概率论与数理统计
第二章 随机变量及其概率分布

金玲飞

复旦大学计算机学院
Email: lfjin@fudan.edu.cn

2018.10.9

目录

第二章 随机变量及其概率分布

- 随机变量的概念
- 随机变量的分布函数及其性质
- 离散型随机变量
- 连续型随机变量
- 随机变量函数的分布。

随机变量的概念

- 在随机现象中很多样本点本身就用数量表示，由于样本点出现的随机性，其数量呈现为随机变量。

- 掷一颗骰子，出现的点数X是一个随机变量。
- 电视寿命T是一个随机变量。

- 对于样本点本身不是数的随机现象，也可以建立随机变量。

随机变量

定义 (随机变量 Random variable)

设$(\Omega, \mathscr{F}, P)$是一个概率空间，X是定义在样本空间Ω上的实值函数。若对任意实数x，$\{X \le x\}$是随机事件，即$\{\omega : X(\omega) \le x\} \in \mathscr{F}$，$X$称为随机变量。常用大写字母$X, Y, Z$ 表示，其取值大小用小写字母 x, y, z 等表示。

随机变量

定义 (随机变量 Random variable)

设$(\Omega, \mathscr{F}, P)$是一个概率空间，X是定义在样本空间Ω上的实值函数。若对任意实数x，$\{X \leq x\}$是随机事件，即$\{\omega : X(\omega) \leq x\} \in \mathscr{F}$，$X$称为随机变量。常用大写字母$X, Y, Z$ 表示，其取值大小用小写字母 x, y, z 等表示。

- 离散型随机变量：取值为有限个或可列无穷多个
- 连续型随机变量: 取值充满一个区间

随机变量

定义 (随机变量 Random variable)

设$(\Omega, \mathscr{F}, P)$是一个概率空间，X是定义在样本空间Ω上的实值函数。若对任意实数x，$\{X \le x\}$是随机事件，即$\{\omega : X(\omega) \le x\} \in \mathscr{F}$，$X$称为随机变量。常用大写字母$X, Y, Z$ 表示，其取值大小用小写字母 x, y, z 等表示。

- 离散型随机变量：取值为有限个或可列无穷多个
- 连续型随机变量: 取值充满一个区间

随机变量的引入使得我们可以利用微积分等数学工具来研究随机现象，全面深刻的揭示随机现象的统计规律。

- 记X为抛一颗骰子出现的点数。事件A="点数不大于3"，可以表示为$A = \{X \leq 3\}$.
- 记T为某个电器使用的寿命。事件C="使用寿命大于40000小于50000小时"，可以表示为$C = \{40000 < T < 50000\}$.

- 记X为抛一颗骰子出现的点数。事件A="点数不大于3"，可以表示为$A = \{X \leq 3\}$.
- 记T为某个电器使用的寿命。事件C="使用寿命大于40000小于50000小时"，可以表示为$C = \{40000 < T < 50000\}$.

用等号或不等号将随机变量X与某些实值连接起来，用来表示事件，如$\{X \leq a\}, \{X < a\}, \{X > b\}, \{a < X \leq b\}$.

- 记X为抛一颗骰子出现的点数。事件A="点数不大于3"，可以表示为$A = \{X \leq 3\}$.
- 记T为某个电器使用的寿命。事件C="使用寿命大于40000小于50000小时"，可以表示为$C = \{40000 < T < 50000\}$.

用等号或不等号将随机变量X与某些实值连接起来，用来表示事件，如$\{X \leq a\}, \{X < a\}, \{X > b\}, \{a < X \leq b\}$.

$$\{X < a\} = \cup_{k=1}^{\infty}\{X \leq a - \frac{1}{k}\}$$
$$\{a < X \leq b\} = \{X \leq b\} - \{X \leq a\}$$
$$\{X > b\} = \Omega - \{X \leq b\}$$

以上事件都可以用$\{X \leq a\}$ 描述。

概率的连续性

设$(\Omega, \mathscr{F}, P)$为概率空间，概率P具有如下性质：

(1) 下连续性：若$\{A_n : n \geq 1\}$是单调增的事件序列，则

$$\lim_{n \to \infty} P(A_n) = P(\cup_{n=1}^{\infty} A_n)$$

(2) 上连续性：若$\{A_n : n \geq 1\}$是单调降的事件序列，则

$$\lim_{n \to \infty} P(A_n) = P(\cap_{n=1}^{\infty} A_n)$$

分布函数

定义 (Distribution function)

设X是$(\Omega, \mathscr{F}, P)$上的随机变量。对任意实数x，

$$F(x) = P(X \leq x), -\infty < x < \infty$$

为X的分布函数，简称分布。并称X服从分布$F(x)$，记为$X \sim F(x)$。

分布函数

定义 (Distribution function)

设X是$(\Omega, \mathscr{F}, P)$上的随机变量。对任意实数x，

$$F(x) = P(X \leq x), -\infty < x < \infty$$

为X的分布函数，简称分布。并称X服从分布$F(x)$，记为$X \sim F(x)$。

- $0 \leq F(x) \leq 1$
- $P(a < X \leq b) = F(b) - F(a)$
- $P(X > a) = 1 - F(a)$

- $P(X < a) = F(a - 0)$
- $P(X = a) = F(a) - F(a - 0)$
- $P(a \leq X \leq b) = F(b) - F(a - 0)$

- $P(X < a) = F(a - 0)$
- $P(X = a) = F(a) - F(a - 0)$
- $P(a \leq X \leq b) = F(b) - F(a - 0)$

- 若$F(x)$是连续函数，则$P(X = a) = 0$。

- $P(X < a) = F(a-0)$
- $P(X = a) = F(a) - F(a-0)$
- $P(a \leq X \leq b) = F(b) - F(a-0)$
- 若$F(x)$是连续函数，则$P(X = a) = 0$。

随机变量落入任何区间的概率都可用分布函数表达。分布函数能完整的描述随机变量的统计规律性。掌握分布函数是关键。

分布函数的性质

定理 (2.1.1)

设随机变量X的分布函数为$F(x)$，则

- 单调性：$F(x)$是x的单调增函数，即$a \leq b$时有$F(a) \leq F(b)$

分布函数的性质

定理 (2.1.1)

设随机变量X的分布函数为$F(x)$，则

- 单调性：$F(x)$是x的单调增函数，即$a \le b$时有$F(a) \le F(b)$
- 有界性：对任意的x，都有$0 \le F(x) \le 1$，

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

- 右连续性：$F(x)$是X的右连续函数

分布函数的性质

定理 (2.1.1)

设随机变量X的分布函数为$F(x)$，则

- 单调性：$F(x)$是x的单调增函数，即$a \le b$时有$F(a) \le F(b)$
- 有界性：对任意的x，都有$0 \le F(x) \le 1$，

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

- 右连续性：$F(x)$是X的右连续函数

- 以上三条是$F(x)$称为某个随机变量的分布函数的充要条件。
- X与Y同分布，不代表"$X = Y$"

例子 (2.1.4)

若随机变量X有连续的分布函数

$$F(x) = \begin{cases} 0, & x < 0; \\ A\sin x, & 0 \le x < \pi/2; \\ 1, & x \ge \pi/2. \end{cases}$$

求常数A和概率$P(|X| < \pi/6)$.

例子 (2.1.1 退化分布，单点分布)

设$(\Omega, \mathscr{F}, P)$为概率空间，X是Ω上恒为常值的随机变量，即存在常数a，使得对任何$\omega \in \Omega$，有 $X(\omega) = a$，则X的分布函数为

$$F(x) = \begin{cases} 0, & x < a \\ 1, & x \geq a \end{cases}$$

例子 (2.1.2 伯努利分布，两点分布)

设X是$(\Omega, \mathscr{F}, P)$上的随机变量，X只取两个值$0, 1$。
设$p = P(X = 1)$，则$P(X = 0) = 1 - p$。此时X的分布函数为

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - p, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

古典概型的分布

例子 (2.1.3 离散均匀分布)

设随机变量X只取有限个值：$x_1, \ldots, x_n$。若对任何$1 \le k \le n$，$P(X = x_k) = \frac{1}{n}$，则称X服从$\{x_1, \ldots, x_n\}$上的 离散均匀分布，其分布函数为

$$F(x) = \frac{\text{集合}\{x_k : x_k \le x, k = 1, \ldots, n\}\text{中元素的个数}}{n}$$

$x \in (-\infty, \infty)$.

2.2 离散型随机变量

离散型随机变量

定义 (Discrete random variable)

设X是概率空间$(\Omega, \mathcal{F}, P)$上的随机变量，如果X只取有限个或可列无穷多个值$x_1, x_2, x_3, \ldots$，则称 X 为离散型随机变量，称

$$p_k = P(X = x_k), k \geq 1$$

为X的分布律 *probability mass function*（或分布列，或概率分布）。

离散型随机变量

定义 (Discrete random variable)

设X是概率空间$(\Omega, \mathcal{F}, P)$上的随机变量，如果X只取有限个或可列无穷多个值$x_1, x_2, x_3, \ldots$，则称 X 为离散型随机变量，称

$$p_k = P(X = x_k), k \geq 1$$

为X的分布律 *probability mass function*（或分布列，或概率分布）。

满足两个条件：

- $p_k \geq 0 \ (k \geq 1)$.
- $\sum_{k \geq 1} p_k = 1$.

—-判断一列实数是否为某一随机变量的分布律的依据。

例子 (2.2.1)

设$p_k = \frac{2^k}{k!}C, k = 1, 2, 3, \ldots$为某一随机变量的分布律。求常数$C$.

分布函数和分布律相互唯一决定：

分布函数和分布律相互唯一决定：

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

$$p_k = P(X \le x_k) - P(X < x_k) = F(x_k) - F(x_k - 0)$$

注： 离散型随机变量，分布律比分布函数更直观，简单。

分布函数和分布律相互唯一决定：

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

$$p_k = P(X \le x_k) - P(X < x_k) = F(x_k) - F(x_k - 0)$$

注：离散型随机变量，分布律比分布函数更直观，简单。

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

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_k & \cdots \\ p_1 & p_2 & \cdots & p_k & \cdots \end{pmatrix}$$

例子 (2.2.2)

设X的分布律为

X	0	1	2
P	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$

求X的分布函数。

二项分布

定义 (Binomial distribution)

设X是概率空间$(\Omega, \mathscr{F}, P)$上的随机变量，如果X的取值范围是$0, 1, \ldots, n$，且其分布律为

$$p_k = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, k = 0, 1, \ldots, n.$$

其中p为常数。则称X服从以n, p为参数的二项分布，记为$X \sim B(n, p)$。

二项分布

定义 (Binomial distribution)

设X是概率空间$(\Omega, \mathscr{F}, P)$上的随机变量，如果X的取值范围是$0, 1, \ldots, n$，且其分布律为

$$p_k = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, k = 0, 1, \ldots, n.$$

其中p为常数。则称X服从以n, p为参数的二项分布，记为$X \sim B(n, p)$。

- 验证二项分布的确是某一随机变量的分布律。

多重伯努利试验的随机变量的分布即是二项分布，记为

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

$$b(k; n, p) = b(n-k; n, 1-p)$$

多重伯努利试验的随机变量的分布即是二项分布，记为

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

$$b(k;n,p) = b(n-k;n,1-p)$$

问：k为何值时 $b(k;n,p)$ 达到最大？

例子 (2.2.3)

从积累的资料看，某厂生产的产品中，一等品为90%。现从该厂生产的*1000*件产品中随机抽取*20*件作检查。试求

- 恰有*18*件一等品的概率
- 一等品不超过*18*件的概率

例子 (2.2.3)

从积累的资料看，某厂生产的产品中，一等品为90%。现从该厂生产的*1000*件产品中随机抽取*20*件作检查。试求

- 恰有*18*件一等品的概率
- 一等品不超过*18*件的概率

解：用放回的抽样近似。
设A :该产品为一等品，$P(A) = 0.9$。
X :抽取的20件中一等品的个数，则$X \sim B(20.0.9)$.

$$P(X = 18) = ?$$