概率论与数理统计

庄玮玮
weizh@ustc.edu.cn

安徽 合肥

2020 年 3 月

第二章　随机变量及概率分布

§2.1 离散型随机变量

定义2.1.1 对随机变量X, 称x的函数

$$F(x) \equiv P(X \leqslant x), \quad -\infty \leqslant x \leqslant \infty$$

为X的概率分布函数, 简称为分布函数(distribution function).

分布函数$F(x)$的常用性质:

(1) F是单调不减的右连续函数;

(2) $\lim\limits_{x\to\infty} F(x) = F(\infty) = 1, \quad \lim\limits_{x\to-\infty} F(x) = F(-\infty) = 0;$

(3) F在点x连续的充分必要条件是$P(X = x) = 0$.

§2.1 离散型随机变量

下面是常用的离散随机变量及其概率分布.

(1) 伯努利分布(Bernoulli分布) $\mathcal{B}(1,p)$: 如果X只取值0或1, 概率分布是

$$P(X=1)=p,\quad P(X=0)=q,\quad p+q=1,$$

则称X服从伯努利分布, 记做$X\sim\mathcal{B}(1,p)$.

任何试验, 当只考虑成功与否时, 都可以用服从伯努利分布的随机变量描述:

$$X=\begin{cases}1, & \text{试验成功},\\ 0, & \text{试验不成功}.\end{cases}$$

§2.1 离散型随机变量

(2) 二项分布(Binomial 分布) $\mathcal{B}(n,p)$: 对于 $p,\ q>0,\ p+q=1$, 如果随机变量X有概率分布

$$P(X=k)=\mathrm{C}_n^k p^k q^{n-k},\ \ k=0,1,\cdots,n,$$

则称X服从二项分布, 记做$X\sim\mathcal{B}(n,p)$.

称为二项分布的原因是$\mathrm{C}_n^k p^k q^{n-k}$为二项展开式

$$(p+q)^n=\sum_{k=0}^{n}\mathrm{C}_n^k p^k q^{n-k}\quad \text{的第 } k+1 \text{ 项}.$$

§2.1 离散型随机变量

图2.1.1 是 $\mathcal{B}(5, 0.6)$ 的概率分布函数，横坐标是x，纵坐标是$F(x) = P(X \leqslant x)$.

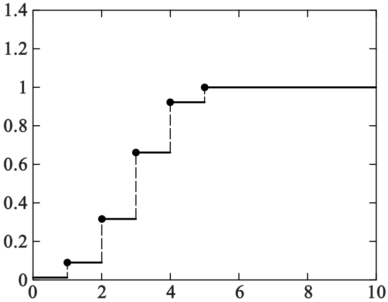


图2.1.1 $\mathcal{B}(5, 0.6)$的分布函数

§2.1 离散型随机变量

二项分布的背景：容易看出$n=1$时的二项分布就是伯努利分布. 对于一般的 n, 设试验S成功的概率为p, 将试验S独立重复n次, 用X表示成功的次数, 下面证明 $X \sim \mathcal{B}(n,p)$.

用A_j表示第j次试验成功, 则A_1, A_2, $\cdots$, A_n相互独立, 且$P(A_j)=p$. 从n次试验中选定k次试验的方法共有C_n^k种. 对第j种选定的$\{j_1,j_2,\cdots,j_k\}$, 用

$$B_j = A_{j_1}A_{j_2}\cdots A_{j_k}\overline{A}_{j_{k+1}}\overline{A}_{j_{k+2}}\cdots\overline{A}_{j_n}$$

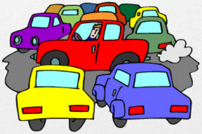
表示第$j_1,j_2,\cdots,j_k$次试验成功, 其余的试验不成功, 则$\{B_j\}$互不相容, 用概率的可加性得到

$$P(X=k)=\mathrm{C}_n^k p^k q^{n-k}.$$

例2.1.0 有一繁忙的汽车站,每天有大量汽车通过,设每辆汽车在一天的某段时间内,出事故的概率为0.0001,在每天的该段时间内有1000 辆汽车通过, 问出事故的次数不小于2的概率是多少?

解　设 1000 辆车通过,出事故的次数为 X, 则 $X \sim B(1000, 0.0001)$,

故所求概率为



$$P\{X \geq 2\} = 1 - P\{X = 0\} - P\{X = 1\}$$

$$= 1 - 0.9999^{1000} - C_{1000}^{1} \cdot 0.0001 \cdot 0.9999^{999} \ ??$$

二项分布 $\xrightarrow{np \to \lambda (n \to +\infty)}$ 泊松分布

§2.1 离散型随机变量

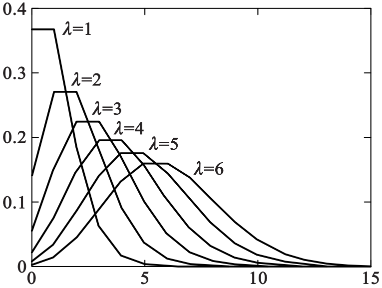
(3) 泊松分布(Poisson 分布) $\mathcal{P}(\lambda)$: 设λ是正常数. 如果随机变量X有概率分布

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

则称X服从参数是λ的泊松分布, 简记为$X\sim\mathcal{P}(\lambda)$.

§2.2 离散型随机变量

下图是$\mathcal{P}(\lambda)$的概率分布折线图. 按最大值由高到低参数依次是$\lambda = 1, 2, \cdots, 6$, 横坐标是$k$, 纵坐标是$P(X = k)$.



$\mathcal{P}(\lambda), \ \lambda = 1, 2, \cdots, 6$

§2.1 离散型随机变量

实际问题中有许多泊松分布的例子.

例2.1.1 1910年, 科学家卢瑟福和盖革观察了放射性物质钋(polonium)放射α 粒子的情况. 他们进行了$N = 2\,608$次观测, 每次观测7.5 s，一共观测到10 094个α 粒子放出, 表2.1.1是观测记录, 其中的Y是服从$\mathcal{P}(3.87)$分布的随机变量,

$$3.87 = 10\,094/2\,608$$

是7.5 s 中放射出α 粒子的平均数.

§2.1 离散型随机变量

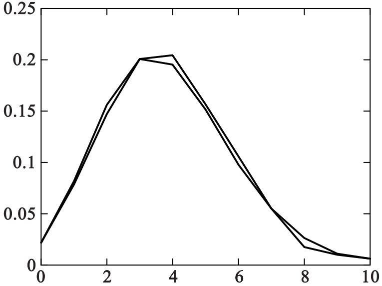
表 2.1.1

观测到的 α 粒子数 k	观测到 k 个粒子的次数 m_k	发生的频率 m_k/N	$P(Y=k)$ $Y\sim\mathcal{P}(3.87)$
0	57	0.022	0.021
1	203	0.078	0.081
2	383	0.147	0.156
3	525	0.201	0.201
4	532	0.204	0.195
5	408	0.156	0.151
6	273	0.105	0.097
7	139	0.053	0.054
8	45	0.017	0.026
9	27	0.010	0.011
10+	16	0.006	0.007
总计	2 608	0.999	1.00

§2.1 离散型随机变量

用X表示这块放射性钋在7.5 s 内放射出的α 粒子数. 表的最后两列表明,事件$\{X=k\}$在$N=2\,608$次重复观测中发生的频率m_k/N和$P(Y=k)$基本相同. 将他们用折线图表示出来就更清楚了. 横坐标是k, 纵坐标是频率m_k/N和$P(Y=k)$. 从下图看到

$$P(X=k)\approx P(Y=k),\ 0\leqslant k<10.$$



例2.1.1中的频率m_k/N 和概率$P(Y=k)$

§2.1 离散型随机变量

下面证明$X \sim \mathcal{P}(\lambda)$. 设想将$t$ 等分成n段, 每段长$\delta_n = \dfrac{t}{n}$. 对充分大的n, 假定:

(a) 在δ_n内最多只有一个α 粒子放出, 并且放出一个粒子的概率是$p_n = \mu\delta_n = \mu t/n$, 这里$\mu$是正常数;

(b) 不同的时间段内是否放射出α 粒子相互独立.

§2.1 离散型随机变量

在以上的假定下，这块放射性物质放射出的粒子数X服从$\mathcal{B}(n, p_n)$. 于是

$$\begin{aligned}
P(X=k) &= \lim_{n\to\infty} \mathrm{C}_n^k p_n^k (1-p_n)^{n-k} \\
&= \lim_{n\to\infty} \frac{n!}{k!(n-k)!}\left(\frac{\mu t}{n}\right)^k \left(1-\frac{\mu t}{n}\right)^{n-k} \\
&= \lim_{n\to\infty} \frac{(\mu t)^k}{k!} \frac{n(n-1)\cdots(n-k+1)}{n^k}\left(1-\frac{\mu t}{n}\right)^{n-k} \\
&= \frac{(\mu t)^k}{k!}\mathrm{e}^{-\mu t}.
\end{aligned}$$

取$\lambda = \mu t$, 得$X \sim \mathcal{P}(\lambda)$.

§2.1 离散型随机变量

由于λ和t 成正比, 所以λ越大, 单位时间内放射出的α 粒子就越多. 另外, 对一般的时间段$(0, t]$, 如果用$N(t)$表示$(0, t]$内观测到的α 粒子数, 则$N(t)$服从泊松分布$\mathcal{P}(\mu t)$.

该推导过程验证了二项分布向泊松分布逼近的事实: 如果 n 很大, p 很小, 则可以用泊松分布来描述二项分布 $\mathcal{B}(n, p)$. 特别是当 n 无法确定时, 就应当使用泊松分布了.

§2.1 离散型随机变量

由于λ和t 成正比，所以λ越大，单位时间内放射出的α 粒子就越多．另外，对一般的时间段$(0, t]$，如果用$N(t)$表示$(0, t]$内观测到的α 粒子数，则$N(t)$服从泊松分布$\mathcal{P}(\mu t)$.

该推导过程验证了二项分布向泊松分布逼近的事实：如果 n 很大，p 很小，则可以用泊松分布来描述二项分布 $\mathcal{B}(n, p)$．特别是当 n 无法确定时，就应当使用泊松分布了．

二项分布的泊松逼近

泊松定理 设λ>0是常数，n是任意正整数，且$\lambda=np$，则对于任意取定的非负整数k，有

$$\lim_{n\to\infty} C_n^k p^k (1-p)^{n-k} = \frac{\lambda^k}{k!}e^{-\lambda}$$

由**泊松定理** 若 n 很大 p 很小时，且 $\lambda=np$，则有

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

$$\boxed{n \ge 20,\ p \le 0.05}$$

接例2.1.0:

$$P\{X \geq 2\} = 1 - P\{X = 0\} - P\{X = 1\}$$

$$= 1 - 0.9999^{1000} - \binom{1000}{1} \cdot 0.0001 \cdot 0.9999^{999}$$

可利用泊松定理 $\lambda = 1000 \times 0.0001 = 0.1,$

$$P\{X \geq 2\} \approx 1 - \frac{e^{-0.1}}{0!} - \frac{0.1 \cdot e^{-0.1}}{1!} = 0.0047$$

$$\boxed{e \approx 2.71828\ldots}$$

§2.1 离散型随机变量

泊松分布可以描述许多有类似背景的随机现象：

- 一部手机一天内收到的短信数X服从泊松分布.这是因为如果只有n个人知道这部手机的号码，每个人在这天发来一条短信的概率是p，则一天内这部手机收到的短信数X服从二项分布 $\mathcal{B}(n,p)$. 因为p比较小，n比较大，而且也不能确定具体的数目，根据二项分布和泊松分布的关系知道用泊松分布描述X的分布更合适；
- 一个E-mail 帐号在一天内收到的E-mail 数服从泊松分布；
- 确定的高速公路段上一周内发生的交通事故数服从泊松分布；
- 一上午某大学上课迟到的总人数也近似服从泊松分布.

§2.1 离散型随机变量

例2.1.2 设一部手机在时间段$[0,t]$内收到的短信数服从泊松分布$\mathcal{P}(\lambda)$，其中$\lambda=\mu t$，μ是正数．每个短信是否广告与其到达时间独立，也与其他短信是否广告独立.假设每个短信是广告的概率$p>0$.

(a) 已知$[0,t]$内收到了n个短信，求其中广告短信数的概率分布$\{h_k\}$;

(b) 计算$[0,t]$内收到的广告短信数的概率分布.

§2.1 离散型随机变量

解 设$[0, t]$内收到的短信数是Y, 其中的广告短信数是X. 每收到一个短信相当于作一次试验, 遇到广告短信是试验成功.

(a) 已知$[0, t]$内收到n个短信, 相当于作了n次独立重复试验, 每次试验成功的概率是p. 根据二项分布知道其中广告短信数的概率分布是

$$h_k = P(X = k|Y = n) = \mathrm{C}_n^k p^k q^{n-k}, \quad 0 \leqslant k \leqslant n, \ q = 1 - p.$$

(b) 因为$\{Y = j\}, j = 0, 1, \cdots$ 是完备事件组. 用全概率公式得到X的概率分布

§2.1 离散型随机变量

$$
\begin{aligned}
P(X=k) &= \sum_{n=k}^{\infty} P(Y=n)P(X=k|Y=n) \\
&= \sum_{n=k}^{\infty} \frac{\lambda^n}{n!}\mathrm{e}^{-\lambda}\mathrm{C}_n^k p^k q^{n-k} \\
&= \sum_{n=k}^{\infty} \frac{(\lambda q)^{n-k}}{k!(n-k)!}\mathrm{e}^{-\lambda}(\lambda p)^k \\
&= \frac{(\lambda p)^k}{k!}\mathrm{e}^{-\lambda}\sum_{j=0}^{\infty}\frac{(\lambda q)^j}{j!} \\
&= \frac{(\lambda p)^k}{k!}\mathrm{e}^{-\lambda}\mathrm{e}^{\lambda q} \\
&= \frac{(\lambda p)^k}{k!}\mathrm{e}^{-\lambda p}, \quad k=0,1,\cdots.
\end{aligned}
$$

说明X服从泊松分布$\mathcal{P}(p\lambda)$.

§2.1 离散型随机变量

从问题的背景也可以想象出上面的结果:

(a) 在互不相交的时间内收到的短信数是相互独立的, 在充分小的时间段$(t, t+\Delta t]$ 内最多只收到一个短信, 收到的短信数和时间段的长度Δt成正比, 于是$[0, t]$内收到的短信数服从泊松分布(参考例2.1).

(b) 同理, 在互不相交的时间内该手机收到的广告短信数是相互独立的, 在充分小的时间段$(t, t+\Delta t]$ 内最多只收到一个广告短信, 收到的广告短信数和时间段的长度Δt成正比, 于是收到的广告短信数也服从泊松分布.

§2.1 离散型随机变量

从问题的背景也可以想象出上面的结果:

(a) 在互不相交的时间内收到的短信数是相互独立的，在充分小的时间段$(t, t+\Delta t]$ 内最多只收到一个短信，收到的短信数和时间段的长度Δt成正比，于是$[0, t]$内收到的短信数服从泊松分布(参考例2.1).

(b) 同理，在互不相交的时间内该手机收到的广告短信数是相互独立的，在充分小的时间段$(t, t+\Delta t]$ 内最多只收到一个广告短信，收到的广告短信数和时间段的长度Δt成正比，于是收到的广告短信数也服从泊松分布.

§2.1 离散型随机变量

(4) 超几何分布(Hypergeometric 分布) $H(N, M, n)$: 如果X的概率分布是

$$P(X=m)=\frac{\mathrm{C}_M^m\ \mathrm{C}_{N-M}^{n-m}}{\mathrm{C}_N^n},\quad m=0,1,\cdots,M,$$

则称X服从超几何分布,记作$X\sim H(N,M,n)$.

注意对$k<0$和$k>n$, 已经约定$\mathrm{C}_n^k=0$. 名称超几何分布来自超几何函数,类似于名称二项分布来自二项展开式.

§2.1 离散型随机变量

例2.1.3 在N件产品中恰有M件次品, 从中任取n件, 用X表示这n 件中的次品数, 则X服从超几何分布.

解 从N件产品中任取n件时, 等可能结果的总数是$|\Omega| = C_N^n$. 事件$X = m$ 表示取到的n件中有m件次品和$n-m$件正品, 对这m件次品有C_M^m种取法, 对这$n-m$件正品有 C_{N-M}^{n-m} 种取法, 所以$|\{X = m\}| = C_M^m C_{N-M}^{n-m}$.

§2.1 离散型随机变量

例2.1.3 在N件产品中恰有M件次品，从中任取n件，用X表示这n 件中的次品数，则X服从超几何分布.

解 从N件产品中任取n件时，等可能结果的总数是$|\Omega| = \mathrm{C}_N^n$. 事件$X = m$ 表示取到的n件中有m件次品和$n-m$件正品，对这m件次品有C_M^m种取法，对这$n-m$件正品有 C_{N-M}^{n-m} 种取法，所以$|\{X = m\}| = \mathrm{C}_M^m \mathrm{C}_{N-M}^{n-m}$.

§2.1 离散型随机变量

例2.1.4 当科学技术发展到今天，任何国家的导弹发射基地都不能躲过敌方的侦察. 为了有效地保存自己的导弹发射装置,A国采用了构建真假导弹发射井的方法. 假设A国的100个发射井中有10个发射井是发射导弹的真井，另外90个是假井. 在对A国的第一波精确打击中，至少要摧毁多少个发射井，才能以90%的概率保证对方的真井都被摧毁.

§2.1 离散型随机变量

解 假设摧毁A国的n个发射井就可以达到目的. 用X表示这n个井中的真井个数, 则$X \sim H(100, 10, n)$. 于是要求n使得

$$p_n = P(X = 10) = \frac{C_{10}^{10} C_{90}^{n-10}}{C_{100}^{n}} \geqslant 0.9.$$

可以计算出

n	99	98	97	96	95	94	$\leqslant 93$
p_n	0.900	0.809	0.726	0.652	0.584	0.522	$\leqslant 0.47$

所以必须摧毁99个发射井才能达到目的.

§2.1 离散型随机变量

(5) 几何分布(Geometric 分布)：设$p+q=1$, $pq>0$. 如果随机变量X有概率分布

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

则称X服从参数为p的几何分布.

例2.1.5 甲向一个目标射击，直到击中为止. 用X表示首次击中目标时的射击次数. 如果甲每次击中目标的概率是p, 则X服从几何分布.

解 用A_j表示甲第j次没击中目标，由$\{A_j\}$的独立性得到

$$\begin{aligned} P(X=k) &= P(A_1A_2\cdots A_{k-1}\overline{A}_k) \\ &= q^{k-1}p, \quad k=1,2,\cdots. \end{aligned}$$

§2.1 离散型随机变量

(5) 几何分布(Geometric 分布)： 设$p+q=1$, $pq>0$. 如果随机变量X有概率分布

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

则称X服从参数为p的几何分布.

例2.1.5 甲向一个目标射击，直到击中为止. 用X表示首次击中目标时的射击次数. 如果甲每次击中目标的概率是p，则X服从几何分布.

解 用A_j表示甲第j次没击中目标，由$\{A_j\}$的独立性得到

$$\begin{aligned} P(X=k) &= P(A_1A_2\cdots A_{k-1}\overline{A}_k) \\ &= q^{k-1}p,\quad k=1,2,\cdots. \end{aligned}$$

§2.1 离散型随机变量

(5) 几何分布(Geometric 分布)： 设$p+q=1$, $pq>0$. 如果随机变量X有概率分布

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

则称X服从参数为p的几何分布.

例2.1.5 甲向一个目标射击, 直到击中为止. 用X表示首次击中目标时的射击次数. 如果甲每次击中目标的概率是p, 则X服从几何分布.

解 用A_j表示甲第j次没击中目标, 由$\{A_j\}$的独立性得到

$$\begin{aligned} P(X=k) &= P(A_1A_2\cdots A_{k-1}\overline{A}_k) \\ &= q^{k-1}p,\quad k=1,2,\cdots. \end{aligned}$$

§2.2 连续型随机变量

定义2.2.1 设$F(x)=P(X\leqslant x)$是随机变量X的分布函数，如果有非负函数$f(x)$使得

$$F(x)=\int_{-\infty}^{x} f(s)\,\mathrm{d}s,\ \ x\in(-\infty,\infty), \tag{2.2.1}$$

则称$f(x)$是X或$F(x)$的**概率密度**(probability density)，简称为**密度**.

容易看出，$F(x)$ 有概率密度时，作为变上限的积分，$F(x)$ 是连续函数. 所以，当分布函数$F(x)$ 不是连续函数时，X没有概率密度. 概率密度通过(2.2.1)唯一决定分布函数.

§2.2 连续型随机变量

定义2.2.1 设$F(x)=P(X\leqslant x)$是随机变量X的分布函数，如果有非负函数$f(x)$使得

$$F(x)=\int_{-\infty}^{x} f(s)\,\mathrm{d}s,\ \ x\in(-\infty,\infty), \qquad (2.2.1)$$

则称$f(x)$是X或$F(x)$的概率密度(probability density)，简称为密度.

容易看出，$F(x)$ 有概率密度时，作为变上限的积分，$F(x)$ 是连续函数. 所以，当分布函数$F(x)$ 不是连续函数时，X没有概率密度. 概率密度通过(2.2.1)唯一决定分布函数.

§2.2 连续型随机变量

定理2.2.1 设X的分布函数$F(x)$连续,且在任何有限区间(a,b)中除去有限个点外有连续的导数，则

$$f(x)=\begin{cases} F'(x), & \text{当} F'(x)\text{存在}, \\ 0, & \text{否则} \end{cases} \qquad (2.2.2)$$

是X的概率密度.

§2.2　连续型随机变量

设$F(x)=P(X\leqslant x)$是X的分布函数，$f(x)$是$F(x)$的概率密度. 对于任何常数$a<b$,

$$\begin{aligned} P(a<X\leqslant b) &= P(X\leqslant b)-P(X\leqslant a) \\ &= \int_{-\infty}^{b} f(s)\,\mathrm{d}s-\int_{-\infty}^{a} f(s)\,\mathrm{d}s \\ &= \int_{a}^{b} f(s)\,\mathrm{d}s. \end{aligned}$$

由此知道当数集A是两个或若干个区间的并时，有

$$P(X\in A)=\int_{A} f(s)\,\mathrm{d}s.$$

§2.2 连续型随机变量

在放射性物质释放α粒子的例2.1.1中，用X表示等待第一个粒子释放的时间，则X是随机变量，取值可以是任何正数. 我们把有类似性质的随机变量称为连续型随机变量.

按照定义2.2，如果X有分布函数$F(x)$和概率密度$f(x)$，则有

$$F(x) = \int_{-\infty}^{x} f(s)\,\mathrm{d}s,\ \ x \in (-\infty, \infty).$$

定义2.2.3 如果随机变量X有概率密度$f(x)$，则称X是连续型随机变量(continuous random variable).

§2.2 连续型随机变量

在放射性物质释放α 粒子的例2.1.1中，用X表示等待第一个粒子释放的时间，则X 是随机变量，取值可以是任何正数．我们把有类似性质的随机变量称为连续型随机变量．

按照定义2.2，如果X有分布函数$F(x)$和概率密度$f(x)$，则有

$$F(x)=\int_{-\infty}^{x} f(s)\,\mathrm{d}s,\ \ x\in(-\infty,\infty).$$

定义2.2.3 如果随机变量X有概率密度$f(x)$，则称X是**连续型随机变量**(continuous random variable)．

§2.2 连续型随机变量

连续型随机变量X及其概率密度$f(x)$有如下的基本性质:

(1) $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$;

(2) $P(X = a) = 0$, 于是$P(a \leqslant X \leqslant b) = P(a < X \leqslant b)$;

(3) 对数集A, $P(X \in A) = \int_A f(x)\,\mathrm{d}x$;

(4) $F(x) = \int_{-\infty}^{x} f(s)\,\mathrm{d}s$是连续函数.

§2.2 连续型随机变量

下面只推导性质(1)和(2). 按照分布函数的定义, 有

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = F(\infty) = 1.$$

所以(1)成立. 再由

$$P(X = a) \leqslant P(X \in (a - \varepsilon, a]) = \int_{a-\varepsilon}^{a} f(x)\,\mathrm{d}x \to 0, \text{ 当 } \varepsilon \to 0,$$

知道(2)成立.

§2.2 连续型随机变量

下面介绍几个常见的连续型随机变量.

(1) 均匀分布(Uniform 分布) $\mathcal{U}(a,b)$: 对$a<b$, 如果X的概率密度是

$$f(x)=\begin{cases}\frac{1}{b-a}, & x\in(a,b),\\ 0, & x\notin(a,b),\end{cases} \tag{2.2.3}$$

则称X服从区间(a,b)上的均匀分布, 记做$X\sim\mathcal{U}(a,b)$. 这里$\mathcal{U}$ 是Uniform的缩写.

明显, 表达式(2.2.3)中的区间(a,b)也可以写成$(a,b]$, $[a,b)$或$[a,b]$. 为了方便, 还可以把X的概率密度简写成

$$f(x)=\frac{1}{b-a},\ x\in(a,b).$$

§2.2 连续型随机变量

容易理解，当X在(a, b)中均匀分布时，X落在(a, b)的子区间内的概率和这个子区间的长度成正比. 反之，如果X落在(a, b)的子区间的概率和这个子区间的长度成正比，则 X 在 (a, b) 中均匀分布.

当我们说在区间(a, b)中任掷一点，用X表示质点的落点时，意指X在(a, b)中均匀分布. 当我们说某人在时间段(a, b)中随机到达时，也是指他的到达时间X 在(a, b)中均匀分布.

§2.2 连续型随机变量

例2.2.1 每天的整点(如9点, 10点, 11点等)甲站都有列车发往乙站. 一位要去乙站的乘客在9点至10点之间随机到达甲站. 计算他候车时间小于30 min的概率.

解 用X表示他的到达时刻, 则X在0至60 min内均匀分布, 有概率密度

$$f(x)=\frac{1}{60},\ x\in(0,60).$$

用Y表示他的候车时间(单位:min), 则 $\{Y<30\}=\{30<X\leqslant 60\}$ 表示该乘客在9:30至10:00之间到达, 于是

$$P(Y<30)=P(30<X\leqslant 60)=\int_{30}^{60} f(x)\,\mathrm{d}x=\frac{1}{2}.$$

§2.2 连续型随机变量

例2.2.1 每天的整点(如9点, 10点, 11点等)甲站都有列车发往乙站. 一位要去乙站的乘客在9点至10点之间随机到达甲站. 计算他候车时间小于30 min的概率.

解 用X表示他的到达时刻, 则X在0至60 min内均匀分布, 有概率密度

$$f(x) = \frac{1}{60}, \quad x \in (0, 60).$$

用Y表示他的候车时间(单位:min), 则 $\{Y < 30\} = \{30 < X \leqslant 60\}$ 表示该乘客在9:30至10:00之间到达, 于是

$$P(Y < 30) = P(30 < X \leqslant 60) = \int_{30}^{60} f(x)\,\mathrm{d}x = \frac{1}{2}.$$

§2.2 连续型随机变量

(2) 指数分布(Exponential 分布) $\mathcal{E}(\lambda)$: 对正常数λ, 如果X的概率密度是

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geqslant 0, \\ 0, & x < 0, \end{cases} \qquad (2.2.4)$$

则称X服从参数为λ的指数分布, 记做$X \sim \mathcal{E}(\lambda)$. 这里$\mathcal{E}$是Exponential的缩写.

通常还把(2.2.4)简记为

$$f(x) = \lambda e^{-\lambda x}, \quad x \geqslant 0.$$

§2.2 连续型随机变量

图2.2.1是概率密度(2.2.4)的图形. 纵轴上从高至低分别是$\lambda=1.2$, $\lambda=0.6$, $\lambda=0.3$的概率密度.

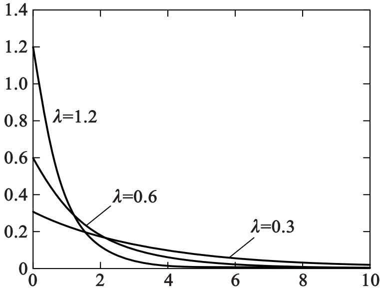


图2.2.1 $\lambda=1.2$, $\lambda=0.6$, $\lambda=0.3$

如果随机变量X使得$P(X<0)=0$, 则称X是非负随机变量. 容易看出, 如果$X\sim\mathcal{E}(\lambda)$, 则$X$是非负随机变量.

§2.2 连续型随机变量

定理2.2.2 设X是连续型非负随机变量, 则X服从指数分布的充分必要条件是对任何$s, t \geqslant 0$, 有

$$P(X > s + t | X > s) = P(X > t). \tag{2.2.5}$$

性质(2.2.5) 称为无记忆性或无后效性, 它是指数分布的特征.

如果X表示某仪器的工作寿命, 无记忆性的解释是: 当仪器工作了s小时后再能继续工作t小时的概率等于该仪器刚开始就能工作t小时的概率. 说明该仪器的剩余寿命不随使用时间的增加发生变化, 或说仪器是“永葆青春”的.

一般来说, 电子元件和计算机软件等具备这种性质, 它们本身的老化是可以忽略不计的, 造成损坏的原因是意外的高电压、计算机病毒等等.

§2.2 连续型随机变量

定理2.2.2 设X是连续型非负随机变量，则X服从指数分布的充分必要条件是对任何$s,t\geqslant 0$，有

$$P(X>s+t|X>s)=P(X>t). \qquad (2.2.5)$$

性质(2.2.5) 称为无记忆性或无后效性，它是指数分布的特征.

如果X表示某仪器的工作寿命，无记忆性的解释是：当仪器工作了s小时后再能继续工作t小时的概率等于该仪器刚开始就能工作t小时的概率. 说明该仪器的剩余寿命不随使用时间的增加发生变化，或说仪器是“永葆青春”的.

一般来说，电子元件和计算机软件等具备这种性质，它们本身的老化是可以忽略不计的，造成损坏的原因是意外的高电压、计算机病毒等等.

§2.2 连续型随机变量

定理2.2.2 设X是连续型非负随机变量，则X服从指数分布的充分必要条件是对任何$s,t\geqslant 0$，有

$$P(X>s+t|X>s)=P(X>t). \qquad (2.2.5)$$

性质(2.2.5) 称为无记忆性或无后效性，它是指数分布的特征.

如果X表示某仪器的工作寿命，无记忆性的解释是：当仪器工作了s小时后再能继续工作t小时的概率等于该仪器刚开始就能工作t小时的概率. 说明该仪器的剩余寿命不随使用时间的增加发生变化，或说仪器是“永葆青春”的.

一般来说，电子元件和计算机软件等具备这种性质，它们本身的老化是可以忽略不计的，造成损坏的原因是意外的高电压、计算机病毒等等.

§2.2 连续型随机变量

例2.2.2 用X_1表示例2.1.1中从开始至观测到第1个α 粒子的等待时间, 证明 X_1 服从指数分布.

证明 用$N(t)$表示时间段$(0, t]$内观测到的α 粒子数. 按例2.1中的推导, $N(t) \sim \mathcal{P}(\mu t)$, 其中$\mu$是正常数. 于是用$\{X_1 > t\} = \{N(t) = 0\}$得到

$$P(X_1 > t) = P(N(t) = 0) = \mathrm{e}^{-\mu t}.$$

所以, 对任何$t > 0$, X_1有分布函数

$$F(t) = P(X_1 \leqslant t) = 1 - P(X_1 > t) = 1 - \mathrm{e}^{-\mu t}. \tag{2.2.6}$$

对$F(t)$求导数得到概率密度$f(t) = \mu \mathrm{e}^{-\mu t}$, $t > 0$. 说明$X_1 \sim \mathcal{E}(\mu)$.

§2.2 连续型随机变量

仅从这个例子的背景也可以体会出等待第一个α 粒子释放的时间X是服从指数分布的. 因为X是从开始观测到有一个粒子释放出的时间, 其概率分布与前一个粒子的释放时间无关, 所以X有无记忆性.

可以想象, 如果X_2是从观测到第一个α 粒子开始到观测到第二个α 粒子的间隔时间, 则X_2也服从指数分布. 这一点可以得到理论证明, 不再赘述.

§2.2 连续型随机变量

仅从这个例子的背景也可以体会出等待第一个α 粒子释放的时间X是服从指数分布的. 因为X是从开始观测到有一个粒子释放出的时间, 其概率分布与前一个粒子的释放时间无关, 所以X有无记忆性.

可以想象, 如果X_2是从观测到第一个α 粒子开始到观测到第二个α 粒子的间隔时间, 则X_2也服从指数分布. 这一点可以得到理论证明, 不再赘述.

§2.2 连续型随机变量

(3) 正态分布(Normal 分布)$N(\mu,\sigma^2)$: 设μ是常数, σ是正常数. 如果X的概率密度是

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\Big(-\frac{(x-\mu)^2}{2\sigma^2}\Big),\quad x\in\mathcal{R}, \tag{2.2.7}$$

则称X服从参数为(μ,σ^2)的正态分布, 记做$X\sim N(\mu,\sigma^2)$. 这里N是Normal的缩写. 特别, 当$X\sim N(0,1)$时, 称X服从标准正态分布(standard normal distribution). 标准正态分布的概率密度有特殊的重要地位, 所以用一个特定的符号φ表示:

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}\exp\Big(-\frac{x^2}{2}\Big),\quad x\in\mathcal{R}.$$

§2.2 连续型随机变量

图2.2.2是 $\mu = 0$ 时正态概率密度 (2.2.7) 的图形. 纵轴上从低至高分别是 $\sigma = 2$, $\sigma = 1$, $\sigma = 0.5$ 的概率密度.

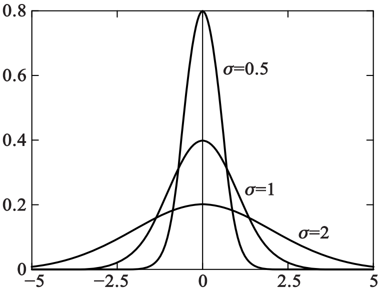


图2.2.2 $\sigma = 2$, $\sigma = 1$, $\sigma = 0.5$

§2.2 连续型随机变量

正态分布在概率论和数理统计中有特殊的重要地位. 事实表明,产品的许多质量指标, 生物和动物的许多生理指标等都服从或近似服从正态分布. 大量相互独立且有相同分布的随机变量的累积也近似服从正态分布(参考二项分布的图形).

正态概率密度有如下的简单性质:

(1) $f(x)$关于$x=\mu$对称;

(2) $f(\mu)=1/\sqrt{2\pi}\sigma$是最大值.

§2.2 连续型随机变量

正态分布在概率论和数理统计中有特殊的重要地位. 事实表明,产品的许多质量指标, 生物和动物的许多生理指标等都服从或近似服从正态分布. 大量相互独立且有相同分布的随机变量的累积也近似服从正态分布(参考二项分布的图形).

正态概率密度有如下的简单性质:

(1) $f(x)$关于$x=\mu$对称;

(2) $f(\mu)=1/\sqrt{2\pi}\sigma$是最大值.

§2.2 连续型随机变量

对$X \sim N(\mu, \sigma^2)$, 在许多的应用问题中会遇到计算概率$P(X \leqslant a)$ 的问题. 为了方便, 人们已经习惯于用

$$\Phi(x) = \int_{-\infty}^{x} \varphi(s)\,\mathrm{d}s$$

表示标准正态分布$N(0,1)$的分布函数. $\Phi(x)$的取值可以查附表得到. 利用$\varphi(x)$的对称性得到

$$\Phi(x) + \Phi(-x) = 1, \quad x \in \mathcal{R}.$$

§2.2 连续型随机变量

并且，只要$X \sim N(\mu, \sigma^2)$，则有

$$\begin{aligned} P(X \leqslant a) &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{a} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(a-\mu)/\sigma} \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x \\ &= \Phi\left(\frac{a-\mu}{\sigma}\right). \end{aligned}$$

于是对任何$a < b$，当$X \sim N(\mu, \sigma^2)$，用

$$P(a < X \leqslant b) = P(X \leqslant b) - P(X \leqslant a)$$

得到公式

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

§2.2 连续型随机变量

例2.2.3 一台机床加工的部件长度服从正态分布 $N(10, 36\times 10^{-6})$. 当部件的长度在10 ± 0.01 内为合格品，求一个部件是合格品的概率.

解 用X 表示部件的长度，则$X\sim N(10, 36\times 10^{-6})$. 事件

$$\{10-0.01\leqslant X\leqslant 10+0.01\}$$

表示这个部件是合格品.

§2.2 连续型随机变量

例2.2.3 一台机床加工的部件长度服从正态分布 $N(10, 36\times 10^{-6})$. 当部件的长度在10 ± 0.01 内为合格品, 求一个部件是合格品的概率.

解 用X 表示部件的长度, 则$X\sim N(10, 36\times 10^{-6})$. 事件

$$\{10-0.01\leqslant X\leqslant 10+0.01\}$$

表示这个部件是合格品.

§2.2 连续型随机变量

$$
\begin{aligned}
& P(10-0.01 \leqslant X \leqslant 10+0.01) \\
= \quad & \Phi\left(\frac{0.01}{6 \times 10^{-3}}\right)-\Phi\left(\frac{-0.01}{6 \times 10^{-3}}\right) \\
= \quad & \Phi(1.67)-\Phi(-1.67) \\
= \quad & 2\Phi(1.67)-1 \\
= \quad & 2 \times 0.9525-1=0.905.
\end{aligned}
$$

这个概率太不令人满意了，说明这台机床的质量有问题. 以后会知道质量问题是由参数σ 控制的. σ越小，质量越好.

§2.2 连续型随机变量

$$
\begin{aligned}
&P(10-0.01 \leqslant X \leqslant 10+0.01) \\
&\quad = \Phi\left(\frac{0.01}{6\times 10^{-3}}\right) - \Phi\left(\frac{-0.01}{6\times 10^{-3}}\right) \\
&\quad = \Phi(1.67) - \Phi(-1.67) \\
&\quad = 2\Phi(1.67) - 1 \\
&\quad = 2\times 0.9525 - 1 = 0.905.
\end{aligned}
$$

这个概率太不令人满意了，说明这台机床的质量有问题. 以后会知道质量问题是由参数σ 控制的. σ越小，质量越好.

§2.2 连续型随机变量

为了介绍伽马分布，先介绍Γ函数. 这里Γ是Gamma 的简写. 伽马函数$\Gamma(\alpha)$由积分

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} \mathrm{e}^{-x}\,\mathrm{d}x, \ \ \alpha > 0$$

定义. 对正数α 和正整数n, 容易验证如下的基本性质:

$$\Gamma(1+\alpha) = \alpha\Gamma(\alpha), \ \Gamma(n) = (n-1)!, \ \Gamma(1/2) = \sqrt{\pi}.$$

§2.2 连续型随机变量

(4) 伽马分布(Gamma 分布) $\Gamma(\alpha, \beta)$: 设α, β 是正常数，如果X的概率密度是

$$f(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \mathrm{e}^{-\beta x}, & x \geqslant 0, \\ 0, & x < 0, \end{cases}$$

则称X服从参数是(α, β)的伽马分布，记做$X \sim \Gamma(\alpha, \beta)$.

§2.2 连续型随机变量

图2.2.3 是 $\Gamma(\alpha, 2)$ 分布的概率密度图，按概率密度的最大值由大到小依次排列的是$\alpha = 1$，2，3, 6, 10时的图形.

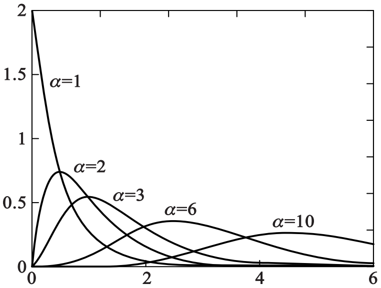


图2.2.3 $\beta = 2$, $\alpha = 1$，2，3，6，10

§2.2 连续型随机变量

英国著名统计学家皮尔逊(Pearson)在研究物理、生物及经济中的随机变量时，发现很多连续型随机变量的分布都不是正态分布. 这些随机变量的特点是只取非负值，于是他致力于这类随机变量的研究.

从1895—1916年间，皮尔逊连续发表了一系列的概率密度曲线，认为这些曲线可以包括常见的单峰分布，其中就有伽马分布. 在气象学中，干旱地区的年、季或月降水量被认为服从伽马分布，指定时间段内的最大风速等也被认为服从伽马分布.

§2.3 随机变量函数的分布

先通过下面的例子学习计算随机变量函数的分布的一般方法.

例2.3.1 设X 有如下的概率分布

X	-2	-1	0	1	2	3
P	0.2	0.2	0.3	0.1	0.1	0.1

求$Y=X^2$的分布.

§2.3 随机变量函数的分布

解 Y的取值是0, 1, 4, 9，而且

$$
\begin{aligned}
P(Y=0) &= P(X=0)=0.3,\\
P(Y=1) &= P(X^2=1)=P(X=-1)+P(X=1)\\
&= 0.2+0.1=0.3,\\
P(Y=4) &= P(|X|=2)=P(X=-2)+P(X=2)\\
&= 0.2+0.1=0.3,\\
P(Y=9) &= P(X=3)=0.1.
\end{aligned}
$$

于是Y有概率分布

Y	0	1	4	9
P	0.3	0.3	0.3	0.1

§2.3 随机变量函数的分布

下面介绍连续型随机变量函数的概率密度的计算方法.

设随机变量X有分布函数$F(x)$, 用$\mathrm{d}x$表示x的微分. 从微积分的知识知道当$F'(x)$在x连续时, 有

$$
\begin{aligned}
P(X=x) &= \lim_{\Delta x \to 0} |F(x)-F(x-\Delta x)| \\
&= \mathrm{d}F(x) \\
&= F'(x)\,\mathrm{d}x.
\end{aligned}
$$

所以我们用

$$P(X=x)=g(x)\,\mathrm{d}x$$

表示X在x有概率密度$g(x)$. 如果D是开区间或开区间的并集, 则称D是开集.

§2.3 随机变量函数的分布

定理2.3.1 如果开集D使得$P(X\in D)=1$, $g(x)$在D中连续, 使得

$$P(X=x)=g(x)\,\mathrm{d}x,\ \ x\in D,$$

则X有概率密度

$$f(x)=g(x),\ \ x\in D.$$

如果$h(y)$在y可微, $F(x)$在$x=h(y)$有连续的导数$f(h(y))$, 则对$h(y)=x$, 用(4.1) 得到

$$P(X=h(y))=|f(h(y))\mathrm{d}h(y)|=f(h(y))|h'(y)|\,\mathrm{d}y. \qquad (2.3.1)$$

在使用定理2.3.1计算随机变量函数的概率密度时,公式(2.3.1)是非常有用的.

§2.3 随机变量函数的分布

定理2.3.1 如果开集D使得$P(X \in D) = 1$, $g(x)$在D中连续，使得

$$P(X = x) = g(x)\,\mathrm{d}x,\ \ x \in D,$$

则X有概率密度

$$f(x) = g(x),\ \ x \in D.$$

如果$h(y)$在y可微，$F(x)$在$x = h(y)$有连续的导数$f(h(y))$，则对$h(y) = x$，用(4.1) 得到

$$P(X = h(y)) = |f(h(y))\mathrm{d}h(y)| = f(h(y))|h'(y)|\,\mathrm{d}y. \qquad (2.3.1)$$

在使用定理2.3.1计算随机变量函数的概率密度时,公式(2.3.1)是非常有用的.

§2.3 随机变量函数的分布

定理2.3.1 如果开集D使得$P(X \in D)=1$, $g(x)$在D中连续,使得

$$P(X=x)=g(x)\,\mathrm{d}x,\ \ x \in D,$$

则X有概率密度

$$f(x)=g(x),\ \ x \in D.$$

如果$h(y)$在y可微, $F(x)$在$x=h(y)$有连续的导数$f(h(y))$, 则对$h(y)=x$, 用(4.1) 得到

$$P(X=h(y))=|f(h(y))\mathrm{d}h(y)|=f(h(y))|h'(y)|\,\mathrm{d}y. \qquad (2.3.1)$$

在使用定理2.3.1计算随机变量函数的概率密度时,公式(2.3.1)是非常有用的.

§2.3 随机变量函数的分布

用定理2.3.1计算随机变量函数的概率密度的方法被称为微分法. 在使用微分法时, 需要遵守以下的约定: 当且仅当$A=B$时, 可用公式$P(A)=P(B)$; 当且仅当$A=\bigcup_{i=1}^{n} A_i$,且$A_1, A_2, \cdots, A_n$作为集合互不相交时,可用公式

$$P(A)=\sum_{i=1}^{n} P(A_i).$$

§2.3 随机变量函数的分布

例2.3.2 设$X \sim N(\mu, \sigma^2)$, 则

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1).$$

解 因为X有概率密度

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right],$$

§2.3 随机变量函数的分布

例2.3.2 设$X \sim N(\mu, \sigma^2)$, 则

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1).$$

解 因为X有概率密度

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2} \right],$$

§2.3　随机变量函数的分布

且对任何z,有

$$\begin{aligned} P(Z=z) &= P\Big(\frac{X-\mu}{\sigma}=z\Big) \\ &= P(X=\sigma z+\mu) \\ &= f_X(\sigma z+\mu)\,|\,\mathrm{d}(\sigma z+\mu)| \\ &= \frac{1}{\sqrt{2\pi}\sigma}\exp\Big[-\frac{(\sigma z+\mu-\mu)^2}{2\sigma^2}\Big]\sigma\,\mathrm{d}z \\ &= \frac{1}{\sqrt{2\pi}}\exp\Big(-\frac{z^2}{2}\Big)\,\mathrm{d}z, \end{aligned}$$

§2.3 随机变量函数的分布

所以Z有概率密度

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right).$$

这正是标准正态分布的概率密度.

例2.3.3 设$X \sim N(0,1)$, $b \neq 0$, 求$Y = a + bX$的概率密度.

解 因为X有概率密度

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right),$$

§2.3 随机变量函数的分布

所以Z有概率密度

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right).$$

这正是标准正态分布的概率密度.

例2.3.3 设$X \sim N(0,1)$, $b \neq 0$, 求$Y = a + bX$的概率密度.

解 因为X有概率密度

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right),$$

§2.3 随机变量函数的分布

且对任何y,有

$$\begin{aligned}
P(Y=y) &= P(a+bX=y) \\
&= P\left(X=\frac{y-a}{b}\right) \\
&= \varphi\left(\frac{y-a}{b}\right)\left|\mathrm{d}\left(\frac{y-a}{b}\right)\right| \\
&= \frac{1}{\sqrt{2\pi}\,|b|}\exp\left[-\frac{(y-a)^2}{2b^2}\right]\mathrm{d}y.
\end{aligned}$$

所以$Y\sim N(a,b^2)$, 有概率密度

$$f_Y(y)=\frac{1}{\sqrt{2\pi}\,|b|}\exp\left[-\frac{(y-a)^2}{2b^2}\right],\ y\in(-\infty,\infty).$$

§2.3 随机变量函数的分布

用X表示某种产品的使用寿命，$F_X(x) = P(X \leqslant x)$是$X$的分布函数. 在产品的可靠性研究中，人们称产品的使用寿命X大于某固定值a的概率

$$P(X > a) = 1 - F_X(a)$$

为该产品的可靠性. 如果产品的使用寿命X服从指数分布$\mathcal{E}(\lambda)$，则已经使用了一段时间的旧产品和新产品有相同的可靠性.

在实际中的大多数场合，人们都不愿意使用旧产品,也就是说，人们并不认为产品的使用寿命服从指数分布. 为了更加合理地描述产品的使用寿命，可以对X进行改造. 设a, b是正数，$X \sim \mathcal{E}(1)$，定义

$$Y = (X/a)^{1/b}.$$

现在用 Y 表示产品的使用寿命，并称 Y 的分布为韦布尔(Weibull) 分布.

§2.3 随机变量函数的分布

实际经验表明许多电子元件和机械设备的使用寿命都可以用韦布尔分布描述. 另外, 凡是由局部部件的失效或故障会引起全局停止运行的设备的寿命也都常用韦布尔分布近似. 特别是金属材料(如轴承等)的疲劳寿命被认为是服从韦布尔分布的.

§2.3 随机变量函数的分布

例2.3.4(韦布尔分布) 设X服从参数为$\lambda=1$的指数分布$\mathcal{E}(1)$, a, b是正常数. 计算$Y=(X/a)^{1/b}$的概率密度.

解 X有概率密度$f_X(x)=\mathrm{e}^{-x}$, $x>0$. 因为$P(Y>0)=1$, 且对$y>0$, 有

$$\begin{aligned} P(Y=y) &= P\big((X/a)^{1/b}=y\big) \\ &= P(X=ay^b) \\ &= \big|f_X(ay^b)\mathrm{d}(ay^b)\big| \\ &= \exp\big(-ay^b\big)aby^{b-1}\,\mathrm{d}y, \end{aligned}$$

所以Y的概率密度是

$$f_Y(y)=aby^{b-1}\exp\big(-ay^b\big),\ \ y>0.$$

这时称Y服从参数为(a,b)的韦布尔分布, 记做$Y\sim W(a,b)$.

§2.3 随机变量函数的分布

例2.3.4(韦布尔分布) 设X服从参数为$\lambda=1$的指数分布$\mathcal{E}(1)$, a, b是正常数. 计算$Y=(X/a)^{1/b}$的概率密度.

解 X有概率密度$f_X(x)=\mathrm{e}^{-x}$, $x>0$. 因为$P(Y>0)=1$, 且对$y>0$, 有

$$
\begin{aligned}
P(Y=y) &= P\big((X/a)^{1/b}=y\big) \\
&= P\big(X=ay^b\big) \\
&= \big|f_X(ay^b)\mathrm{d}(ay^b)\big| \\
&= \exp\big(-ay^b\big)aby^{b-1}\,\mathrm{d}y,
\end{aligned}
$$

所以Y的概率密度是

$$f_Y(y)=aby^{b-1}\exp\big(-ay^b\big),\ \ y>0.$$

这时称Y服从参数为(a,b)的韦布尔分布, 记做$Y\sim W(a,b)$.

§2.3 随机变量函数的分布

图2.3.1和2.3.2是韦布尔密度的图形. 横轴是y, 纵轴是$f_Y(y)$.

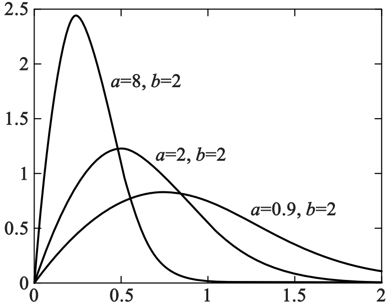


图2.3.1 $a = 8, 2, 0.9,\ b = 2$

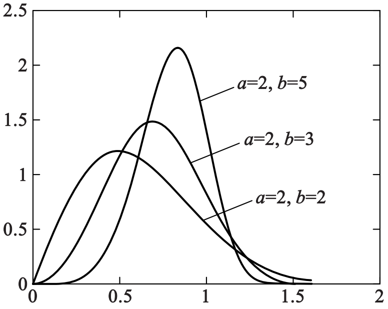


图2.3.2 $a = 2,\ b = 2, 3, 5$

§2.3 随机变量函数的分布

设$X \sim N(\mu, \sigma^2)$, 则称$Y = \mathrm{e}^X$的分布为对数正态分布. 在产品的使用寿命研究中, 人们经常用到对数正态分布. 实践表明, 在研究因化学或物理化学的缓慢变化造成的断裂或失效时(如绝缘体等), 用对数正态分布描述使用寿命是合适的.

按照定理2.3.1, 如果开集D使得$P(X \in D) = 1$, D中的连续函数$g(x)$使得

$$\frac{P(X = x)}{\mathrm{d}x} = g(x),\quad x \in D,$$

则X的概率密度是$g(x)$, $x \in D$.

§2.3 随机变量函数的分布

设$X \sim N(\mu, \sigma^2)$, 则称$Y = \mathrm{e}^X$的分布为对数正态分布. 在产品的使用寿命研究中, 人们经常用到对数正态分布. 实践表明, 在研究因化学或物理化学的缓慢变化造成的断裂或失效时(如绝缘体等), 用对数正态分布描述使用寿命是合适的.

按照定理2.3.1, 如果开集D使得$P(X \in D) = 1$, D中的连续函数$g(x)$使得

$$\frac{P(X = x)}{\mathrm{d}x} = g(x), \ \ x \in D,$$

则X的概率密度是$g(x)$, $x \in D$.

§2.3 随机变量函数的分布

例2.3.5 (对数正态分布) 设$X \sim N(\mu, \sigma^2)$, 计算$Y = \mathrm{e}^X$的概率密度.

解 易见$P(Y > 0) = 1$, 对$y > 0$, 利用

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

得到Y的概率密度

§2.3 随机变量函数的分布

例2.3.5 (对数正态分布) 设$X \sim N(\mu, \sigma^2)$, 计算$Y = \mathrm{e}^X$的概率密度.

解 易见$P(Y > 0) = 1$, 对$y > 0$, 利用

$$f_X(x) = \frac{1}{\sqrt{2\pi}\ \sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

得到Y的概率密度

§2.3 随机变量函数的分布

$$
\begin{aligned}
f_Y(y) &= \frac{P(Y=y)}{\mathrm{d}y} = \frac{P(\mathrm{e}^X=y)}{\mathrm{d}y} \\
&= \frac{P(X=\ln y)}{\mathrm{d}y} \\
&= \frac{\left|f_X(\ln y)\mathrm{d}(\ln y)\right|}{\mathrm{d}y} \\
&= \frac{1}{y} f_X(\ln y) \\
&= \frac{1}{\sqrt{2\pi}\,\sigma y} \exp\left[-\frac{(\ln y-\mu)^2}{2\sigma^2}\right], \quad y>0.
\end{aligned}
$$

这时称Y服从参数为(μ,σ^2)的对数正态分布.

§2.3 随机变量函数的分布

图2.3.3和2.3.4 是对数正态密度的图形, 横轴是y, 纵轴是$f_Y(y)$.

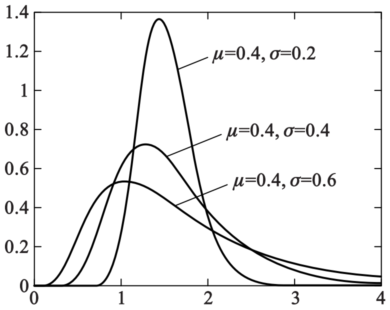


图2.3.3 $\mu = 0.4$, $\sigma = 0.2, 0.4, 0.6$

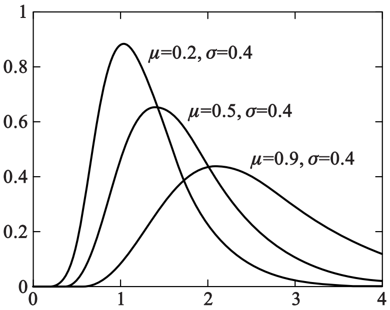


图2.3.4 $\mu = 0.2, 0.5, 0.9$, $\sigma = 0.4$

§2.3 随机变量函数的分布

例2.3.6 设$X \sim N(0,1)$, 计算$Y = X^2$的概率密度.

解 X有概率密度

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

因为$P(Y > 0) = 1$, 且对$y > 0$, 有

$$\begin{aligned} P(Y = y) &= P(X = \sqrt{y}) + P(X = -\sqrt{y}) \\ &= \left|\varphi(\sqrt{y})\mathrm{d}\sqrt{y}\right| + \left|\varphi(-\sqrt{y})\mathrm{d}\sqrt{y}\right| \\ &= \frac{1}{\sqrt{y}}\varphi(\sqrt{y})\,\mathrm{d}y. \end{aligned}$$

所以Y有概率密度

$$f_Y(y) = \frac{1}{\sqrt{y}}\varphi(\sqrt{y}) = \frac{1}{\sqrt{2\pi y}}\,\mathrm{e}^{-y/2}, \quad y > 0.$$