

第**4**章 离散型概率分布

- 1. 离散均衡分布
- 2. 二项分布
- 3. 超几何分布
- 4. 泊松分布
- 5. 几何分布
- 6. 负二项分布*
- 7. 多项分布*

2/31

2

자료구조및실습 1장. 자료구조와 알고리즘

1. 离散均衡分布



[定义 6-1] 离散均衡分布(discrete uniform distribution) n 个结果值以均衡的概率出项的概率分布

• pdf
$$f(x) = \frac{1}{n}, x = 1, 2, ..., n$$

• 期望
$$E(X) = \sum_{x=1}^{n} x \left(\frac{1}{n}\right) = \frac{1}{n} \times \frac{n(n+1)}{2} = \frac{n+1}{2}$$

•
$$\vec{j}$$
 $\not\equiv$
$$E(X^2) = \sum_{x=1}^n x^2 \left(\frac{1}{n}\right) = \frac{1}{n} \times \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}$$
$$\Rightarrow Var(X) = E(X^2) - E(X)^2$$
$$= \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{(n+1)(n-1)}{12}$$

3/31

3

1. 离散均衡分布



[例 6-1] 装着20个标有1到20号的球, 从中随机抽取一个,出现的号码X

- (1) X的概率分布函数 $f(x) = \frac{1}{20}, x = 1, 2, ..., 20$
- (2) X的期望与方差 $E(X) = \frac{20+1}{2} = 10.5$

$$Var(X) = \frac{(20+1)(20-1)}{12} = 33.25$$

(3) 大于15号以上的概率

$$P(X \ge 15) = \sum_{x=15}^{20} \frac{1}{20} = \frac{6}{20} = 0.3$$

4/31

2. 二项分布



[定义 6-2] 伯努利分布(Bernoulli distribution) X~B(1,p)

:在一次成功概率一定的试验中,出现的成功次数的概率分布

• pdf
$$f(x) = \begin{cases} p, & x = 1 \\ 1-p, & x = 0 \end{cases} \Leftrightarrow f(x) = p^x (1-p)^{1-x}, x = 0,1$$

• 期望
$$E(X) = \sum_{x=0}^{1} xf(x) = 0 \times (1-p) + 1 \times p = p$$

• 方差
$$E(X^2) = \sum_{x=0}^{1} x^2 f(x) = 0^2 \times (1-p) + 1^2 \times p = p$$

$$\Rightarrow Var(X) = E(X^2) - E(X)^2 = p - p^2 = p(1-p)$$

■ MGF
$$m(t) = E(e^{tX}) = \sum_{x=0}^{1} e^{tx} f(x) = e^{0} \times (1-p) + e^{t} \times p = 1-p+pe^{t}$$

$$\Rightarrow E(X) = m'(0) = pe^{0} = p \qquad E(X^{2}) = m''(0) = pe^{0} = p$$

$$\Rightarrow Var(X) = E(X^{2}) - E(X)^{2} = p - p^{2} = p(1-p)$$

5/21

5

2. 二项分布



[定义 6-3] 二项分布(binomial distribution) $X \sim B(n, p)$

:在n次成功概率一定的试验中,出现成功次数的概率分布

• Pdf
$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \ x = 0,1,2,...,n$$

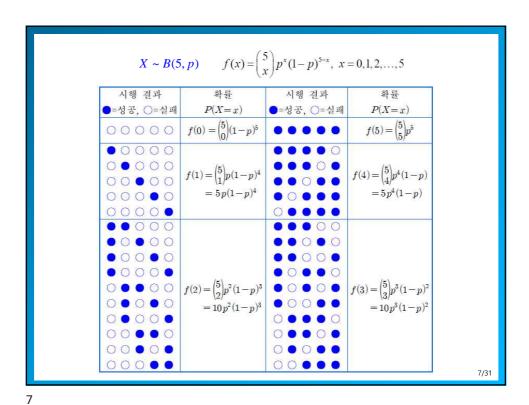
$$X_1, X_2, \dots, X_n \stackrel{iid}{\sim} B(1, p) \implies X \equiv \sum_{i=1}^n X_i \sim B(n, p)$$

• 期望
$$E(X) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} p = np$$

• 方差
$$Var(X) = \sum_{i=1}^{n} Var(X_i) = \sum_{i=1}^{n} p(1-p) = np(1-p)$$

• MGF
$$m(t) = m_{X_1}(t) \times m_{X_2}(t) \times \cdots \times m_{X_n}(t) = (1 - p + pe^t)^n$$

6/31

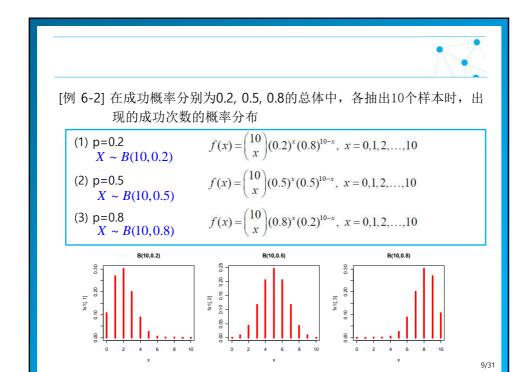


2. 二项分布

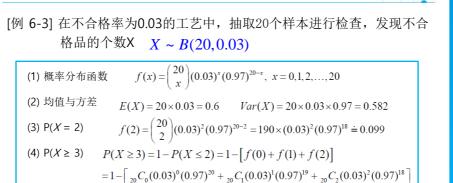


- # 二项分布 R函数
- # 概率分布函数 (size=n=样本大小, prob=p=成功概率) dbinom(x, size, prob)
- # Excel 函数 = BINOM.DIST(x, size, prob, FALSE)
- # 累积分布函数 (q=分位数, lower.tail=TRUE=由下向上累积) pbinom(q, size, prob, lower.tail = TRUE)
- # Excel 函数 = BINOM.DIST(x, size, prob, TRUE)
- # 分位数 (p=累积概率) qbinom(p, size, prob, lower.tail = TRUE)
- # Excel 函数 = BINOM.INV(size, prob, p)
- # 二项随机变量(n=随机数的个数) rbinom(n, size, prob)
- # 虽然没有相应的Excel函数, 但可用下面的指令生成一个随机数 = BINOM.INV(size, prob, RAND())

8/31



9



 $=1-\left\lceil {_{20}C_0(0.03)^0(0.97)^{20}} + {_{20}C_1(0.03)^1(0.97)^{19}} + {_{20}C_2(0.03)^2(0.97)^{18}} \right\rceil$ $=1-\left\lceil (0.97)^{20}+20\times 0.03\times (0.97)^{19}+190\times (0.03)^2(0.97)^{18}\right\rceil$

=1-(0.544+0.336+0.099)=1-0.979=0.021

dbinom(0:2, 20, 0.03)

2. 二项分布

[1] 0.54379434 0.33636763 0.09882967

1-sum(dbinom(0:2, 20, 0.03)); pbinom(2, 20, 0.03, lower=F)

[1] 0.02100836 [1] 0.02100836

4. 泊松分布



: 单位时间内发生的少数事件数的概率分布

$$f(x) = \lambda^{x} \frac{e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \dots$$

$$\lim_{n \to \infty} f(x) = \lim_{n \to \infty} \binom{n}{x} p^{x} (1-p)^{n-x} = \lim_{n \to \infty} \frac{n!}{x!(n-x)!} p^{x} (1-p)^{x}$$

$$\lambda = np \Rightarrow p = \frac{\lambda}{n} = \frac{1}{x!} \lim_{n \to \infty} \frac{n!}{(n-x)!} \left(\frac{\lambda}{n}\right)^{x} \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{\lambda^{x}}{x!} \lim_{n \to \infty} \frac{n(n-1)\cdots(n-x+1)}{n^{x}} \left(1 - \frac{\lambda}{n}\right)^{n} / \left(1 - \frac{\lambda}{n}\right)^{x}$$

$$= \lambda^{x} \frac{e^{-\lambda}}{x!}$$

11/31

11

4. 泊松分布



• 概率分布函数的条件

$$\sum_{x=0}^{\infty} f(x) = \sum_{x=0}^{\infty} \lambda^{x} \frac{e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^{x}}{x!} = e^{-\lambda} e^{\lambda} = 1$$

• MGF, 均值与方差

$$\begin{split} m(t) &= E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx} \lambda^{x} \frac{e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^{t})^{x}}{x!} = e^{-\lambda} e^{\lambda e^{t}} = e^{\lambda(e^{t}-1)} \\ m'(t) &= \lambda e^{t} e^{\lambda(e^{t}-1)} \implies E(X) = m'(0) = \lambda e^{0} e^{0} = \lambda \\ m''(t) &= \lambda e^{t} e^{\lambda(e^{t}-1)} + (\lambda e^{t})^{2} e^{\lambda(e^{t}-1)} \implies E(X^{2}) = m''(0) = \lambda + \lambda^{2} \\ \implies Var(X) &= E(X^{2}) - E(X)^{2} = \lambda \end{split}$$

12/31

4. 泊松分布



- # 泊松分布 R函数
- # 概率分布函数 (lambda=期望) dpois(x, lambda)
- # Excel函数 = POISSON.DIST(x, lambda, FALSE)
- # 累积分布函数 (q=分位数, lower.tail=TRUE=由下往上累积) ppois(q, lambda, lower.tail = TRUE)
- # Excel函数 = POISSON.DIST(x, lambda, TRUE)
- # 分位数 (p=累积概率) qpois(p, lambda, lower.tail = TRUE)
- # 泊松随机变量(n=随机数的个数) rpois(n, lambda)

13/31

13

4. 泊松分布



[例 6-6] 从单位时间内平均发生次数分别为2次、5次、8次的三个无限总体中抽取一定单位的样本时,泊松概率分布

(1) 单位时间内平均发生次数为2的情况 Poi(2)

$$f(x) = 2^x \frac{e^{-2}}{x!}, \ x = 0, 1, 2, \dots$$

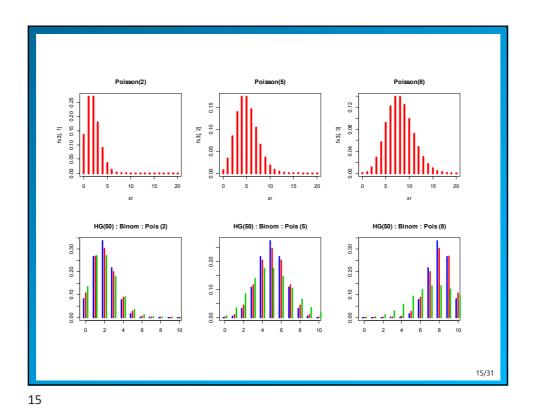
(2) 单位时间内平均发生次数为5的情况 Poi(5)

$$f(x) = 5^x \frac{e^{-5}}{x!}, x = 0,1,2,...$$

(3) 单位时间内平均发生次数为8的情况 Poi(8)

$$f(x) = 8^x \frac{e^{-8}}{x!}, x = 0,1,2,...$$

14/31



4. 泊松分布

[例 6-7] 对单位时间内平均瑕疵数为1.5的产品生产流程进行抽样 $X \sim Poi(1.5)$

- (1) 概率分布函数 $f(x) = (1.5)^x \frac{e^{-1.5}}{x!}, x = 0,1,2,...$
- $E(X) = \lambda = 1.5$ (2) 期望
- (3) 方差
- $Var(X) = \lambda = 1.5$ $f(2) = (1.5)^{2} \frac{e^{-1.5}}{2!} \doteq 0.251$ (4) P(2个)
- (5) P(3个以上) $P(X \ge 3) = 1 P(X \le 2) = 1 e^{-1.5} \times (1 + 1.5 + 1.5^2 / 2) = 0.191$ (6) P(10到20个) $Y \sim Poi(15) \implies P(Y = 20) = 15^{20} \times \frac{e^{-15}}{20!} = 0.042$

dpois(0:2, 1.5)

[1] 0.2231302 0.3346952 0.2510214

1-sum(dpois(0:2, 1.5)); ppois(2, 1.5, lower=F)

[1] 0.1911532 [1] 0.1911532



:在成功概率一定的试验中,直到出现第一次成功为止,所进行的试验次 数的概率分布

	f(x) = 0	$(1-p)^{x-1}p$	$x = 1, 2, \dots$
--	----------	----------------	-------------------

첫 번째 성공이 나오기까지 시행 시나리오	총 시행횟수	확률
(● : 성공 ○ : 실패)	(확률변수 <i>X</i>)	P(X=x)
•	X=1	p
0	X=2	(1-p)p
00	X=3	$(1-p)^2p$
000	X=4	$(1-p)^{3}p$
0000	X=5	$(1-p)^4p$
00000	X=6	$(1-p)^{5}p$
÷	:	÷
0000000000 0	X = x	$(1-p)^{x-1}p$

17

5. 几何分布



• 概率分布函数的条件

$$\sum_{x=1}^{\infty} f(x) = \sum_{x=1}^{\infty} (1-p)^{x-1} p = \frac{p}{1-(1-p)} = 1$$

累积分布函数

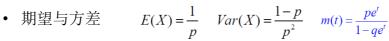
$$F(x) = P(X \le x) = \sum_{v=1}^{x} (1-p)^{v-1} p = \frac{p[1-(1-p)^{x}]}{1-(1-p)} = 1-(1-p)^{x}$$

• 动差生成函数

$$m(t) = E(e^{tX}) = \sum_{x=1}^{\infty} e^{tx} (1-p)^{x-1} p$$

$$x-1 \to y$$

$$= pe^{t} \sum_{y=0}^{\infty} (e^{t} q)^{y} = \frac{pe^{t}}{1-qe^{t}}, \ qe^{t} < 1$$



$$\begin{split} m'(t) &= \frac{pe'(1-qe') + pe'qe'}{(1-qe')^2} = \frac{pe'}{1-qe'} + \frac{pqe^{2t}}{(1-qe')^2} \\ \Rightarrow E(X) &= m'(0) = \frac{p}{p} + \frac{p(1-p)}{p^2} = \frac{1}{p} \\ m''(t) &= \frac{pe'(1-qe') + pe'qe'}{(1-qe')^2} + \frac{2pqe^{2t}(1-qe')^2 + 2pqe^{2t}q(1-qe')}{(1-qe')^4} \\ \Rightarrow E(X^2) &= m''(0) = \frac{p^2 + p(1-p)}{p^2} + \frac{2pqp^2 + 2pq^2p}{p^4} \\ &= \frac{p}{p^2} + \frac{2p(1-p) + 2(1-p)^2}{p^2} = \frac{2-p}{p^2} \\ \Rightarrow Var(X) &= E(X^2) - E(X)^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2} \end{split}$$

19/31

19

5. 几何分布



- # 几何分布 R函数
- # Excel函数不存在
- # 概率分布函数 (x=失败次数, prob=p=成功概率) dgeom(x, prob)
- # 累积分布函数 (q=分位数, lower.tail=TRUE=由下至上累积) pgeom(q, prob, lower.tail = TRUE)
- # 分位数 (p=累积概率) qgeom(p, prob, lower.tail = TRUE)
- # 几何随机变量(n=随机数的个数) rgeom(n, prob)

20/31



[例 6-8] 在成功概率分别为0.1, 0.2, 0.3, 0.5的四种无限总体中,第一次获得成功的试验

$$f(x) = p(1-p)^{x-1}$$

(1) 成功概率为0.1的情况 $f(x) = 0.1 \times 0.9^{x-1}, x = 1, 2, 3, ...$

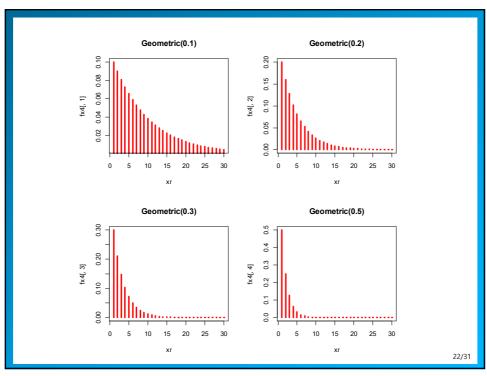
(2) 成功概率为0.2的情况 $f(x) = 0.2 \times 0.8^{x-1}, x = 1, 2, 3, ...$

(3) 成功概率为0.3的情况 $f(x) = 0.3 \times 0.7^{x-1}, x = 1, 2, 3, ...$

(4) 成功概率为0.5的情况 $f(x) = 0.5 \times 0.5^{x-1}, x = 1, 2, 3, ...$

21/31

21





[例 6-9] 反复掷一个骰子,直到出现点数'6'为止,其总试验次数X

- (1) X的概率分布函数 $f(x) = \frac{1}{6} \left(\frac{5}{6}\right)^{x-1}, x = 1, 2, 3, ...$
- (2) X的期望与方差 $E(X) = \frac{1}{p} = 6$ $Var(X) = \frac{q}{p^2} = \frac{5/6}{(1/6)^2} = 30$
- (3) 3次试验内出现 '6'的概率 $F(3) = 1 \left(\frac{5}{6}\right)^3 = \frac{91}{216} \doteq 0.421$
- 无记忆性 (memoryless property)

$$P(X > x + y \mid X > x) = P(X > y)$$

$$P(X > x + y \mid X > x) = \frac{P(X > x + y)}{P(X > x)} = \frac{q^{x+y}}{q^x} = q^y = P(X > y)$$

23/31