 意义

X“等可能”地取区间[a,b]中的值，这里的“等可能”理解为：X落在区间[a,b]中任意等长度的子区间内的可能性是相同的。或者说它落在子区间内的概率只依赖于子区间的长度而与子区间的位置无关。

$$P\{c < X \le d\} = \int_c^d f(x)dx$$

$$= \int_c^d \frac{1}{b-a}dx = \frac{d-c}{b-a}$$

思考 设ξ在[-1，5]上服从均匀分布，求方程

$$x^2 + 2\xi x + 1 = 0$$

有实根的概率。

解 方程有实数根 $\Longleftrightarrow$ $4\xi^2 - 4 \ge 0$

即 $|\xi| \ge 1$

而 ξ 的密度函数为 $f(x) = \begin{cases} \frac{1}{6} & (-1 \le x \le 5) \\ 0 & \text{其它} \end{cases}$

所求概率为 $P\{|\xi| \ge 1\} = \int_{-\infty}^{-1} f(x)dx + \int_1^{+\infty} f(x)dx = \frac{2}{3}$

2. 指数分布

■ 定义　若连续型随机变量X具有密度函数

$$f(x)=\begin{cases}0 & x\le 0\\ \lambda e^{-\lambda x} & 0<x\end{cases}\quad(\lambda>0\text{为常数})$$

则称X服从参数为 λ 的指数分布.

$$X\sim E(\lambda)$$

■ 分布函数

$$F(x)=\begin{cases}0 & x<0\\ 1-e^{-\lambda x} & 0\le x\end{cases}$$

53

例　设X服从参数为3的指数分布，求它的密度函数

及 $P\{X\ge 1\}$ 和 $P\{-1<X\le 2\}$

解　X的概率密度 $f(x)=\begin{cases}0 & x\le 0\\ 3e^{-3x} & 0<x\end{cases}$

$$P\{x_1<X\le x_2\}=\int_{x_1}^{x_2} f(x)dx$$

$$P\{X\ge 1\}=\int_1^{+\infty} f(x)dx=\int_1^{+\infty} 3e^{-3x}dx=e^{-3}$$

$$P\{-1<X\le 2\}=\int_0^2 3e^{-3x}dx=1-e^{-6}$$

54

例 设X服从指数分布，$s, t > 0$求

$$P\{X > s+t \mid X > s\} \qquad P\{X > t\}$$

解 X的概率密度 $f(x) = \begin{cases} 0 & x \le 0 \\ \lambda e^{-\lambda x} & 0 < x \end{cases}$

$$P\{X > t\} = \int_t^{+\infty} \lambda e^{-\lambda x} dx = -e^{-\lambda x}\Big|_t^{+\infty} = e^{-\lambda t}$$

$$P\{X > s+t \mid X > s\} = \frac{P\{(X > s+t) \cap (X > s)\}}{P\{X > s\}}$$

$$= \frac{P\{X > s+t\}}{P\{X > s\}} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t}$$

55

3. 正态分布

- 若连续型随机变量X的密度函数为

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

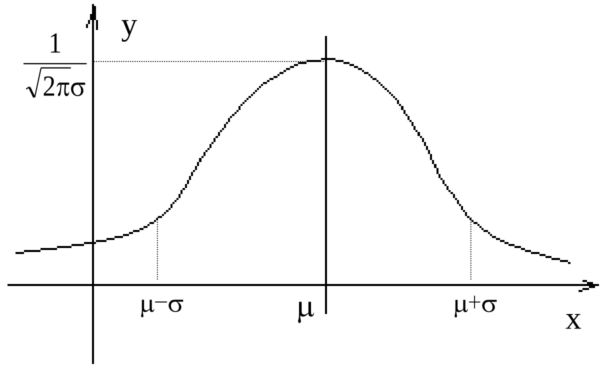
$\mu, \sigma (> 0)$为常数

则称X服从参数为 μ, σ^2 正态分布,记为

$$X \sim N(\mu, \sigma^2)$$

56

正态分布的密度函数的性质与图形



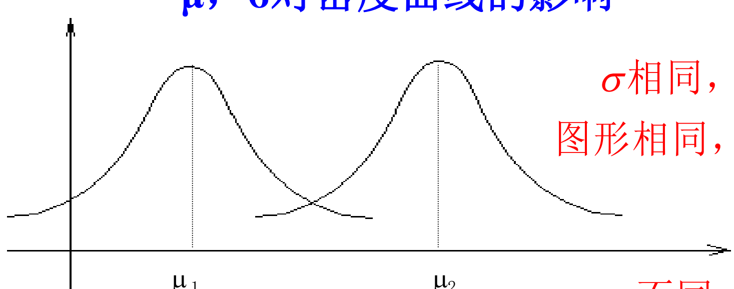
中间高

两边低

- **对称性** 关于 $x=\mu$ 对称，并在 $x=\mu$ 达到最大值
- **单调性** （− ∞，μ）升，（μ，+∞ ）降
- **拐点** $\left(\mu\pm\sigma,\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}}\right)$; $f_{最大}(\mu)=\frac{1}{\sqrt{2\pi}\sigma}$

57

μ，σ对密度曲线的影响



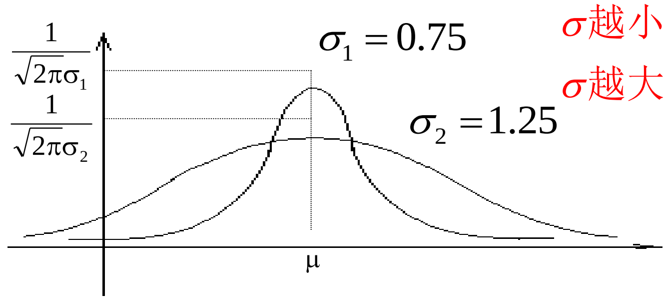
σ相同，μ不同

图形相同，位置平移

σ不同，μ相同

σ越小，图形越陡；

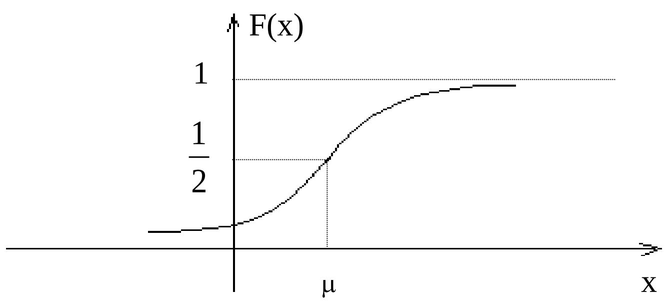
σ越大，图形越平缓



58

正态分布的分布函数

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



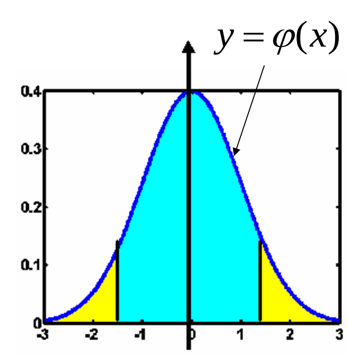
59

标准正态分布
Standard Normal distribution

■ 定义 U ~ N（0，1^2）分布称为标准正态分布

■ 密度函数

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$ 偶函数



■ 分布函数

$$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$$

$\mu=0 \quad \sigma=1$

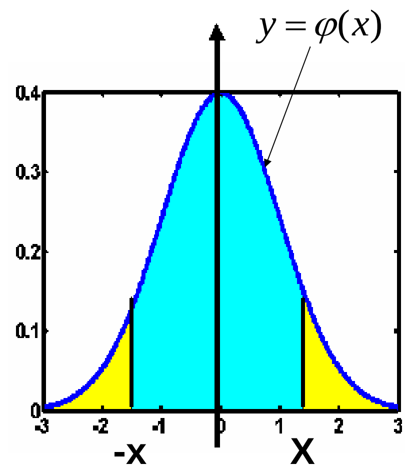
60

标准正态分布的概率计算

■ 分布函数

$$\Phi(x) = P\{U \le x\}$$

$$= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$



$$\boxed{\Phi(-x) = 1 - \Phi(x)}$$

$$\Phi(0) = 0.5$$

61

标准正态分布的概率计算

■ 公式

$$P\{a \le U \le b\} = \Phi(b) - \Phi(a)$$

$$P\{U \le b\} = \Phi(b) \qquad P\{U \ge a\} = 1 - \Phi(a)$$

■ 查表　$x \ge 0$ 时，$\Phi(x)$的值可以查表

$x < 0$ 时，$\Phi(-x) = 1 - \Phi(x)$

■ 例　$U \sim N(0,1)$

$$P\{1 \le U \le 2\} = \Phi(2) - \Phi(1) = 0.9772 - 0.8413 = 0.1359$$

$$P\{U \le -1\} = \Phi(-1) = 1 - \Phi(1) = 1 - 0.8413 = 0.1587$$

$$P\{|U| \le 1\} = \Phi(1) - \Phi(-1) = 2\Phi(1) - 1 = 0.6826$$

62

一般正态分布的标准化

- 定理

如果 $X \sim N(\mu,\sigma^2)$, 则 $F(x)=\Phi\left(\dfrac{x-\mu}{\sigma}\right)$

- 概率计算

若 $X \sim N(\mu,\sigma^2)$

$$P\{a \le X \le b\} = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

查标准正态
分布表

63

一般正态分布的概率计算

$X \sim N(\mu,\sigma^2)$ $\Phi(x)$为标准正态分布函数

$$P\{a < X < b\} = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$P\{X < b\} = \Phi\left(\frac{b-\mu}{\sigma}\right)$$

$$P\{X > a\} = 1 - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

64

例 设X~N（1，4），求 P{0<X<1.6}

解 $\mu=1,\ \sigma=2$

$$P\{0<X<1.6\}=\Phi(\frac{1.6-1}{2})-\Phi(\frac{0-1}{2})$$
$$=\Phi(0.3)-\Phi(-0.5)$$
$$=\Phi(0.3)-[1-\Phi(0.5)]$$
$$=0.6179-1+0.6915=0.3094$$

$$P\{a<X<b\}=\Phi(\frac{b-\mu}{\sigma})-\Phi(\frac{a-\mu}{\sigma})$$

65

正态分布的实际应用

某单位招聘155人，按考试成绩录用，共有526人报名，假设报名者的考试成绩 $X\sim N(\mu,\sigma^2)$

已知90分以上的12人，60分以下的83人，若从高分到低分依次录取，某人成绩为78分，问此人能否被录取？

- 分析

首先求出 μ 和 σ

然后根据录取率或者分数线确定能否录取

66

解 成绩X服从 $N\left(\mu,\sigma^2\right)$

$$P\left\{X>90\right\}=\frac{12}{526}\approx 0.0228 \qquad P\left\{X<60\right\}=\frac{83}{526}\approx 0.1588$$

$$0.9772=1-0.0228\approx P\{X\le 90\}=F(90)=\Phi\left(\frac{90-\mu}{\sigma}\right)$$

查表得 $\frac{90-\mu}{\sigma}\approx 2.0$

$P\left\{X<60\right\}=\Phi\left(\frac{60-\mu}{\sigma}\right)\approx 0.1588$ 不能直接在表中查出

所以 $\Phi\left(\frac{\mu-60}{\sigma}\right)\approx 1-0.1588=0.8412$

查表得 $\frac{\mu-60}{\sigma}\approx 1.0$

解得 $\mu=70$ ， $\sigma=10$

67

故 $X\sim N\left(70,10^2\right)$

方法1 设录取的最低分为 x 录取率为 $\frac{155}{526}\approx 0.2947$

则应有 $P\left\{X\ge x\right\}=0.2947$

$$\Phi\left(\frac{x-70}{10}\right)=P\left\{X<x\right\}\approx 1-0.2947=0.7053$$

所以 $\frac{x-70}{10}\approx 0.54$ $\quad x=75.4$

此人78分，可被录取。

方法2 $P\{X\ge 78\}=1-P\{X<78\}=1-\Phi\left(\frac{78-70}{10}\right)$

$=1-\Phi(0.8)\approx 1-0.7881=0.2119 \quad <0.2947$

因此此人可被录取.

68

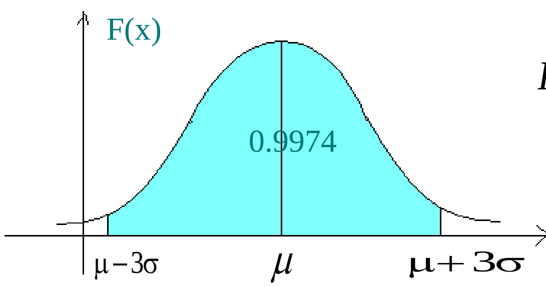
3σ准则

$$X \sim N(\mu,\sigma^2)$$

X的取值几乎都落入以μ为中心，以3σ为半径的区间内。这是因为：

$$P\{\mu-3\sigma \le X \le \mu+3\sigma\} = \Phi(3)-\Phi(-3)$$

$$= \Phi(3)-[1-\Phi(3)] = 2\Phi(3)-1 = 0.9974$$



$$P\{|X-\mu|>3\sigma\} < 0.3\%$$

是小概率事件

69

4. 伽马分布

若随机变量X的密度函数为

$$f(x)=\begin{cases} \dfrac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} & x>0 \\ 0 & x \le 0 \end{cases}$$

则称X服从伽马分布，记为 $X \sim Ga(\alpha,\lambda)$

其中$\alpha>0$为形状参数，$\lambda>0$为尺度参数，$\Gamma(\alpha)$为伽马函数，其定义为

$$\Gamma(\alpha)=\int_0^\infty x^{\alpha-1}e^{-x}dx$$

性质：$(1)\,\Gamma(1)=1,\ \Gamma\left(\dfrac{1}{2}\right)=\sqrt{\pi}$　$(2)\,\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$

70

5. 贝塔分布

若随机变量X的密度函数为

$$f(x)=\begin{cases}\dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1} & 0<x<1\\ 0 & else\end{cases}$$

则称X服从**贝塔分布**，记为 $X \sim Be(a,b)$

其中 $a,b>0$ 都为形状参数.

课后练习

P55 习题2
10--22

作业

15，21

如何查"标准正态分布函数值表"

例　求 $\Phi(0.64)$

$$0.64 = 0.6 + 0.04$$

u	…	0.04	…
…	…	…	…
0.6	…	0.7389	…
…	…	…	…

因此 $\Phi(0.64) = 0.7389$

返回

第五节

一维随机变量函数的分布

随机变量的函数的分布

■ 背景

在许多实际问题中,常常需要研究随机变量的函数的分布问题, 例:

☆ 测量圆轴截面的直径d,而关心的却是截面积:

$$S=\frac{1}{4}\pi d^2$$

d为随机变量, S 就是随机变量d的函数。

☆ 在统计物理中,已知分子的运动速度x的分布,求其动能:

$$y=\frac{1}{2}mx^2$$ 的分布。

一般地，设X是一个随机变量，y=g(x)是定义在X的取值域上的实值连续函数，则称Y=g(X) 为随机变量X的函数。

75

例 设随机变量X的分布律为

X	−1	0	1	2
p_k	0.2	0.3	0.4	0.1

求Y=2X² +1的分布律.

分析 1. Y为离散型随机变量;

2. Y的可能取值为1,3,9;

3. {Y=1}={X=0}

所以P{Y=1}=P{X=0}=0.3

{Y=3}={X=-1}U{X=1}

所以P{Y=3}=P{X=-1}+P{X=1}=0.2+0.4=0.6

{Y=9}={X=2}

所以P{Y=9}=P{X=2}=0.1

所以，$y=2x^2+1$ 的分布律为

y	1	3	9
p_k	0.3	0.6	0.1

76

离散随机变量的函数的分布

若X为离散型 随机变量, 其分布律为

X	x_1	x_2	x_3	. . . .	x_n	. . . .
p_k	p_1	p_2	p_3	. . . .	p_n	. . . .

则随机变量X的函数 Y= g (X) 的分布律为

Y	$g(x_1)$	$g(x_2)$	$g(x_3)$	. . . . .	$g(x_n)$	. . . .
p_k	p_1	p_2	p_3	. . . . . .	p_n	. . . .

如果$g(x_i)$与$g(x_j)$相同，此时将两项合并，对应概率相加

77

例 设随机变量X的分布律为

X	−1	0	1	2
p_k	0.2	0.3	0.4	0.1

求$Y=2X^2+1$的分布律.

解 由题设可得如下表格

x	-1	0	1	2
$Y=2x^2+1$	3	1	3	9
概率	0.2	0.3	0.4	0.1

所以，$y=2x^2+1$的分布律为

y	1	3	9
p_k	0.3	0.6	0.1

78

例 设圆半径X的分布律为

X	9.5	10	10.5	11
p_k	0.06	0.5	0.4	0.04

求周长及面积的分布律.

解 由题设可得如下表格

x	9.5	10	10.5	11
周长	19π	20π	21π	22π
面积	90.25π	100π	110.25π	121π
概率	0.06	0.5	0.4	0.04

79

解 所以，周长的分布律为

周长	19π	20π	21π	22π
概率	0.06	0.5	0.4	0.04

面积的分布律为

面积	90.25π	100π	110.25π	121π
概率	0.06	0.5	0.4	0.04

80

连续型随机变量的函数的分布

设 X 为连续型随机变量，其密度函数为 $f(x)$.
y = $g(x)$为一个连续函数，求随机变量Y=g(X)的密度函数.

- 一般方法

(1) 求Y的分布函数 $F_Y(y)$

$$F_Y(y) \overset{\text{根据分布函数的定义}}{=\!=\!=} P\{Y \le y\} = P\{g(X) \le y\}$$
$$= P\{X \le G(y)\}$$

(2) $F_Y(y)$ 对y求导，得到 $f_Y(y)$

$$f_Y(y) = F_Y'(y)$$

81

例　设随机变量X的密度函数为

$$f(x) = \begin{cases} \frac{x}{8}, 0 < x < 4 \\ 0, 其它 \end{cases}$$

求随机变量Y=2X+8的概率密度。

解（1） 先求Y=2X+8的分布函数 $F_Y(y)$.

$$F_Y(y) = P\{Y \le y\} = P\{2X + 8 \le y\}$$
$$= P\{X \le \frac{y-8}{2}\} = F_X(\frac{y-8}{2})$$

82

（2） 求Y=2X+8的概率密度

$$f_Y(y)=F_Y'(y) = f\left(\frac{y-8}{2}\right)\left(\frac{y-8}{2}\right)'$$

$$=\begin{cases}\frac{1}{8}\left(\frac{y-8}{2}\right)\cdot\frac{1}{2}, & 0<\frac{y-8}{2}<4\\ 0, & 其它\end{cases}$$

$$=\begin{cases}\frac{y-8}{32}, & 8<y<16\\ 0, & 其它\end{cases}$$

83

例 设随机变量X服从正态分布 $N\left(\mu,\sigma^2\right)$

求 $Y=aX+b$ 的概率密度。

解 先求分布函数 $F_Y(y)$。

$$F_Y(y)=P\{Y\le y\}=P\{aX+b\le y\}$$

(i)当 $a>0$ 时，

$$F_Y(y)=P\{X\le\frac{y-b}{a}\}=F_X\left(\frac{y-b}{a}\right)$$

所以，

$$f_Y(y)=f(\frac{y-b}{a})\cdot\frac{1}{a}=\frac{1}{\sqrt{2\pi}\sigma a}\cdot e^{-\frac{(y-b-a\mu)^2}{2(a\sigma)^2}}$$

84

(ii)当 $a<0$ 时， $F_Y(y)=P\{X\geq\frac{y-b}{a}\}$

$$F_Y(y)=P\{X\geq\frac{y-b}{a}\}=1-P\{X<\frac{y-b}{a}\}$$

$$=1-F_X(\frac{y-b}{a})$$

$$f_Y(y)=-f(\frac{y-b}{a})\cdot\frac{1}{a}=\frac{-1}{\sqrt{2\pi}\sigma a}\cdot e^{-\frac{(y-b-a\mu)^2}{2(a\sigma)^2}}$$

所以， $Y\sim N\left(a\mu+b,a^2\sigma^2\right)$

85

■ **定理** 正态分布的线性函数仍服从正态分布

设$X\sim N(\mu,\sigma^2)$, $Y=aX+b(a\neq 0)$, 则 $Y\sim N(a\mu+b,a^2\sigma^2)$

■ **推论**

若$X\sim N(\mu,\sigma^2)$, 则$\frac{X-\mu}{\sigma}\sim N(0,1)$

正态分布的标准化

86

例 设U ~ N(0，1)，其密度函数为:

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},\quad -\infty<x<+\infty$$

则$Y=U^2$的概率密度函数为

$$f_Y(y)=\begin{cases}\frac{1}{\sqrt{2\pi y}}e^{-\frac{y}{2}}, & y>0\\ 0, & y\le 0\end{cases}$$

此时称Y服从自由度为1的 χ^2分布,记作 $Y\sim\chi^2(1)$

结论：若 $U\sim N(0,1)$ 则 $U^2\sim\chi^2(1)$

87

定理 若随机变量X和随机变量Y=g(X)的密度函数分别为 $f_X(x)$ $f_Y(y)$, 当 g(x) 是严格单调函数，则

$$f_Y(y)=f_X[(G(y))]\left|G'(y)\right|$$

其中 $x=G(y)$ 为 $y=g(x)$ 的反函数

88

例 设随机变量X服从[90，110]上的均匀分布,求Y=0.1X+10的密度函数。

解 X的密度函数为

$$f_X(x)=\begin{cases}\dfrac{1}{20}, & 90\le x\le 110\\ 0, & \text{其它}\end{cases}$$

Y=0.1X+10的密度函数为

$$f_Y(y)=\frac{1}{0.1}f_x\left(\frac{y-10}{0.1}\right)=\begin{cases}\dfrac{1}{2}, & 19\le y\le 21\\ 0, & \text{其它}\end{cases}$$

即 Y 服从[19，21]上的均匀分布.

89

例 设球的半径X的概率密度为

$$f(x)=\begin{cases}6x(1-x), & x\in(0,1)\\ 0, & \text{其它}\end{cases}$$ 试求体积的概率密度。

解 体积 $Y=\frac{4}{3}\pi X^3$的分布函数为

$$F_Y(y)=P\left\{\frac{4}{3}\pi X^3\le y\right\}=P\left\{X\le\sqrt[3]{\frac{3y}{4\pi}}\right\}=F_X\left(\sqrt[3]{\frac{3y}{4\pi}}\right)$$

所以体积的概率密度为

$$f_Y(y)=f_X\left(\sqrt[3]{\frac{3y}{4\pi}}\right)\cdot\left(\sqrt[3]{\frac{3y}{4\pi}}\right)'$$

$$=f_X\left(\sqrt[3]{\frac{3y}{4\pi}}\right)\cdot\frac{1}{3}\cdot\left(\frac{3y}{4\pi}\right)^{-2/3}\cdot\frac{3}{4\pi}$$

90

所以体积的概率密度为

$$f_Y(y)=f_X\left(\sqrt[3]{\frac{3y}{4\pi}}\right)\cdot\left(\sqrt[3]{\frac{3y}{4\pi}}\right)'$$

$$=f_X\left(\sqrt[3]{\frac{3y}{4\pi}}\right)\cdot\frac{1}{3}\cdot\left(\frac{3y}{4\pi}\right)^{-2/3}\cdot\frac{3}{4\pi}$$

即 $$f_Y(y)=\begin{cases}\dfrac{3}{2\pi}\left(\sqrt[3]{\dfrac{4\pi}{3y}}-1\right), & y\in\left(0,\dfrac{4}{3}\pi\right)\\ 0, & \text{其它}\end{cases}$$

91

课后练习

P55 习题2
23--26

作业

26

92

本章作业

1, 7, 15, 21, 26

下节课交作业，请同学们做好作业，下次课不要忘了带过来

93