# 随机变量及其分布

Didnelpsun

## 目录

# 1 一维随机变量

## 1.1 随机变量概念

定义：随机变量就是其值会随机而定的变量。设随机试验 $E$ 的样本空间 $\Omega = \omega$，如果对每一个 $\omega$ 都有唯一的实数 $X(\omega)$ 与之对应，并且对任意实数 $x$，$\{\omega|X(\omega) \leqslant x, \omega \in \Omega\}$ 是随机事件，则称定义在 $\Omega$ 上的实值单值函数 $X(\omega)$ 为**随机变量**，记为随机变量 $X$。

## 1.2 分布函数

### 1.2.1 概念

定义：设 $X$ 为随机变量，$x$ 为任意实数，称函数 $F(x) = P\{X \leqslant x\}$（$x \in R$ 且取遍所有实数）为随机变量 $X$ 的分布函数，或称 $X$ 服从分布 $F(x)$，记为 $X \sim F(x)$。（随着 $x$ 从 $-\infty$ 到 $+\infty$，$X(\omega)$ 到 $\varnothing$ 到 $\Omega$）

### 1.2.2 性质

同样是分布函数的充要条件：

- $F(x)$ 是 $x$ 的单调不减函数，即对任意实数 $x_1 < x_2$，有 $F(x_1) \leqslant F(x_2)$。
- $F(x)$ 是 $x$ 的右连续函数，即对任意 $x_0 \in R$，有 $\lim\limits_{x \to x_0^+} F(x) = F(x_0 + 0) = F(x_0)$。（左空心右实心）
- $F(-\infty) = \lim\limits_{x \to -\infty} F(x) = 0$，$F(+\infty) = \lim\limits_{x \to +\infty} F(x) = 1$。

### 1.2.3 应用

- $P\{X \leqslant a\} = F(a)$。
- $P\{X < a\} = F(a - 0)$。是指分布函数下该点左极限的概率值。
- $P(X = a) = F(a) - F(a - 0)$。$\because P\{X \leqslant a\} = P\{X < a \cup X = a\} = P\{X < a\} + P\{X = a\}$，$\therefore P\{X = a\} = P\{X \leqslant a\} - P\{X < a\} = F(a) - F(a - 0)$。

# 2 一维离散随机变量

定义：若随机变量 $X$ 只可能取有限个或可列各值 $x_1, x_2, \cdots$，则称 $X$ 为离散型随机变量。

## 2.1 分布律

定义：$P\{X=x_i\}=p_i$，$i=1,2,\cdots$ 为 $X$ 的**分布列**、**分布律**或**概率分布**，记为 $X \sim p_i$。

概率分布常用表格或矩阵表示：

| $X$ | $x_1$ | $x_2$ | $\cdots$ |
|---|---|---|---|
| $P$ | $p_1$ | $p_2$ | $\cdots$ |

或 $X \sim \begin{pmatrix} x_1 & x_2 & \cdots \\ p_1 & p_2 & \cdots \end{pmatrix}$。

## 2.2 性质

数列 $\{p_i\}$ 是离散型随机变量的概率分布的充要条件是 $p_i \geqslant 0$（$i=1,2,\cdots$）且 $\sum\limits_{i=1}^{n} p_i = 1$。

设离散型随机变量 $X$ 的概率分布为 $P\{X=x_i\}=p_i$，则 $X$ 的分布函数 $F(x)=P\{X \leqslant x\}=\sum\limits_{x_i \leqslant x} P\{X=x_i\}$，即离散型随机变量的分布律函数是一个左实右空的阶梯形函数。

$P\{X=x_i\}=P\{X \leqslant x_i\}-P\{X<x_i\}=F(x_i)-F(x_i-0)$，即某点的概率值为该点分布律值减去该点左极限的分布律值。

对实数轴上的任一集合 $B$ 有 $P\{X \in B\}=\sum\limits_{x_i \in B} P\{X=x_i\}$，特别地 $P\{a<X \leqslant b\}=P\{X \leqslant b\}-P\{X \leqslant a\}=F(b)-F(a)$。

## 2.3 应用

**例题**：已知随机变量 $X$ 的概率分布为：

| $X$ | $1$ | $2$ | $3$ |
|---|---|---|---|
| $P\{X=k\}$ | $\theta^2$ | $2\theta(1-\theta)$ | $(1-\theta)^2$ |

且 $P\{X \geqslant 2\}=\dfrac{3}{4}$，求未知参数 $\theta$ 与 $X$ 的分布函数 $F(x)$。

解：$\because P\{X \geqslant 2\}=\dfrac{3}{4}$，$\therefore 2\theta(1-\theta)+(1-\theta)^2=\dfrac{3}{4}$，解得 $\theta=\dfrac{1}{2}$，$-\dfrac{1}{2}$ 舍。

## 2.4 分布

### 2.4.1 0-1 分布

定义：若 $X$ 的概率分布为 $X \sim \begin{pmatrix} 1 & 0 \\ p & 1-p \end{pmatrix}$，即 $P\{X=1\}=p$，$P\{X=0\}=1-p$，则称 $X$ 服从参数为 $p$ 的 0-1 分布，记为 $X \sim B(1,p)$（$0<p<1$）。

0-1 分布基于一次伯努利试验，$X$ 也称为伯努利计数变量。

### 2.4.2 二项分布

定义：如果 $X$ 的概率分布为 $P\{X=k\}=C_n^k p^k(1-p)^{n-k}$（$k=0,1,\cdots,n$，$0<p<1$），则称 $X$ 服从参数为 $(n,p)$ 的**二项分布**，记为 $X \sim B(n,p)$。

二项分布基于 $n$ 重伯努利试验。

二项分布的分布律计算，总共进行试验 $n$ 次，已知成功的概率为 $p$，若成功了 $k$ 次，则 $k$ 次成功概率为 $p^k$，则失败次数为 $n-k$，从而 $n-k$ 失败概率为 $(1-p)^{n-k}$，因为 $n$ 次试验都是相互独立的，所以将成功的概率与失败的概率乘在一起。又在 $n$ 次中成功 $k$ 次就可以了，进行排列，所以还乘上 $C_n^k$。

### 2.4.3 泊松分布

定义：如果 $X$ 的概率分布为 $P\{X=k\}=\dfrac{\lambda^k}{k!}e^{-\lambda}$（$k=0,1,\cdots,n$，$\lambda>0$），则称 $X$ 服从参数为 $\lambda$ 的**泊松分布**，记为 $X \sim P(\lambda)$。

泊松分布基于某场合某单位时间内源源不断的质点来流的个数 $X=k$，$\lambda$ 代表质点流动到来的强度。也可以代表稀有事件发生的概率。

### 2.4.4 几何分布

定义：如果 $X$ 的概率分布为 $P\{X=k\}=(1-p)^{k-1}p$（$k=0,1,\cdots,n$，$0<p<1$），则称 $X$ 服从参数为 $p$ 的**几何分布**，记为 $X \sim G(p)$。

几何分布与几何无关，代表的是 $n$ 重伯努利试验首次成功就停止试验，试验次数可以为无穷。设 $X$ 表示伯努利试验中事件 $A$ 首次放生所需要的试验次数，则 $X \sim G(p)$，其中 $p=P(A)$。

从而根据意义，几何分布要求前 $k-1$ 次都失败，从而概率为 $(1-p)^{k-1}$，最后一次成功，所以再乘上 $p$。

### 2.4.5 超几何分布

定义：如果 $X$ 的概率分布为 $P\{X=k\}=\dfrac{C_M^k C_{N-M}^{n-k}}{C_N^n}$（$\max\{0,n-N+M\}\leqslant k\leqslant\min\{MM,n\}$，$M,N,n$ 为正整数且 $M\leqslant N$，$n\leqslant N$，$k$ 为整数），则称 $X$ 服从参数为 $(n,N,M)$ 的**超几何分布**，记为 $X\sim H(n,N,M)$。

超几何分布考的可能性很小，事件数就是古典概型的一个特例。

如有 $N$ 件产品，其中 $M$ 件正品，从而 $N-M$ 件次品，任取 $n$ 个，则取出 $k$ 件正品的概率就是超几何分布。

# 3 一维连续随机变量

定义：若随机变量 $X$ 的分布函数可以表示为 $F(x)=\int_{-\infty}^{x} f(t)\,\mathrm{d}t$（$x\in R$ 且取遍所有实数），其中 $f(x)$ 是非负可积函数，则 $X$ 为**连续型随机变量**。

## 3.1 概率密度

定义：$f(x)$ 称为 $X$ 的**概率密度函数**，简称**概率密度**，记为 $X\sim f(x)$。

## 3.2 性质

改变 $f(x)$ 有限各点的值 $f(x)$ 仍是概率密度，$f(x)$ 为某一随机变量 $X$ 的概率密度的充分必要条件：$f(x)\geqslant 0$，且 $\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x=1$。

若 $X$ 为连续型随机变量，$X\sim f(x)$，则对任意实数 $c$ 有 $P\{X=c\}=0$。

对实数轴上的任一集合 $B$ 有 $P\{X\in B\}=\int_B f(x)\,\mathrm{d}x$，特别地 $P\{a<X<b\}=P\{a\leqslant X<b\}=P\{a<X\leqslant b\}=P\{a\leqslant X\leqslant b\}=\int_a^b f(x)\,\mathrm{d}x=F(b)-F(a)$。

## 3.3 应用

**例题**：已知随机变量 $X$ 的概率密度为 $\begin{cases} Ax, & 1<x<2 \\ B, & 2\leqslant x<3 \\ 0, & \text{其他} \end{cases}$，且 $P\{1<X<2\}=P\{2<X<3\}$，求常数 $AB$，分布函数 $F(x)$ 以及概率 $P\{2<X<4\}$。

解：由于归一性 $\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x=1$，$\therefore \int_1^2 Ax\,\mathrm{d}x+\int_2^2 B\,\mathrm{d}x=1$。

$\therefore \dfrac{3}{2}A+B=1$。又 $P\{1<X<2\}=P\{2<X<3\}$。

$\therefore \int_1^2 Ax\,\mathrm{d}x = \int_2^3 B\,\mathrm{d}x$，即 $\therefore \int_1^2 Ax\,\mathrm{d}x = \int_2^2 B\,\mathrm{d}x = \frac{1}{2}$，$A = \frac{1}{3}$，$B = \frac{1}{2}$。

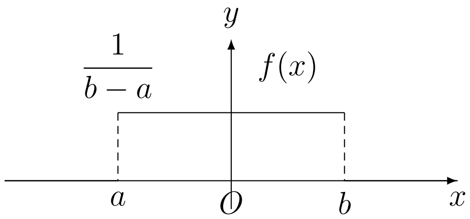
$f(x) = \begin{cases} \frac{1}{3}x, & 1 < x < 2 \\ \frac{1}{2}, & 2 \leqslant x < 3 \\ 0, & \text{其他} \end{cases}$，$\because F(x) = \int_{-\infty}^x f(t)\,\mathrm{d}t$。

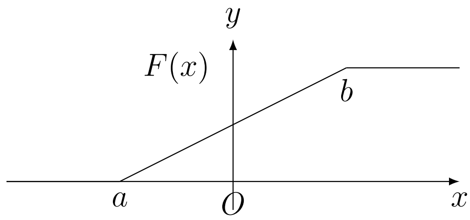
$$\therefore F(x) = \begin{cases} 0, & x < 1 \\ \int_1^x \frac{1}{3}t\,\mathrm{d}t = \frac{x^2}{6} - \frac{1}{6}, & 1 \leqslant x < 2 \\ \int_1^2 \frac{1}{3}x\,\mathrm{d}x + \int_2^x \frac{1}{2}\,\mathrm{d}x = \frac{1}{2}x - \frac{1}{2}, & 2 \leqslant x < 3 \\ 1, & x \geqslant 3 \end{cases}$$

## 3.4 分布

### 3.4.1 均匀分布

**定义：**如果 $X$ 的概率密度或分布函数分别为 $f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{其他} \end{cases}$，$F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leqslant x < b \\ 1, & x \geqslant b \end{cases}$，则称 $X$ 在区间 $(a, b)$ 上服从**均匀分布**，记为 $X \sim U(a, b)$。




### 3.4.2 指数分布

### 3.4.3 正态分布

# 4 一维随机变量函数分布

## 4.1 离散型

## 4.2 连续性