概率论与数理统计

第二章 · 随机变量

2020 年 2 月 12 日

■ 暨南大学数学系 ■ 吕荐瑞

第一节	随机变量的定义
第二节	离散型随机变量
第三节	连续型随机变量
第四节	随机变量函数的分布

随机变量

随机试验的结果通常可以用数量来表示：

随机变量

随机试验的结果通常可以用数量来表示：

- 扔一个硬币所得的结果；
- 掷一颗骰子所得的点数；
- 抽查样品时的废品个数；
- 广州每日的平均气温；
- 某电子管的使用寿命；

随机变量

随机试验的结果通常可以用数量来表示：

- 扔一个硬币所得的结果；
- 掷一颗骰子所得的点数；
- 抽查样品时的废品个数；
- 广州每日的平均气温；
- 某电子管的使用寿命；

将试验结果数值化，就产生了随机变量的概念.

随机变量

定义 1　设 Ω 是某随机试验的样本空间．如果对于每个 $\omega \in \Omega$，都有一个实数 $X(\omega)$ 与其对应，这样就得到一个定义在 Ω 上的函数

$$X = X(\omega),$$

称该函数为随机变量(random variable).

随机变量一般用大写英文字母 X、Y、Z 或小写希腊字母 ξ、η、γ 来表示.

随机变量的分类

1. 离散型：只能取有限个或者可列个值

随机变量的分类

1. 离散型：只能取有限个或者可列个值
   - 抽取到的次品数

随机变量的分类

1. 离散型：只能取有限个或者可列个值
   - 抽取到的次品数
   - 收到的呼叫次数

随机变量的分类

1. 离散型：只能取有限个或者可列个值
   - 抽取到的次品数
   - 收到的呼叫次数
2. 连续型：取得某一区间内的任何数值

随机变量的分类

1. 离散型：只能取有限个或者可列个值
   - 抽取到的次品数
   - 收到的呼叫次数
2. 连续型：取得某一区间内的任何数值
   - 电视机的寿命

随机变量的分类

1. 离散型：只能取有限个或者可列个值
   - 抽取到的次品数
   - 收到的呼叫次数
2. 连续型：取得某一区间内的任何数值
   - 电视机的寿命
   - 测量的误差

离散型随机变量

定义 1　只能取有限个或者可列个值的随机变量，称为离散型随机变量.

- 骰子的点数
- 抽取的次品数
- 收到的呼叫次数

第二节	离散型随机变量
A	离散型随机变量的概率分布
B	常见的离散型随机变量
C	离散型随机变量的分布函数

若离散型随机变量 X 的所有可能值为 $\{x_k\}$，分别对应概率 $\{p_k\}$，则称

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

为 X 的概率分布.

若离散型随机变量 X 的所有可能值为 $\{x_k\}$，分别对应概率 $\{p_k\}$，则称

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

为 X 的概率分布.

离散型随机变量也常用下面的概率分布表给出：

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

若离散型随机变量 X 的所有可能值为 $\{x_k\}$，分别对应概率 $\{p_k\}$，则称

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

为 X 的概率分布.

离散型随机变量也常用下面的概率分布表给出：

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

概率分布 $\{p_k\}$ 的性质：$\begin{cases} p_k \geqslant 0, \quad k = 1, 2, \cdots \\ \sum\limits_k p_k = 1 \end{cases}$

例 1　100 件产品中，有 98 件是正品，2 件是次品，今从中随机地抽取一件，若规定

$$X=\begin{cases}1, & \text{取到正品；}\\ 0, & \text{取到次品；}\end{cases}$$

例 1　100 件产品中，有 98 件是正品，2 件是次品，今从中随机地抽取一件，若规定

$$X=\begin{cases}1, & \text{取到正品;}\\ 0, & \text{取到次品;}\end{cases}$$

则随机变量 X 的概率分布表为

X	0	1
P	0.02	0.98

.

例 1 100 件产品中，有 98 件是正品，2 件是次品，今从中随机地抽取一件，若规定

$$X=\begin{cases}1, & \text{取到正品;}\\ 0, & \text{取到次品;}\end{cases}$$

则随机变量 X 的概率分布表为

X	0	1
P	0.02	0.98

.

定义 2 若随机变量 X 的概率分布为

X	0	1
P	$1-p$	p

,

则称 X 服从参数为 p 的两点分布（或 0-1 分布）.

例 2　已知袋中装有 2 白 3 黑共 5 个球．每次从袋中任取一个，直到取到白球为止．求取球次数 X 的概率分布．

(1) 每次都不放回的情形．

(2) 每次都放回的情形．

例 2　已知袋中装有 2 白 3 黑共 5 个球．每次从袋中任取一个，直到取到白球为止．求取球次数 X 的概率分布．

(1) 每次都不放回的情形．……不放回抽样

(2) 每次都放回的情形．……放回抽样

例 2　已知袋中装有 2 白 3 黑共 5 个球．每次从袋中任取一个，直到取到白球为止．求取球次数 X 的概率分布．

(1) 每次都不放回的情形．……不放回抽样

(2) 每次都放回的情形．……放回抽样

定义 3　若随机变量 X 的概率分布为

$$P\{X=k\}=p(1-p)^{k-1},\quad k=1,2,3\cdots,$$

则称 X 服从参数为 p 的几何分布，记为 $X\sim G(p)$.

第二节	离散型随机变量
A	离散型随机变量的概率分布
B	常见的离散型随机变量
C	离散型随机变量的分布函数

常见的离散型随机变量

1. 两点分布 ✓
2. 几何分布 ✓
3. 二项分布
4. 泊松分布

离散型随机变量：二项分布

例 3　若某射手每次射击命中的概率均为 p，现进行 n 次独立射击，求恰有 k 次命中的概率.

离散型随机变量：二项分布

例 3　若某射手每次射击命中的概率均为 p，现进行 n 次独立射击，求恰有 k 次命中的概率.

解答　先研究射击次数 $n = 4$ 的特殊情形．此时有

$k = 0$	✗✗✗✗
$k = 1$	✓✗✗✗, ✗✓✗✗, ✗✗✓✗, ✗✗✗✓
$k = 2$	✓✓✗✗, ✓✗✓✗, ✓✗✗✓, ✗✓✓✗, ✗✓✗✓, ✗✗✓✓
$k = 3$	✓✓✓✗, ✓✓✗✓, ✓✗✓✓, ✗✓✓✓
$k = 4$	✓✓✓✓

离散型随机变量：二项分布

定义 4　只有两种可能结果的试验称为伯努利试验. 将一伯努利试验独立重复 n 次称为n 重伯努利试验.

离散型随机变量：二项分布

定义 4 只有两种可能结果的试验称为**伯努利试验**. 将一伯努利试验独立重复 n 次称为**n 重伯努利试验**.

定理 1 (伯努利定理) 设一次试验中事件 A 发生的概率为 $p(0 < p < 1)$，则 n 重伯努利试验中，事件 A 恰好发生 $k(0 \leqslant k \leqslant n)$ 次的概率为

$$b(k; n, p) = C_n^k p^k (1-p)^{n-k}.$$

离散型随机变量：二项分布

定义 5 如果随机变量 X 服从以下分布律

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

其中 $0<p<1$，$0\leqslant k\leqslant n$，则称 X 服从参数为 n，p 的二项分布，记为 $X\sim B(n,p)$.

离散型随机变量：二项分布

注记　在物品总数很大，而抽取数目较小时，不放回抽样可以看成放回抽样，从而可按伯努利试验来处理.

离散型随机变量：二项分布

注记　在物品总数很大，而抽取数目较小时，不放回抽样可以看成放回抽样，从而可按伯努利试验来处理.

例 4　有一大批灯泡，已知其次品率是 0.2．随机抽出 20 只灯泡做寿命试验，求这 20 只灯泡中恰有 3 只是次品的概率．.....0.205

例 5　设某城市平均每天发生火灾 λ 次，试研究一天内发生火灾次数 X 的概率分布.

例 5　设某城市平均每天发生火灾 λ 次，试研究一天内发生火灾次数 X 的概率分布.

解答　假定该城市的火灾事故满足下列条件：

1. 在短时段内不会发生超过一次火灾，

例 5　设某城市平均每天发生火灾 λ 次，试研究一天内发生火灾次数 X 的概率分布.

解答　假定该城市的火灾事故满足下列条件：

1. 在短时段内不会发生超过一次火灾，
2. 在短时段内发生火灾的概率与该时段长度成正比，

例 5 设某城市平均每天发生火灾 λ 次，试研究一天内发生火灾次数 X 的概率分布.

解答 假定该城市的火灾事故满足下列条件：

1. 在短时段内不会发生超过一次火灾，
2. 在短时段内发生火灾的概率与该时段长度成正比，
3. 在不相交时段内发生的火灾相互独立，

例 5 设某城市平均每天发生火灾 λ 次，试研究一天内发生火灾次数 X 的概率分布.

解答 假定该城市的火灾事故满足下列条件：

1. 在短时段内不会发生超过一次火灾，
2. 在短时段内发生火灾的概率与该时段长度成正比，
3. 在不相交时段内发生的火灾相互独立，

则概率分布为 $P\{X=k\}=\dfrac{\lambda^k}{k!}\mathrm{e}^{-\lambda},\ k=0,1,2\cdots.$

例 5 设某城市平均每天发生火灾 λ 次，试研究一天内发生火灾次数 X 的概率分布.

解答 假定该城市的火灾事故满足下列条件：

1. 在短时段内不会发生超过一次火灾，
2. 在短时段内发生火灾的概率与该时段长度成正比，
3. 在不相交时段内发生的火灾相互独立，

则概率分布为 $P\{X=k\}=\dfrac{\lambda^k}{k!}\mathrm{e}^{-\lambda},\ k=0,1,2\cdots.$

注记 这三个条件分别称为普通性，平稳性和独立性.

定义 6 如果随机变量 X 服从以下分布律

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

其中 $\lambda>0$，则称 X 服从参数为 λ 的泊松分布，记为 $X\sim P(\lambda)$.

定义 6 如果随机变量 X 服从以下分布律

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

其中 $\lambda>0$，则称 X 服从参数为 λ 的泊松分布，记为 $X\sim P(\lambda)$.

例子 泊松分布常与单位时间（或单位面积、单位产品等）上的计数过程相联系：

- 某地区每天发生火灾的次数.
- 某地区每年发生暴雨的次数.
- 某种玻璃每平方米内的气泡数.

离散型随机变量：泊松分布

例 6　某一城市每天发生火灾的次数 X 服从参数为 0.8 的泊松分布．求该城市一天内发生 3 次或以上火灾的概率.

离散型随机变量：泊松分布

例 6　某一城市每天发生火灾的次数 X 服从参数为 0.8 的泊松分布．求该城市一天内发生 3 次或以上火灾的概率．.....0.047423

离散型随机变量：泊松分布

例 6　某一城市每天发生火灾的次数 X 服从参数为 0.8 的泊松分布．求该城市一天内发生 3 次或以上火灾的概率．……0.047423

例 7　某工厂生产的床单，每条床单上含有疵点的个数 X 服从参数 $\lambda = 1$ 的泊松分布．质检部门规定：

- 床单上无疵点或只有一个疵点的为一等品，
- 有 2 到 4 个疵点的为二等品，
- 有 5 个或 5 个以上疵点的为次品．

试求工厂生产的床单为一、二等品和次品的概率．

二项分布与泊松分布

历史上，泊松分布是作为二项分布的近似，于 1837 年由法国数学家泊松（Poisson）引入的.

二项分布与泊松分布

历史上，泊松分布是作为二项分布的近似，于 1837 年由法国数学家泊松（Poisson）引入的.

对二项分布 $B(n,p)$，当 n 充分大、p 很小时，对任意固定的非负整数 k，有近似式

$$b(k;n,p) \approx \frac{\lambda^k}{k!}e^{-\lambda},$$

其中 $\lambda = np$.

离散型随机变量：泊松分布

例 8　某公司生产的一种产品，根据历史生产纪录知，产品的次品率为 0.01，问该种产品 300 件中次品数大于 5 的概率是多少？

离散型随机变量：泊松分布

例 8　某公司生产的一种产品，根据历史生产纪录知，产品的次品率为 0.01，问该种产品 300 件中次品数大于 5 的概率是多少？ 0.0840

离散型随机变量：泊松分布

例 8　某公司生产的一种产品，根据历史生产纪录知，产品的次品率为 0.01，问该种产品 300 件中次品数大于 5 的概率是多少？ 0.0840

练习 1　若一个人注射疫苗之后有不良反应的概率为 0.001．确定 3000 个人注射疫苗后

(1) 恰有 3 个有不良反应的概率；

(2) 多于 2 个有不良反应的概率.

离散型随机变量：泊松分布

例 8　某公司生产的一种产品，根据历史生产纪录知，产品的次品率为 0.01，问该种产品 300 件中次品数大于 5 的概率是多少？ 0.0840

练习 1　若一个人注射疫苗之后有不良反应的概率为 0.001．确定 3000 个人注射疫苗后

(1) 恰有 3 个有不良反应的概率；

(2) 多于 2 个有不良反应的概率.

第二节	离散型随机变量
A	离散型随机变量的概率分布
B	常见的离散型随机变量
C	离散型随机变量的分布函数

定义 7 设离散型随机变量 X 的概率分布为

$$P\{X = x_k\} = p_k, \qquad k = 1, 2, \cdots.$$

则定义 X 的**分布函数**为

$$F(x) = P\{X \leqslant x\} = \sum_{x_k \leqslant x} p_k,$$

这里的和式是对所有满足 $x_k \leqslant x$ 的 p_k 求和.

定义 7 设离散型随机变量 X 的概率分布为

$$P\{X = x_k\} = p_k, \qquad k = 1, 2, \cdots.$$

则定义 X 的分布函数为

$$F(x) = P\{X \leqslant x\} = \sum_{x_k \leqslant x} p_k,$$

这里的和式是对所有满足 $x_k \leqslant x$ 的 p_k 求和.

例 9 求随机变量 X 的分布函数，其中 X 的概率分布如下

X	1	2	3	5
P	0.4	0.3	0.1	0.2

复习与提高

选择　某人向同一目标独立重复射击，每次射击命中目标的概率为 p $(0 < p < 1)$，则此人第 4 次射击恰好第 2 次命中目标的概率为 . . . . . .(　　)

(A) $3p(1-p)^2$　　(B) $6p(1-p)^2$

(C) $3p^2(1-p)^2$　　(D) $6p^2(1-p)^2$

复习与提高

二项分布与两点分布的关系：

1. 两点分布 $B(1,p)$ 是二项分布 $B(n,p)$ 在 $n=1$ 时的特殊情况.

复习与提高

二项分布与两点分布的关系：

1. 两点分布 $B(1,p)$ 是二项分布 $B(n,p)$ 在 $n=1$ 时的特殊情况.
2. 设在某试验中事件 A 的概率为 p，将该试验独立地进行 n 次. 记 X 为 n 次试验中事件 A 发生的总次数，X_i 为第 i 次试验中事件 A 发生的次数，则有

$$X = X_1 + X_2 + \cdots + X_n.$$

复习与提高

复习 1　甲乙两人独立地向目标射击，他们击中目标的概率都为 0.7，求击中目标的人数 X 的概率分布.

复习与提高

复习 1　甲乙两人独立地向目标射击，他们击中目标的概率都为 0.7，求击中目标的人数 X 的概率分布.

复习 2　已知随机变量 X 有如下分布律

X	1	2	3	4	5
P	c	$4c$	$6c$	$4c$	c

求常数 c 并计算

$$P\{X > 2|X < 5\},\ P\{X^2 + 4 = 5X\}.$$

复习与提高

复习 3　假设每分钟通过某路口的车流量服从泊松分布，已知在一分钟内无车通过的概率是 0.2，求在一分钟内至少通过 2 辆车的概率.

连续型随机变量

定义 1　取得某一区间内的任意数值的随机变量，称为连续型随机变量.

- 学校男生的身高
- 电视机的寿命
- 测量的误差
- 在公交站等车的时间

第三节	连续型随机变量
A	连续型随机变量的概率密度
B	常见的连续型随机变量

问题　研究学校男生的身高 X 的概率分布.

问题　研究学校男生的身高 X 的概率分布.

连续型随机变量 X 的取值范围为一个区间. 此时

问题　研究学校男生的身高 X 的概率分布.

连续型随机变量 X 的取值范围为一个区间. 此时

- X 取某个确定值的概率总是等于零.

问题　研究学校男生的身高 X 的概率分布.

连续型随机变量 X 的取值范围为一个区间. 此时

- X 取某个确定值的概率总是等于零.
  - 事件 $X = a$ 可能发生，但 $P\{X = a\} = 0$.

问题　研究学校男生的身高 X 的概率分布.

连续型随机变量 X 的取值范围为一个区间. 此时

- X 取某个确定值的概率总是等于零.
  - 事件 $X = a$ 可能发生，但 $P\{X = a\} = 0$.
- X 在某个小区间内取值的概率可以大于零.

问题　研究学校男生的身高 X 的概率分布.

连续型随机变量 X 的取值范围为一个区间. 此时

- X 取某个确定值的概率总是等于零.
  - 事件 $X = a$ 可能发生，但 $P\{X = a\} = 0$.
- X 在某个小区间内取值的概率可以大于零.

将区间分为若干段，研究 X 在各小区间取值的概率.

问题　研究学校男生的身高 X 的概率分布.

连续型随机变量 X 的取值范围为一个区间. 此时

- X 取某个确定值的概率总是等于零.
  - 事件 $X = a$ 可能发生，但 $P\{X = a\} = 0$.
- X 在某个小区间内取值的概率可以大于零.

将区间分为若干段，研究 X 在各小区间取值的概率.

- 分组的频率直方图刻画随机变量的概率分布.

问题　研究学校男生的身高 X 的概率分布.

连续型随机变量 X 的取值范围为一个区间. 此时

- X 取某个确定值的概率总是等于零.
  - 事件 $X = a$ 可能发生，但 $P\{X = a\} = 0$.
- X 在某个小区间内取值的概率可以大于零.

将区间分为若干段，研究 X 在各小区间取值的概率.

- 分组的频率直方图刻画随机变量的概率分布.
- 分组越细，频率直方图就越接近一条连续曲线.

定义 2 如果存在非负函数 $f(x)$，满足

$$P\{X \leqslant x\} = \int_{-\infty}^{x} f(t)\,\mathrm{d}t,$$

定义 2 如果存在非负函数 $f(x)$，满足

$$P\{X \leqslant x\} = \int_{-\infty}^{x} f(t)\,\mathrm{d}t,$$

则称 X 为连续型随机变量，称 $f(x)$ 为 X 的概率密度函数，记为 $X \sim f(x)$.

定义 2 如果存在非负函数 $f(x)$，满足

$$P\{X \leqslant x\} = \int_{-\infty}^{x} f(t)\,\mathrm{d}t,$$

则称 X 为连续型随机变量，称 $f(x)$ 为 X 的概率密度函数，记为 $X \sim f(x)$.

性质 1 (1) $f(x) \geqslant 0$, (2) $\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1$.

定义 3 对任意随机变量 X（离散或连续），称函数

$$F(x) := P\{X \leqslant x\}, \quad x \in \mathbb{R}$$

为 X 的**分布函数**.

定义 3 对任意随机变量 X（离散或连续），称函数

$$F(x) := P\{X \leqslant x\}, \quad x \in \mathbb{R}$$

为 X 的**分布函数**.

性质 2 随机变量的分布函数 $F(x)$ 满足以下性质：

1. 对任意实数 $a < b$，总有 $F(a) \leqslant F(b)$；

定义 3 对任意随机变量 X（离散或连续），称函数

$$F(x) := P\{X \leqslant x\}, \quad x \in \mathbb{R}$$

为 X 的**分布函数**.

性质 2 随机变量的分布函数 $F(x)$ 满足以下性质：

1. 对任意实数 $a < b$，总有 $F(a) \leqslant F(b)$；
2. $0 \leqslant F(x) \leqslant 1$；

定义 3 对任意随机变量 X（离散或连续），称函数

$$F(x) := P\{X \leqslant x\}, \quad x \in \mathbb{R}$$

为 X 的**分布函数**.

性质 2 随机变量的分布函数 $F(x)$ 满足以下性质：

1. 对任意实数 $a < b$，总有 $F(a) \leqslant F(b)$；
2. $0 \leqslant F(x) \leqslant 1$；
3. $F(-\infty) = 0$，$F(+\infty) = 1$.

分布函数与密度函数

性质 3　连续型随机变量的分布函数总是连续的，而且有

1. $F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t.$
2. $f(x) = \begin{cases} F'(x), & \text{若导数存在;} \\ 0, & \text{若导数不存在.} \end{cases}$

任意区间上概率的计算

由密度函数的定义可知,

$$P\{a < X \leqslant b\} = F(b) - F(a) = \int_a^b f(x)\,\mathrm{d}x.$$

任意区间上概率的计算

由密度函数的定义可知,

$$P\{a < X \leqslant b\} = F(b) - F(a) = \int_a^b f(x)\,\mathrm{d}x.$$

在上式中令 $a \to b$, 可以得到 $P\{X = b\} = 0$.

任意区间上概率的计算

由密度函数的定义可知,

$$P\{a < X \leqslant b\} = F(b) - F(a) = \int_a^b f(x)\,\mathrm{d}x.$$

在上式中令 $a \to b$, 可以得到 $P\{X = b\} = 0$.

因此将 $a < X \leqslant b$ 改为 $a < X < b$, $a \leqslant X < b$ 或 $a \leqslant X \leqslant b$ 后, 第一个等式仍成立.

连续型随机变量

例 1　设连续型随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<0\\ Ax^2, & 0\leqslant x<1\\ 1, & x\geqslant 1\end{cases}.$$

求系数 A，概率 $P\{0.3<X<0.7\}$，及概率密度 $f(x)$.

连续型随机变量

例 2　设连续型随机变量 X 的概率密度为

$$f(x)=\begin{cases} kx+1, & 0\leqslant x\leqslant 2; \\ 0, & \text{其他}. \end{cases}$$

求系数 k 及分布函数 $F(x)$, 并计算 $P\{0.5<X<2\}$.

连续型随机变量

练习 2 设连续型随机变量 X 的概率密度为

$$f(x)=\begin{cases} k|x|, & -1\leqslant x\leqslant 1; \\ 0, & \text{其他}. \end{cases}$$

求系数 k 及分布函数 $F(x)$, 并计算

$$P\{-2<X<0.5\}.$$

第三节	连续型随机变量
A	连续型随机变量的概率密度
B	常见的连续型随机变量

常见的连续型随机变量

1. 均匀分布
2. 指数分布
3. 正态分布

定义 4 若随机变量 X 有概率密度

$$f(x) = \begin{cases} \frac{1}{b-a} & x \in [a,b] \\ 0 & \text{otherwise} \end{cases}, \qquad (a < b)$$

则称 X 服从区间 $[a,b]$ 上的均匀分布，记为

$$X \sim U[a,b].$$

定义 4 若随机变量 X 有概率密度

$$f(x) = \begin{cases} \frac{1}{b-a} & x \in [a,b] \\ 0 & \text{otherwise} \end{cases}, \qquad (a < b)$$

则称 X 服从区间 $[a,b]$ 上的均匀分布，记为

$$X \sim U[a,b].$$

例子 均匀分布有如下这些例子：

- 四舍五入时产生的误差.
- 查看当前时间时的分钟值.

例 3　某医院平均每小时出生 λ 个婴儿，试研究婴儿出生的时间间隔 X 的概率分布.

例 3　某医院平均每小时出生 λ 个婴儿，试研究婴儿出生的时间间隔 X 的概率分布.

解答　婴儿出生事件同样满足这三个条件：①普通性；②平稳性；③独立性.

例 3　某医院平均每小时出生 λ 个婴儿，试研究婴儿出生的时间间隔 X 的概率分布.

解答　婴儿出生事件同样满足这三个条件：①普通性；②平稳性；③独立性. 因此

- 1 小时内出生婴儿个数 $N(1)$ 服从泊松分布 $P(\lambda)$.

例 3　某医院平均每小时出生 λ 个婴儿，试研究婴儿出生的时间间隔 X 的概率分布.

解答　婴儿出生事件同样满足这三个条件：①普通性；②平稳性；③独立性. 因此

- 1 小时内出生婴儿个数 $N(1)$ 服从泊松分布 $P(\lambda)$.
- t 小时内出生婴儿个数 $N(t)$ 服从泊松分布 $P(\lambda t)$.

例 3 某医院平均每小时出生 λ 个婴儿，试研究婴儿出生的时间间隔 X 的概率分布.

解答 婴儿出生事件同样满足这三个条件：①普通性；②平稳性；③独立性. 因此

- 1 小时内出生婴儿个数 $N(1)$ 服从泊松分布 $P(\lambda)$.
- t 小时内出生婴儿个数 $N(t)$ 服从泊松分布 $P(\lambda t)$.

当 $t < 0$ 时，有 $F(t) = P\{X \leqslant t\} = 0$.

例 3 某医院平均每小时出生 λ 个婴儿，试研究婴儿出生的时间间隔 X 的概率分布.

解答 婴儿出生事件同样满足这三个条件：①普通性；②平稳性；③独立性. 因此

- 1 小时内出生婴儿个数 $N(1)$ 服从泊松分布 $P(\lambda)$.
- t 小时内出生婴儿个数 $N(t)$ 服从泊松分布 $P(\lambda t)$.

当 $t<0$ 时，有 $F(t)=P\{X\leqslant t\}=0$. 当 $t\geqslant 0$ 时，

$$\begin{aligned} F(t) &= P\{X\leqslant t\} = 1-P\{X>t\} \\ &= 1-P\{N(t)=0\} = 1-\mathrm{e}^{-\lambda t}. \end{aligned}$$

例 3 某医院平均每小时出生 λ 个婴儿，试研究婴儿出生的时间间隔 X 的概率分布.

解答 婴儿出生事件同样满足这三个条件：①普通性；②平稳性；③独立性. 因此

- 1 小时内出生婴儿个数 $N(1)$ 服从泊松分布 $P(\lambda)$.
- t 小时内出生婴儿个数 $N(t)$ 服从泊松分布 $P(\lambda t)$.

当 $t<0$ 时，有 $F(t)=P\{X\leqslant t\}=0$. 当 $t\geqslant 0$ 时，

$$\begin{aligned}F(t)&=P\{X\leqslant t\}=1-P\{X>t\}\\&=1-P\{N(t)=0\}=1-\mathrm{e}^{-\lambda t}.\end{aligned}$$

因此，当 $t<0$ 时，$f(t)=0$；

例 3 某医院平均每小时出生 λ 个婴儿，试研究婴儿出生的时间间隔 X 的概率分布.

解答 婴儿出生事件同样满足这三个条件：①普通性；②平稳性；③独立性. 因此

- 1 小时内出生婴儿个数 $N(1)$ 服从泊松分布 $P(\lambda)$.
- t 小时内出生婴儿个数 $N(t)$ 服从泊松分布 $P(\lambda t)$.

当 $t<0$ 时，有 $F(t)=P\{X\leqslant t\}=0$. 当 $t\geqslant 0$ 时，

$$\begin{aligned}F(t)&=P\{X\leqslant t\}=1-P\{X>t\}\\&=1-P\{N(t)=0\}=1-\mathrm{e}^{-\lambda t}.\end{aligned}$$

因此，当 $t<0$ 时，$f(t)=0$；当 $t\geqslant 0$ 时，$f(t)=\lambda\mathrm{e}^{-\lambda t}$.

连续型随机变量：指数分布

定义 5 如果随机变量 X 有以下概率密度

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geqslant 0 \\ 0 & \text{otherwise} \end{cases},$$

其中 $\lambda > 0$，则称 X 服从参数为 λ 的指数分布，记为

$$X \sim EP(\lambda).$$

其分布函数为

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

连续型随机变量：指数分布

指数分布经常作为时间间隔或等待时间的分布：

- 婴儿出生的时间间隔
- 客户来电的时间间隔
- 商品销售的时间间隔
- 网站访问的时间间隔
- 取号排队的等待时间
- 电子产品的寿命长度

例 4　设某电子管的使用寿命 X（单位：小时）服从参数 $\lambda = 0.0002$ 的指数分布．求电子管使用寿命超过 3000 小时的概率．

例 4　设某电子管的使用寿命 X（单位：小时）服从参数 $\lambda = 0.0002$ 的指数分布．求电子管使用寿命超过 3000 小时的概率．

练习 3　某元件寿命 X 服从参数为 $\lambda = 0.001$ 的指数分布．3 个这样的元件使用 1000 小时后，都没有损坏的概率是多少？

连续型随机变量：指数分布

例 5 (指数分布的无记忆性)　随机变量 X 服从参数为 λ 的指数分布．设 $s, t > 0$，求以下条件概率

$$P\{X > s + t | X > s\}.$$

连续型随机变量：正态分布

定义 6 如果随机变量 X 有以下概率密度

$$\varphi_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

其中 μ, σ 为常数且 $\sigma > 0$，则称 X 服从正态分布，简记为

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

连续型随机变量：正态分布

定义 6 如果随机变量 X 有以下概率密度

$$\varphi_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

其中 μ, σ 为常数且 $\sigma > 0$，则称 X 服从正态分布，简记为

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

称 $N(0,1)$ 为标准正态分布，并将其概率密度函数 $\varphi_{0,1}(x)$ 简写为 $\varphi(x)$. 即有

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

标准正态分布

标准正态分布的概率密度函数 $\varphi(\cdot)$ 有以下性质：

1. 无穷次可微；
2. 偶函数；
3. 在零点取得最大值；
4. 有拐点 ± 1；
5. 有水平渐近线（x 轴）.

正态分布的分布函数

正态分布的分布函数为

$$\Phi_{\mu,\sigma^2}(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\cdot\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \, dt,$$

该函数不是初等函数.

正态分布的分布函数

正态分布的分布函数为

$$\Phi_{\mu,\sigma^2}(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}\cdot\sigma}\mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}}\,\mathrm{d}t,$$

该函数不是初等函数.

标准正态分布的分布函数简记为 $\Phi(x)$.

$$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t,$$

正态分布的分布函数

标准正态分布的分布函数的性质

$$\Phi(-x) = 1 - \Phi(x).$$

正态分布的分布函数

标准正态分布的分布函数的性质

$$\Phi(-x) = 1 - \Phi(x).$$

例 6　已知 $X \sim N(0,1)$，求

$P(X \leqslant 1.5)$,　　$P(X \leqslant -1.5)$,　$P(|X| \leqslant 1.5)$,

$P(X \leqslant 3.2)$,　　$P(-1 < X \leqslant 2)$.

正态分布的分布函数

标准正态分布的分布函数的性质

$$\Phi(-x) = 1 - \Phi(x).$$

例 6　已知 $X \sim N(0,1)$，求

$P(X \leqslant 1.5)$,　　$P(X \leqslant -1.5)$,　$P(|X| \leqslant 1.5)$,

$P(X \leqslant 3.2)$,　　$P(-1 < X \leqslant 2)$.

一般正态分布的分布函数 $\Phi_{\mu,\sigma^2}(x)$ 可以表示为

$$\Phi_{\mu,\sigma^2}(x) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

正态分布

例 7　已知某台机器生产的螺栓长度 X (单位：厘米) 服从参数 $\mu = 10.05, \sigma = 0.06$ 的正态分布．若规定螺栓长度在 10.05 ± 0.12 内为合格品，试求螺栓为合格品的概率.

正态分布

例 7　已知某台机器生产的螺栓长度 X (单位：厘米) 服从参数 $\mu = 10.05, \sigma = 0.06$ 的正态分布．若规定螺栓长度在 10.05 ± 0.12 内为合格品，试求螺栓为合格品的概率.

例 8　假设某地区成年男性的身高 (单位: cm) $X \sim N(170, 7.69^2)$, 求该地区成年男性的身高超过 175 cm 的概率.

正态分布

练习 4　$X \sim N(8, 0.25)$，求

$$P(|X-8|<1), \quad P(X \leqslant 10).$$

正态分布

练习 4 $X \sim N(8, 0.25)$，求

$$P(|X-8|<1), \quad P(X \leqslant 10).$$

练习 5 $X \sim N(\mu, \sigma^2)$，且

$$P(X \leqslant 3) = 0.6179, \quad P(X \leqslant -5) = 0.0446.$$

求 μ 和 σ.

正态分布

例 9　$X \sim N(\mu, \sigma^2)$，求

$P(|X-\mu|<\sigma)$，$P(|X-\mu|<2\sigma)$，$P(|X-\mu|<3\sigma)$.

正态分布

例 9　$X \sim N(\mu, \sigma^2)$，求

$P(|X-\mu|<\sigma)$, $P(|X-\mu|<2\sigma)$, $P(|X-\mu|<3\sigma)$.

3σ 原则：在应用中，对 $X \sim N(\mu, \sigma^2)$，通常认为 X 只取 $(\mu-3\sigma, \mu+3\sigma)$ 中的值．计算可知，使用该原则犯错误的概率不到千分之三.

正态分布

设随机变量 $X \sim N(0,1)$，对给定的 $\alpha \in (0,1)$，称满足条件

$$P\{X > Z_\alpha\} = \alpha$$

的点 Z_α 为标准正态分布的上 α 分位点.

正态分布

设随机变量 $X \sim N(0,1)$，对给定的 $\alpha \in (0,1)$，称满足条件

$$P\{X > Z_\alpha\} = \alpha$$

的点 Z_α 为标准正态分布的上 α 分位点．易知

$$\Phi(Z_\alpha) = 1 - \alpha$$

例 10　求 $Z_{0.05}, Z_{0.025}, Z_{0.005}$.

复习与提高

选择　设 $F_1(x)$，$F_2(x)$ 为两个分布函数，其相应的概率密度 $f_1(x)$ 与 $f_2(x)$ 是连续函数，则必为概率密度的是 . . . . . (　　)

(A) $f_1(x)f_2(x)$

(B) $2f_2(x)F_1(x)$

(C) $f_1(x)F_2(x)$

(D) $f_1(x)F_2(x)+f_2(x)F_1(x)$

复习与提高

复习 1　已知连续型随机变量 X 的密度函数为

$$f(x)=\begin{cases}-6x^2+6x+C, & 0\leqslant x\leqslant 1;\\ 0, & \text{otherwise.}\end{cases}$$

求常数 C 并计算 $P\{X\mid -0.5<X\leqslant 0.5\}$.

复习与提高

复习 2　已知连续型随机变量 X 的密度函数为

$$f(x)=\begin{cases}12x^2-12x+C, & 0\leqslant x\leqslant 1;\\ 0, & \text{otherwise.}\end{cases}$$

求常数 C 并计算 $P\{X\leqslant 0.2\mid 0.1<X\leqslant 0.5\}$.

第一节	随机变量的定义
第二节	离散型随机变量
第三节	连续型随机变量
第四节	随机变量函数的分布

随机变量函数的分布

在实际问题中，有时我们关心的随机变量 Y 不容易直接测量，而是要测量另外一个随机变量 X，把 Y 表示为 X 的函数 $Y = g(X)$.

随机变量函数的分布

在实际问题中，有时我们关心的随机变量 Y 不容易直接测量，而是要测量另外一个随机变量 X，把 Y 表示为 X 的函数 $Y = g(X)$.

由此引出的问题是：已知 X 的分布，如何求 Y 的分布？

随机变量函数的分布

在实际问题中，有时我们关心的随机变量 Y 不容易直接测量，而是要测量另外一个随机变量 X，把 Y 表示为 X 的函数 $Y = g(X)$.

由此引出的问题是：已知 X 的分布，如何求 Y 的分布？

例如：已知圆球直径 D 的分布，求圆球体积 $V = \dfrac{\pi D^3}{6}$ 的分布.

第四节	随机变量函数的分布
A	离散型随机变量函数的分布
B	连续型随机变量函数的分布

离散型随机变量函数的分布

例 1　设随机变量 X 有如下概率分布：

X	-1	0	1	2
P	0.2	0.3	0.1	0.4

求 $Y=(X-1)^2$ 的概率分布.

离散型随机变量的函数的分布

总结：设离散型随机变量 X 的概率分布为

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

令 $Y = g(X)$，则 Y 也是一个离散型随机变量，其分布可按如下步骤求得

1. 根据函数关系列出 Y 的所有可能值；
2. 对 Y 的每个可能值 y，$P\{Y = y\}$ 等于所有满足 $g(x_k) = y$ 的 p_k 之和.

离散型随机变量的函数的分布

练习 1 假设随机变量 X 的分布为

X	-1	0	1	2
P	0.2	0.25	0.3	0.25

求 $Y = X^2 + 1$ 以及 $Z = X^2 - X$ 的概率分布.

第四节	随机变量函数的分布
A	离散型随机变量函数的分布
B	连续型随机变量函数的分布

连续型随机变量函数的分布

对连续型随机变量 X，求 $Y = g(X)$ 的密度函数的基本方法是

1. 根据函数关系先求 Y 的分布函数

$$F_Y(y) = P\{Y \leqslant y\} = P\{g(X) \leqslant y\}$$

连续型随机变量函数的分布

对连续型随机变量 X，求 $Y = g(X)$ 的密度函数的基本方法是

1. 根据函数关系先求 Y 的分布函数

$$F_Y(y) = P\{Y \leqslant y\} = P\{g(X) \leqslant y\}$$

2. 然后对 $F_Y(y)$ 求导可得 Y 的概率密度.

连续型随机变量函数的分布

例 2　设随机变量 X 有概率密度函数

$$f_X(x) = \begin{cases} x/8, & 0 < x < 4 \\ 0, & \text{else} \end{cases}$$

求随机变量 $Y = 2X - 1$ 的概率密度函数.

连续型随机变量函数的分布

例 2　设随机变量 X 有概率密度函数

$$f_X(x) = \begin{cases} x/8, & 0 < x < 4 \\ 0, & \text{else} \end{cases}$$

求随机变量 $Y = 2X - 1$ 的概率密度函数.

练习 2　设随机变量 X 有概率密度函数

$$f_X(x) = \begin{cases} 2x, & x \in [0, 1] \\ 0, & \text{else} \end{cases}$$

求随机变量 $Y = -3X + 2$ 的概率密度函数.

连续型随机变量函数的分布

例 3　设随机变量 X 有概率密度函数

$$f_X(x) = \begin{cases} |x|, & x \in (-1, 1) \\ 0, & \text{else} \end{cases}$$

求随机变量 $Y = 2X + 1$ 的概率密度函数.

连续型随机变量函数的分布

例 3　设随机变量 X 有概率密度函数

$$f_X(x) = \begin{cases} |x|, & x \in (-1, 1) \\ 0, & \text{else} \end{cases}$$

求随机变量 $Y = 2X + 1$ 的概率密度函数.

例 4　设 $X \sim U[0, 1]$，求 $Y = e^X$ 的概率密度.

连续型随机变量函数的分布

例 5　设 $X \sim N(0,1)$，求 $Y = e^X$ 的概率密度.

连续型随机变量函数的分布

例 5　设 $X \sim N(0,1)$，求 $Y = e^X$ 的概率密度.

练习 3 (1995 数四)　假设随机变量 X 的概率密度为

$$f_X(x) = \begin{cases} 2e^{-2x}, & x > 0 \\ 0, & x \leqslant 0 \end{cases}$$

证明：$Y = 1 - e^{-2X}$ 服从区间 $(0,1)$ 上的均匀分布.

一般正态分布的标准化

例 6　设 $X \sim N(\mu, \sigma^2)$，求 $Y = \dfrac{X-\mu}{\sigma}$ 的概率密度.

一般正态分布的标准化

例 6　设 $X \sim N(\mu, \sigma^2)$，求 $Y = \dfrac{X-\mu}{\sigma}$ 的概率密度.

解答　$Y = \dfrac{X-\mu}{\sigma} \sim N(0,1)$.

一般正态分布的标准化

例 6　设 $X \sim N(\mu, \sigma^2)$，求 $Y = \dfrac{X-\mu}{\sigma}$ 的概率密度.

解答　$Y = \dfrac{X-\mu}{\sigma} \sim N(0,1)$.

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

$$Y \sim N(a\mu + b, a^2\sigma^2).$$

连续型随机变量函数的分布

定理 1 设随机变量 X 有密度 $f_X(x), x \in (\alpha, \beta)$. 如果 $g(x)$ 在 (α, β) 上是严格单调的连续函数，存在唯一的反函数 $x = h(y), y \in (a, b)$，并且 $h'(y)$ 存在且连续，那么 $Y = g(X)$ 也是连续型随机变量，具有密度函数

$$f_Y(y) = \begin{cases} f_X[h(y)] \cdot |h'(y)|, & y \in (a, b) \\ 0, & \text{otherwise.} \end{cases}$$

连续型随机变量函数的分布

当 $g(x)$ 不是严格单调函数时，不能套用前面定理来得到 $Y = g(X)$ 的概率密度，而只能用我们刚开始介绍的方法.

连续型随机变量函数的分布

当 $g(x)$ 不是严格单调函数时，不能套用前面定理来得到 $Y = g(X)$ 的概率密度，而只能用我们刚开始介绍的方法.

例 7　设 X 有概率密度 $f_X(x)$，求 $Y = X^2$ 的概率密度函数.

连续型随机变量函数的分布

当 $g(x)$ 不是严格单调函数时，不能套用前面定理来得到 $Y = g(X)$ 的概率密度，而只能用我们刚开始介绍的方法.

例 7　设 X 有概率密度 $f_X(x)$，求 $Y = X^2$ 的概率密度函数.

例 8　设 $X \sim U[-\frac{\pi}{2}, \frac{\pi}{2}]$，求 $Y = 2\cos X$ 的概率密度函数.