3.3 常用的离散型随机变量

本节介绍几种常见且重要的离散型随机变量

3.3.1 离散型均匀分布

定义3.5. 设随机变量X的取值$a_1, a_2, \cdots, a_n$, 且$P(X = a_i) = 1/n$, 称X服从离散型均匀分布.

由定义可知

$$\begin{aligned} E(X) &= \frac{1}{n}\sum_{i=1}^{n} a_i, \\ \mathrm{Var}(X) &= \frac{1}{n}\sum_{i=1}^{n} a_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} a_i\right)^2. \end{aligned}$$

例3.8 (德国坦克数量问题)**.** 二战期间假德国坦克编号为 $1, 2, \cdots, N$, 共生产 N 辆. 盟军战斗中随机击毁了 k 辆, 且所击毁的最大编号为 m. 如何估计 N 的大小.

解. 上述问题本质上是从 $1, 2, \cdots, N$ 中以'不放回'方式抽取 k 个数, 观察到 k 个数中最大数为 m, 如何利用 m 和 k 估计 N.

随机变量$X =${抽取 k 个数中的最大数} $(X = k, k+1, \cdots, N)$, 有

$$Pr[X = i] = \binom{i-1}{k-1}\Big/\binom{N}{k}, \quad \text{从而得到} \quad E(X) = \sum_{i=k}^{N} i\binom{i-1}{k-1}\Big/\binom{N}{k}.$$

现考虑从 $N+1$ 个元素中取 $k+1$ 个元素, 共有多少种不同的取法, 此问题可等价于按所抽取 $k+1$ 个元素中最大元进行分类, 即

$$\binom{N+1}{k+1} = \sum_{i=k}^{N}\binom{i}{k} = \sum_{i=k}^{N}\frac{i}{k}\binom{i-1}{k-1}$$

代入 $E(X)$ 可得

$$E(X) = \sum_{i=k}^{N} i\binom{i-1}{k-1}\Big/\binom{N}{k} = k\binom{N+1}{k+1}\Big/\binom{N}{k} = \frac{k}{k+1}(N+1).$$

由于仅做了一次观察, 可以将一次观察的最大值 m 看作为 $E[X]$ 的近似, 即

$$m = E(X) = \frac{k}{k+1}(N+1) \quad \Rightarrow \quad N = m(1 + k^{-1}) - 1$$

从而完成 N 的估计. □

例如, 如果观察到被击毁坦克编号分别为 $17, 68, 94, 127, 135, 212$, 根据上面的推到可估计出 $N = 212(1 + \frac{1}{6}) - 1 \approx 246$. 下表给出这种方法的估计效果:

由此可见: 统计估计比情报估计准确得多, 接近德国的实际产量.

时间	统计估计	英国情报估计	德国实际产量
1940-06	169	1000	122
1941-06	244	1550	271
1942-08	327	1550	342

3.3.2 0-1分布

定义3.6. *随机变量 X 的取值为 $\{0,1\}$, 其分布列 $P(X=1)=p$, $P(X=0)=1-p$, 称 X 服从参数为 p 的 $0-1$ 分布, 又称两点分布, 或Bernoulli分布, 记 $X\sim Ber(p)$.*

由定义可知

$$E(X)=p, \qquad Var(X)=p-p^2=p(1-p).$$

$0-1$分布是很多概率模型的基础.

3.3.3 二项分布

Bernoulli试验E只有两个结果: A 和 $\bar{A}$, 将试验E独立重复地进行 n 次, 称为 n 重Bernoulli试验. n 重Bernoulli试验是一种很重要的数学模型, 具有广泛的应用.

定义3.7. *在 n 重Bernoulli试验, 设事件A发生的概率为p, 用随机变量 X 表示 n 次独立试验中事件 A 发生的次数, X 的取值为$0,1,2,\ldots,n$, 其分布列为 $P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$ $(k=0,1,2,\cdots,n)$, 称 X 服从二项分布(binomial distribution), 记$X\sim B(n,p)$.*

由于分布列类似于二项展开式$(a+b)^n=\sum_{k=1}^{n}\binom{n}{k}a^kb^{n-k}$, 因此称为二项分布. 容易验证

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}\geq 0, \quad \sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}=1,$$

从而得到二项分布是一个分布列.

性质3.4. *若 $X\sim B(n,p)$, 则有*

$$E(X)=np, \qquad Var(X)=np(1-p).$$

Proof. 由定义可知:

$$E(X)=\sum_{k=0}^{n}k\binom{n}{k}p^k(1-p)^{n-k}=(1-p)^n\sum_{k=0}^{n}\binom{n}{k}k\left(\frac{p}{1-p}\right)^k.$$

为计算 $\sum_{k=0}^{n}\binom{n}{k}k\left(\frac{p}{1-p}\right)^k$, 考虑

$$(1+x)^n=\sum_{k=0}^{n}\binom{n}{k}x^k \Longrightarrow n(1+x)^{n-1}=\sum_{k=0}^{n}\binom{n}{k}x^{k-1}k \Longrightarrow nx(1+x)^{n-1}=\sum_{k=0}^{n}\binom{n}{k}x^kk$$

将 $x = p/(1-p)$带入上式可得

$$E(X) = (1-p)^n \sum_{k=0}^{n} \binom{n}{k} k \left(\frac{p}{1-p}\right)^k = (1-p)^n n \frac{p}{1-p} \frac{1}{(1-p)^{n-1}} = np.$$

对于方差, 首先计算

$$E(X^2) = \sum_{k=0}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=1}^{n} nk \binom{n-1}{k-1} p \cdot p^{k-1} (1-p)^{n-k}$$

令$i = k-1$, 带入上式可得

$$\begin{aligned} E(X^2) &= np \sum_{i=0}^{n-1} (i+1) \binom{n-1}{i} \cdot p^i (1-p)^{n-1-i} \\ &= np \sum_{i=0}^{n-1} i \binom{n-1}{i} \cdot p^i (1-p)^{n-1-i} + np \sum_{i=0}^{n-1} \binom{n-1}{i} \cdot p^i (1-p)^{n-1-i} \\ &= np(n-1)p + np. \end{aligned}$$

从而得到$\mathrm{Var}(X) = E[X^2] - (E[X])^2 = np(1-p)$. □

例3.9. *一张试卷上有5道选择题, 每道选择题有4个答案, 只有一个是正确的, 某学生靠随机猜测至少能做对4个题的概率是多少?*

解. 每答一道题相当于一次Bernoulli试验, 事件 $A = \{$答对一题$\}$, 有 $P(A) = 1/4$. 5道题相当于5重Bernoulli试验. 用 X 表示学生猜对的题数, 则 $X \sim B(5, 1/4)$, 从而得到

$$P(X \geq 4) = P(X = 4) + P(X = 5) = \binom{5}{4} \left(\frac{1}{4}\right)^4 \cdot \frac{3}{4} + \left(\frac{1}{4}\right)^5 = \frac{1}{64}.$$

□

3.3.4 几何分布

定义3.8. *在多重Bernoulli试验, 设事件A发生的概率为p, 用随机变量 X 表示事件 A 首次发生时的试验次数, X 的取值为$1, 2, \ldots$, 其分布列为 $P(X = k) = (1-p)^{k-1} p$ $(k = 1, 2, \ldots)$, 称 X 服从参数为 p 的几何分布, 记$X \sim G(p)$.*

容易得到 $P(X = k) = (1-p)^{k-1} p \geq 0$, 以及

$$\sum_{k=1}^{\infty} P(X = k) = p \sum_{k=1}^{\infty} (1-p)^{k-1} = p \frac{1}{1-(1-p)} = 1,$$

从而验证了几何分布构成一个分布列.

性质3.5. *若随机变量 $X \sim G(P)$, 则有*

$$E(X) = \frac{1}{p}, \quad Var(X) = \frac{1-p}{p^2}.$$

Proof. 根据期望的定义有

$$E(X)=\sum_{k=1}^{\infty}kP(X=k)=\sum_{k=1}^{\infty}kp(1-p)^{k-1}=p\sum_{k=1}^{\infty}k(1-p)^{k-1}.$$

根据公式

$$\sum_{k=0}^{\infty}x^k=\frac{1}{1-x}\quad\Rightarrow\quad\sum_{k=1}^{\infty}kx^{k-1}=\frac{1}{(1-x)^2}.$$

令$x=1-p$可得$E(X)=1/p$. 对于随机变量 X 的方差, 首先计算

$$E(X^2)=\sum_{k\geq 1}k^2p(1-p)^{k-1}=p\sum_{k\geq 1}k^2(1-p)^{k-1}=p\cdot\frac{2-p}{p^3}=\frac{2-p}{p^2}.$$

根据公式

$$\sum_{k=0}^{\infty}x^k=\frac{1}{1-x}\quad\Rightarrow\quad\sum_{k=1}^{\infty}kx^{k-1}=\frac{1}{(1-x)^2}\quad\Rightarrow\quad\sum_{k=1}^{\infty}kx^k=\frac{x}{(1-x)^2}\quad\Rightarrow\quad\sum_{k=1}^{\infty}k^2x^{k-1}=\frac{1+x}{(1-x)^3}.$$

令 $x=1-p$ 可得 $E(X^2)=(2-p)/p^2$, 于是有 $\mathrm{Var}(X)=E(X^2)-(EX)^2=(1-p)/p^2$. □

下面研究几何分布的特性: 无记忆性(memoryless property).

定理3.4. *设$X\sim G(P)$, 对任意正整数m,n, 有*

$$P(X>m+n|X>m)=P(X>n).$$

直观理解: 假设已经经历了 m 次失败, 从当前起直至成功的次数与 m 无关.

Proof. 对任何正整数 k, 根据几何分布的定义有

$$P(X>k)=\sum_{i=k+1}^{\infty}p(1-p)^{i-1}=p\sum_{i=k+1}^{\infty}(1-p)^{i-1}=p\frac{(1-p)^k}{1-(1-p)}=(1-p)^k.$$

根据条件概率的定义有

$$P(X>m+n|X>m)=\frac{P(X>m+n)}{P(X>m)}=\frac{(1-p)^{m+n}}{(1-p)^m}=(1-p)^n=P(X>n)$$

这里利用$\{X>m+n\}\cap\{X>m\}=\{X>m+n\}$. □

例3.10. *古代有一村落因为资源限制不能养太多的人, 每个家庭都很重视男性, 于是制定一条规则: 每个家庭可以生一个男孩, 如果没有男孩则可以继续生育直至有一个男孩为止. 若已有一个男孩, 则不再生育. 若生男孩的概率为p, 问多年后该村的男女比例是多少?*

解. 对一个家庭而言, 用随机变量 X 表示该家庭的小孩个数, 则 $X=1,2,\ldots$, 以及

$$P(X=i)=p(1-p)^{i-1},$$

即 $X\sim G(p)$. 于是一个家庭的小孩数的期望是 $E[X]=1/p$, 从而得到一个家庭中小孩的男/女比例为 $1:(1/p-1)$, 特别地, 当 $p=1/2$ 时男女比例 $1:1$. □

3.3.5 泊松分布

定义3.9. 如果随机变量X的分布列为

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

其中 $\lambda>0$ 是一个给定的常数, 称随机变量X服从参数为λ的泊松分布, 记为 $X\sim P(\lambda)$.

容易验证$P(X=k)=\lambda^k e^{-\lambda}/k!\geq 0$, 根据泰勒展式$e^\lambda=\sum_{k=0}^{\infty}\lambda^k/k!$有

$$\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=e^{-\lambda}\cdot e^{\lambda}=1.$$

泊松分布用于描述大量试验中稀有事件出现次数的概率模型, 例如:

- 电话在一段时间内收到的呼叫次数,
- 放射物在一段时间内放射的粒子数,
- 一段时间内通过某路口的出租车数,
- 一本书中一页出现的语法错误数,
- 一天内道一所银行办理业务的顾客数.

性质3.6. 对任意给定的 $\lambda>0$, 若 $X\sim P(\lambda)$, 则

$$E(X)=\lambda, \qquad Var(X)=\lambda.$$

Proof. 根据期望的定义有

$$E(X)=\sum_{i=0}^{\infty}i\cdot P(X=i)=\sum_{i=1}^{\infty}i\cdot\frac{\lambda^i}{i!}e^{-\lambda}=\lambda\sum_{i=1}^{\infty}\frac{\lambda^{i-1}}{(i-1)!}e^{-\lambda}=\lambda.$$

对于随机变量的方差, 首先计算

$$\begin{aligned} E[X^2] &= \sum_{i=0}^{\infty}i^2P(X=i)=\sum_{i=1}^{\infty}i^2\frac{\lambda^i}{i!}e^{-\lambda}=\lambda\sum_{i=1}^{\infty}i\cdot\frac{\lambda^{i-1}}{(i-1)!}e^{-\lambda} \\ &= \lambda\sum_{i=1}^{\infty}(i-1)\frac{\lambda^{i-1}}{(i-1)!}e^{-\lambda}+\lambda=\lambda\sum_{k=0}^{\infty}k\frac{\lambda^k}{k!}e^{-\lambda}+\lambda=\lambda^2+\lambda. \end{aligned}$$

从而得到$\mathrm{Var}(X)=E[X^2]-(E[X])^2=\lambda.$ □

例3.11. 设随机变量X服从参数为λ的泊松分布, 且$P(X=1)=P(X=2)$, 求$P\{X\geq 4\}$.

解. 根据泊松分布的定义可知 $P(X=k)=\lambda^k e^{-\lambda}/k!$, 由此可得

$$P(X=1)=P(X=2) \quad \Rightarrow \quad \frac{\lambda e^{-\lambda}}{1!}=\frac{\lambda^2}{2!} \quad \Rightarrow \quad \lambda=2,$$

进一步得到

$$\begin{aligned} P\{X\geq 4\} &= 1-P(X=0)-P(X=1)-P(X=2)-P(X=3) \\ &= 1-e^{-2}-4e^{-2}-\frac{4}{3}e^{-2}=1-\frac{19}{3}e^{-2}. \end{aligned}$$

□

下面研究二项分布于泊松分布的关系: 即可以用泊松分布逼近二项分布.

定理3.5 (泊松定理). *对任意给定的常数$\lambda>0$, n为任意正整数, 设$np_n=\lambda$, 则对任意给定的非负整数k, 有*

$$\lim_{n\to\infty}\binom{n}{k}p_n^k(1-p_n)^{n-k}=\frac{\lambda^k}{k!}e^{-\lambda}.$$

Proof. 由 $p_n=\lambda/n$, 有

$$\begin{aligned} \binom{n}{k}p_n^k(1-p_n)^{n-k} &= \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k}{k!}(1-\frac{1}{n})\cdots(1-\frac{k-1}{n})\left(1-\frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k}{k!}(1-\frac{1}{n})\cdots(1-\frac{k-1}{n})\left(1-\frac{\lambda}{n}\right)^{\frac{n}{\lambda}\frac{n-k}{n}\lambda} \end{aligned}$$

当$n\to\infty$, 有$(1-\frac{\lambda}{n})^{\frac{n}{\lambda}}\to e$以及$\frac{n-k}{n}\to 1$, 从而完成证明. □

泊松分布的应用: 若 $X\sim B(n,p)$, 当 n 比较大, p 比较小时, 令 $\lambda=np$, 有

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

即利用泊松分布近似计算二项分布.

例3.12. *设每次射击命中目标的概率为0.002, 现射击1000次, 求命中目标在500次与600次之间的概率. (用泊松近似计算)*

解. 将1000次射击可看作1000重Bernoulli试验, 设随机变量 X 表示1000次设计中射中目标的次数, 则 $X\sim B(1000,0.002)$, 利用泊松分布近似, 则可以看作 $X\sim P(2)$, 于是有

$$P(500\leq X\leq 600)=\sum_{k=500}^{600}\binom{1000}{k}(0.002)^k 0.998^{1000-k}\approx\sum_{k=500}^{600}\frac{2^k}{k!}e^{-2}.$$

□

例3.13. *有80台同类型设备, 各台独立工作, 发生故障的概率是0.01, 一台设备发生故障时只能1人处理, 考虑方案: I) 由4人维护, 每人单独负责20台; II) 由三人共同维护80台. 哪种方案更为可取?*

解. 对方案 I): 令 X 为第1人负责20台中同一时刻发生故障的台数, 则$X \sim B(20, 0.01)$. 设A_i表示第i人负责的设备发生故障不能及时维修的事件, 则

$$P(A_1 \cup A_2 \cup A_3 \cup A_4) \geq P(A_1) = P\{X \geq 2\} = 1 - P(X=0) - P(X=1) \approx 1 - \sum_{k=0}^{n} \frac{(0.2)^k}{k!} e^{-0.2} \approx 0.0175.$$

对方案 II): 设随机变量Y为80台设备中同一时刻发生故障的台数, 则$Y \sim (B, 0.01)$, 则有设备发生故障不能及时维修的概率为

$$P(Y \geq 4) = 1 - \sum_{k=1}^{3} \binom{80}{k} 0.01^k 0.99^{80-k} \approx 1 - \sum_{k=0}^{3} \frac{(0.8)^k}{k!} e^{-0.8} \approx 0.0091.$$

由此可知: 方案 II)更优. □

4 连续型随机变量

4.1 概念与性质

4.1.1 分布函数

定义4.1 (分布函数)**.** 任意给定随机变量$X \in \mathbb{R}$, 函数$F(x) = P(X \le x)$称为X的分布函数.

对任意实数$x_1 < x_2$, 有$P(x_1 < X \le x_2) = P(X \le x_2) - P(X \le x_1) = F(x_2) - F(x_1)$. 分布函数$F(x)$具有如下性质:

i) 单调性: 若$x_1 < x_2$, 则$F(x_1) \le F(x_2)$;

ii) 规范性: $F(x) \in [0, 1]$, 且$F(-\infty) = \lim_{x \to -\infty} F(x) = 0$, $F(+\infty) = \lim_{x \to +\infty} F(x) = 1$;

iii) 右连续性: $F(x + 0) = \lim_{\Delta x \to 0^+} F(x + \Delta x) = F(x)$.

任一分布函数必满足以上三性质, 反之亦成立. 满足以上三性质的函数必定是某随机变量的分布函数. 利用分布函数$F(x)$表示随机事件的概率:

$$
\begin{aligned}
P(X > a) &= 1 - F(a) \\
P(X < a) &= F(a - 0) = \lim_{x \to a^-} F(x) \\
P(X = a) &= F(a) - F(a - 0) \\
P(X \ge a) &= 1 - F(a - 0) \\
P(a \le X \le b) &= F(b) - F(a - 0).
\end{aligned}
$$

例4.1. 随机变量X的分布列为$\begin{array}{c|ccc} X & -1 & 2 & 3 \\ \hline P & 1/4 & 1/2 & 1/4 \end{array}$. 求$X$的分布函数.

解. 当$x < -1$时, $F(x) = P(X \le x) = P(\emptyset) = 0$,
当$-1 \le x < 2$时, $F(x) = P(X \le x) = P(X = -1) = \frac{1}{4}$,
当$2 \le x < 3$时, $F(x) = P(X \le x) = P(X = -1) + P(X = 2) = \frac{3}{4}$,
当$x \ge 3$时, $F(x) = 1$. □

例4.2. 在$[0, 1]$区间随机抛一个点, X表示落点的坐标, 假设X落入$[0, 1]$区间内任一子区间的概率与区间长度成正比, 求X的分布函数.

解. 由$X \in [0, 1]$, 当$x < 0$时, $F(x) = 0$, 当$x > 1$时, $F(x) = 1$. 当$x \in [0, 1]$时, $F(x) = P(X \le x) = kx$, 由$F(1) = 1$, 所以$k = 1$. 综上

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

□

例4.3. 随机变量X的分布函数$F(X) = A + B\arctan x$, $x \in \mathbb{R}$, 求A, B.

解. 由分布函数的性质有

$$0 = F(-\infty) = \lim_{x\to-\infty} A + B\arctan x = A - \frac{\pi}{2}B,$$

$$1 = F(+\infty) = \lim_{x\to+\infty} A + B\arctan x = A + \frac{\pi}{2}B.$$

所以$A = 1/2, B = 1/\pi$. □

4.1.2 密度函数

定义4.2. 给定随机变量 X 的分布函数 $F(x)$, 如果存在可积函数 $f(x)$, 对任意 $x \in \mathbb{R}$, 有 $F(x) = \int_{-\infty}^{x} f(t)dt$, 称 X 为连续型随机变量, 函数 $f(x)$ 为随机变量 X 的概率密度函数, 简称概率密度.

性质4.1. 密度函数$f(x)$满足*i)* 非负性: 对任意$x \in \mathbb{R}$, 有$f(x) \geq 0$; *ii)* 规范性: $\int_{-\infty}^{+\infty} f(t)dt = 1$;

概率密度满足性质 i) 和 ii), 反之亦成立. 若$f(x)$满足性质 i) 和 ii), 引入$G(x) = \int_{-\infty}^{x} f(t)dt$是随机变量$X$的分布函数.

对任意$x_1 \leq x_2$, 有

$$P(x_1 < X \leq x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(t)dt$$

几何解释: X落入区间$[x_1, x_2]$的概率等于x轴, $x = x_1$, $x = x_2$, $y = f(x)$所围成的曲边梯形的面积.

定理4.1. 对于连续性随机变量X, 其分布函数$F(x)$在整个实数域上连续; 若$f(x)$在x点连续, 则$F(x)$在x点可导, 且$F'(x) = f(x)$.

Proof. 根据函数的积分性质: 若$f(x)$在$[a, b]$上可积, 则$\phi(x) = \int_a^x f(t)dt$在$[a, b]$上连续. 若$f(x)$在$[a, b]$上连续, 则$\phi(x) = \int_a^x f(t)dt$在$[a, b]$上可导. □

性质4.2. 对于任意 $x \in \mathbb{R}$, 以及任意连续型随机变量X, 有$P(X = x) = 0$;

Proof. 根据定义有

$$P(X = x) = \lim_{\Delta x\to 0} P(x - \Delta x \leq X \leq x + \Delta x) = \lim_{\Delta x\to 0} \int_{x-\Delta x}^{x+\Delta x} f(t)dt \leq \lim_{\Delta x\to 0} 2|f(x)|\Delta x \to 0.$$

□

由此可知: 连续型随机变量无需考虑端点, 即

$$P(a \leq X \leq b) = P(a < X < b) = P(a \leq X < b),$$

以及概率密度函数不是概率, 即$P(X = x) = 0 \neq f(x)$.

若$f(x)$在点x连续, 由连续性定义有

$$\lim_{\Delta x\to 0}\frac{P(x\le X\le x+\Delta x)}{\Delta x}=\lim_{\Delta x\to 0}\frac{\int_x^{x+\Delta x}f(t)dt}{\Delta x}=\lim_{\Delta x\to 0}\frac{\Delta x f(\xi)}{\Delta x}=f(x)\quad(\xi\in(x,x+\Delta x)),$$

由此可得$P(x\le X\le x+\Delta x)\approx f(x)\Delta x$, 若概率密度$f(x)$越大, 则$X$在$x$附近取值的概率越大.

例4.4. 设X是连续型随机变量, 密度函数$f(x)=\begin{cases}c(4x-2x^2) & 0<x<2\\ 0 & \text{其它}\end{cases}$, 求$P(X>1)$.

解.

$$F(\infty)=1=\int_{-\infty}^{+\infty}f(t)dt=\int_0^2 c(4t-2t^2)dt=\frac{8}{3}c,$$

得到$c=3/8$, 所以

$$P(X>1)=\int_1^{\infty}f(t)dt=\int_1^2 f(t)dt=\int_1^2\frac{8}{3}(4t-2t^2)dt=\frac{1}{2}.$$

□

例4.5 (由密度函数求分布函数)**.** 设X的密度函数$f(x)=\begin{cases}x & 0\le x\le 1\\ a-x & 1<x<2\\ 0 & \text{其它}\end{cases}$, 求$F(x)$.

解. 根据密度函数的规范性有

$$1=\int_{-\infty}^{+\infty}f(t)dt=\int_0^1 tdt+\int_1^2(a-t)dt=\frac{1}{2}+a-2+\frac{1}{2}=a-1\Rightarrow a=2.$$

进一步得到

$$f(x)=\begin{cases}x & 0\le x\le 1\\ 2-x & 1<x<2\\ 0 & \text{其它}\end{cases}$$

于是, 当$x\le 0$时, $F(x)=0$; 当$0<x\le 1$时, $F(x)=\int_0^x f(t)dt=x^2/2$; 当$1<x\le 2$时, $F(x)=\int_0^1 f(t)dt+\int_1^x f(t)dt=1/2+\int_1^x(2-t)dt=-x^2/2+2x-1$; 当$x\ge 2$时, $F(x)=1$. 最后得到

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

□

例4.6 (由分布函数计算密度函数)**.** 一个靶半径为$2m$的圆盘, 设击中靶上任一同心圆盘上的点的概率与该圆盘的面积成正比. 假设射击都能击中靶. X表示击中点与圆心的距离, 求X的密度函数.

解. 当$x < 0$时, $F(x) = 0$; 当$0 \le x \le 2$时, $F(x) = P(X \le x) = P(0 \le X \le x) = kx^2$. 由$F(2) = 1 = 4k$可得$k = 1/4$, 此时$F(x) = x^2/4$; 当$x > 2$时, $F(x) = 1$. 于是有X的概率密度

$$f(x) = \begin{cases} x/2 & 0 \le x \le 2 \\ 0 & \text{其它} \end{cases}$$

□

4.1.3 连续型随机变量的期望和方差

定义4.3. *设连续型随机变量X的密度函数为$f(x)$, 若积分$\int_{-\infty}^{+\infty} tf(t)dt$绝对收敛, 称$\int_{-\infty}^{+\infty} tf(t)dt$为$X$的期望, 记为$E(X)$, $E(X) = \int_{-\infty}^{+\infty} tf(t)dt$.*

与离散性随机变量一致, 有如下性质:

性质4.3. *i) 对任意任意常数a, b和随机变量X, 有$E(aX + b) = aE(X) + b$;*
ii) 若X的密度函数为$f(x)$, 且$\int_{-\infty}^{+\infty} g(t)f(t)dt$绝对可积, 则

$$E(g(X)) = \int_{-\infty}^{+\infty} g(t)f(t)dt;$$

iii) 对连续函数$g_1(x), \ldots, g_n(x)$和常数$c_1, \ldots, c_n$), 有

$$E\left(\sum_{i=1}^{n} c_i g_i(X)\right) = \sum_{i=1}^{n} c_i E(g_i(X)).$$

例4.7. *设随机变量X的密度函数为$f(x) = \begin{cases} cx & x \in [0, 1] \\ 0 & \text{其它} \end{cases}$, 求$E(X^m)$, 其中$m$为正整数.*

解. 由$\int_{-\infty}^{+\infty} f(t)dt = 1$, 可解$c = 2$, 所以

$$E(X^m) = \int_0^1 t^m \cdot 2tdt = 2\int_0^1 t^{m+1}dt = \frac{2}{m+2}.$$

由此可得$E(X) = 2/3$以及$E(X^2) = 1/2$. □

定理4.2. *若随机变量X非负, 有$E[X] = \int_0^{\infty} P(X > t)dt$.*

Proof. 容易得到

$$X = \int_0^X 1dt = \int_0^{+\infty} \mathbb{I}[t < X]dt,$$

这里$\mathbb{I}[\cdot]$表示指示函数, 如果论断为真, 其值为1, 否则为0. 于是得到

$$E[X] = \underbrace{E[\int_0^{+\infty} \mathbb{I}[t < X]dt] = \int_0^{+\infty} E[\mathbb{I}[t < X]]dt}_{\text{积分换序}} = \int_0^{+\infty} P(X > t)dt.$$

这里使用积分换序

$$E[\int_0^{+\infty}\mathbb{I}[t<X]dt]=\int_0^{\infty}f(y)(\int_0^{+\infty}\mathbb{I}[t<X]dt)dy$$
$$=\int_0^{\infty}(\int_0^{+\infty}f(y)\mathbb{I}[t<X]dy)dt=\int_0^{+\infty}E[\mathbb{I}[t<X]]dt.$$

□

习题. 利用此定理计算上例中的期望$E(X)=2/3$.

推论4.1. 对任意非负函数$g(x):\mathbb{R}\to\mathbb{R}$和随机变量$X$，有$E[g(X)]=\int_0^{+\infty}Pr[g(X)>t]dt$.

定义4.4. 对于连续型随机变量X，设其密度函数为$f(x)$，若$\int_{-\infty}^{+\infty}(t-E(X))^2f(t)dt$收敛，称$\int_{-\infty}^{+\infty}(t-E(X))^2f(t)dt$为$X$的方差，记为$Var(X)$，即

$$Var(X)=E(X-E(X))^2=\int_{-\infty}^{+\infty}(t-E(X))^2f(t)dt.$$

性质4.4. *i)* 对任意随机变量X，有$Var(X)=E[X-E(X)]^2=E[X^2]-(E(X))^2$; *ii)* 对任意参数$a,b$和随机变量$X$，有$Var(aX+b)=a^2Var(X)$.

定理4.3 (Bhatia-Davis不等式)**.** 对任意随机变量$X\in[a,b]$，有

$$Var(X)\le(b-E(X))\,(E(X)-a)\le(b-a)^2/4.$$

Proof. 对任意$X\in[a,b]$，有

$$(b-X)(X-a)\ge 0\Rightarrow x^2\le(a+b)x-ab$$

两边同时取期望有$E(X^2)\le(a+b)E(X)-ab$, 从而得到

$$Vax(X)=E(X^2)-(E(X))^2\le(a+b)E(X)-ab-E(X^2)=(b-E(X))(E(X)-a).$$

设函数$g(t)=(b-t)(t-a)$ $(t\in(a,b))$, 根据二次函数的性质求解最大值可得$f(t)\le(b-a)^2/4$. □

4.2 常用连续性随机变量