

04 CHAPTER

随机变量与随机分布



1

第4章 随机变量与随机分布

- 4.1 随机变量与概率分布的概念
- 4.2 联合概率分布
- 4.3 边缘概率分布
- 4.4 条件概率分布
- 4.5 随机变量的独立
- 4.6 随机变量的变换*

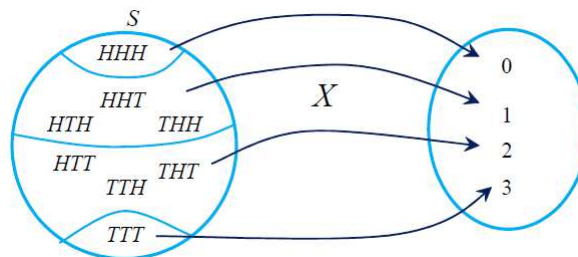
2/31

2

4.1 随机变量与随机分布的概念

[定义 4-1] 随机变量(random variable)

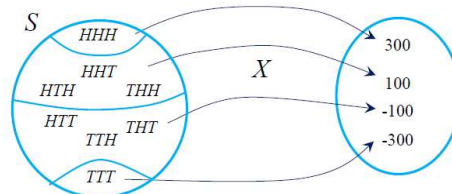
- 用实数来表示样本空间的各元素的一种函数
- 具有随机分布特点
- 扔三次硬币的试验中出现的样本空间
 $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$
- 硬币反面出现的次数 \rightarrow 随机变量



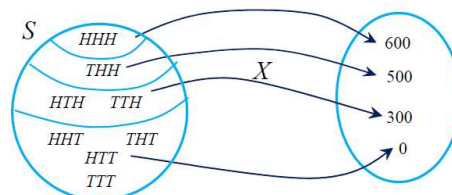
3/31

3

[例 4-1] 根据三次扔硬币出现的结果(正面的次数-反面的次数), 每次给出或得到100元, 在这游戏中产生的收益 \rightarrow 随机变量 X



- 第一次正面 \rightarrow 100, 第二次正面 \rightarrow 200, 第三次正面 \rightarrow 300
- 反面出现的话 \rightarrow 归0



4/31

4



- 离散型样本空间(discrete sample space)
 - 由有限个或者可数的无限个元素组成的样本空间
 - 离散型随机变量(discrete random variable)
 - ✓ 一直扔硬币，直到出现正面为止的次数
 - ✓ 100个产品中不合格品的数量
- 连续型样本空间(continuous sample space)
 - 可用实线内任意区间来表示的样本空间
 - 连续型随机变量(continuous random variable)
 - ✓ 人的身高和体重
 - ✓ 产品的寿命

5/31

5

4.1.2 概率分布



- [定义 4-2] 离散型概率分布(discrete probability distribution)
- 由离散型样本空间的随机变量形成的概率分布
 - 概率质量函数(probability mass function)

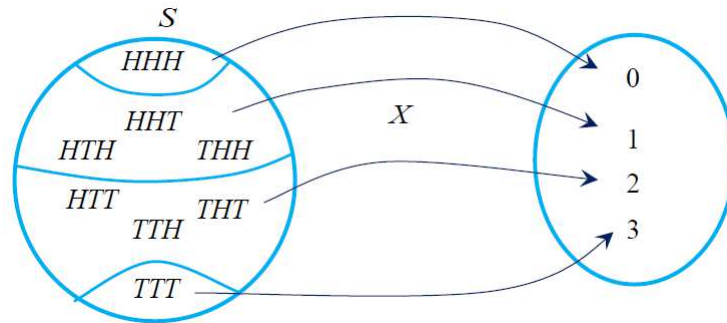
$$P(X = x) = f(x)$$

$$\sum_x f(x) = 1, 0 \leq f(x) \leq 1$$

6/31

6

[例 4-2] 在三次扔硬币的试验中，反面的个数X的概率分布



$$\Rightarrow f(0) = \frac{1}{8}, f(1) = \frac{3}{8}, f(2) = \frac{3}{8}, f(3) = \frac{1}{8}$$

$$\Rightarrow \sum_x f(x) = 1$$

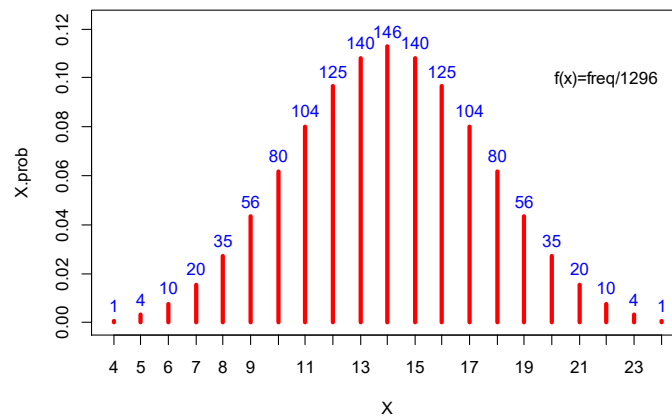
7/31

7

[例 4-3] 在四次掷骰子的试验中出现的点数和X的概率分布

随机变量 X 的分布图 [图 4-3]

주사위 4개 눈의 합 확률분포



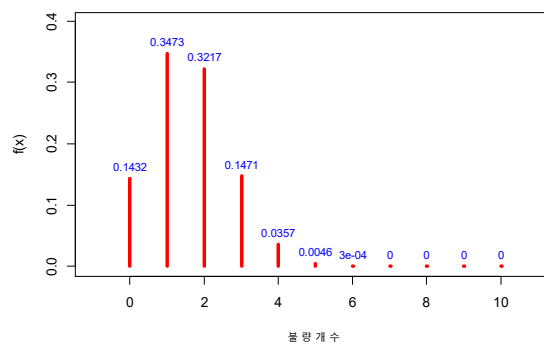
8/31

8

[例 4-4] 在装有50个产品 (其中有8个不合格品) 的箱子中随机抽取10个产品时, 发现不合格品的个数X的相关概率分布

$$f(x) = P(X = x) = \frac{{}_8C_x \times {}_{42}C_{10-x}}{{}_{50}C_{10}}, x = 0, 1, \dots, 8$$

(50개 중 8개 불량)에서 10개 추출 확률분포



9/31

9

[定义 4-3] 连续型概率分布(continuous probability distribution)

- 具有连续的(不可数的)值的随机变量的概率分布
- 概率分布函数f(x)是用来求概率P(a<X<b)的概率密度函数

$$P(a < X < b) = \int_a^b f(x)dx$$

$$f(x) \geq 0, \int_{-\infty}^{\infty} f(x)dx = 1$$

[例 4-5] 密度函数: $f(x) = 2e^{-2x}, 0 < x < \infty$

(1) 概率分布? $f(x) \geq 0 \quad \int_{-\infty}^{\infty} f(x)dx = \int_0^{\infty} 2e^{-2x} dx = [-e^{-2x}]_0^{\infty} = 1$

(2) $P(0 < X < 1) \quad P(0 < X < 1) = \int_0^1 2e^{-2x} dx = [-e^{-2x}]_0^1 = 1 - e^{-2} \approx 0.8647$

10/31

10

[定义 4-4] 累积分布函数(cumulative distribution function)

- 随机变量 X 低于特定值 x 的概率

$$F(x) = P(X \leq x)$$

[例 4-6] 在三次扔硬币的随机试验中, 出现反面的个数 X 的相关累积分布函数

$$f(0) = \frac{1}{8}, f(1) = \frac{3}{8}, f(2) = \frac{3}{8}, f(3) = \frac{1}{8}$$

$$\Rightarrow F(x) = \begin{cases} 0, & x < 0 \\ 1/8, & 0 \leq x < 1 \\ 1/2, & 1 \leq x < 2 \\ 7/8, & 2 \leq x < 3 \\ 1, & x \geq 3 \end{cases}$$

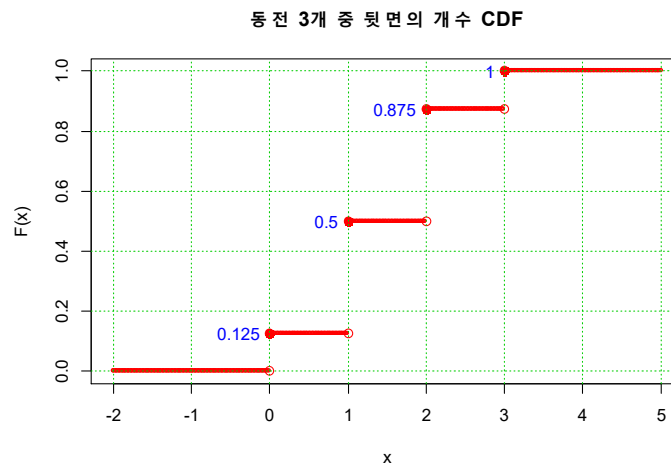
```
# 累积分布函数 F(x) 定义
Fx <- function(x) {
  if (x < 0) {y <- 0}
  } else if (x < 1) {y <- 1/8}
  } else if (x < 2) {y <- 1/2}
  } else if (x < 3) {y <- 7/8}
  } else {y <- 1}
  return(y) }
```

```
# 累积分布函数向量化  $\Rightarrow$  函数 Vectorize()
VFx <- Vectorize(Fx, "x")
```

11/31

11

累积分布函数 $F(x)$ 图



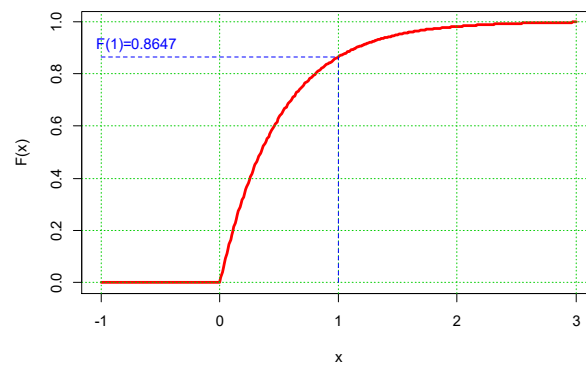
12/31

12

[例 4-7] 连续型概率分布的CDF $f(x) = 2e^{-2x}, 0 < x < \infty$

$$F(x) = \int_{-\infty}^x f(y)dy = \int_0^x 2e^{-2y} dy = [-e^{-2y}]_0^x = 1 - e^{-2x}, x > 0$$

연속형 누적확률분포함수 예



13/31