概　率　论

Probability

已学知识点

- 第一章 事件与概率
  - 随机现象与统计规律性

    ① 概率的频率解释依然是当今最通行的解释.

    ② 描述频率趋近于概率的大数定律总是概率论的第一大数定律.

    ③ 实际当中用频率作为概率的估计是十分自然的.
  - 样本空间与事件

符号	集合论含义	概率论含义
Ω	空间或全集	样本空间或必然事件
Φ	空集	不可能事件
ω	元素	样本点
A	子集	随机事件
$\omega \in A$	ω 是 A 的元素	事件 A 包含样本点 ω
$A \subset B$	A 是 B 的子集	A 发生则 B 发生
$AB = \Phi$	A, B 不相交	A, B 不可能同时发生
$A \cup B$	并集	A, B 至少有一个发生
$A \cap B$	交集	A, B 同时发生
$A - B$	差集	A 发生而 B 不发生
$\bar{A}$	余集	A 不发生

已学知识点

- 第一章 事件与概率
  - 古典概型 (等可能概率模型)：(1) 样本空间样本点有限；(2) 每个样本点等可能出现.
    - 计数方法：排列组合.
    - 三个基本性质：非负性、规范性、有限可加性.
  - 几何概率：以等可能性定义概率，处理无限场合，概率是几何体的测度之比.
    - 三个基本性质：非负性、规范性、可列可加性.
  - 概率空间：$\left(\Omega, \mathscr{F}, P\right)$
    - 难点和要点：事件域 $\mathscr{F}$ 的选择，太小不能满足需要，太大难以定义概率.
    - 选择包含我们关注的所有事件的 σ 域，保证事件对交、并、逆、差作可列次运算的封闭性.
    - 在这种 σ 域上，能定义满足非负、规范和可列可加性的概率测度.

已学知识点

- 第二章 条件概率与统计独立性

  - 条件概率的定义：$P(A\,|\,B)=\dfrac{P(AB)}{P(B)}$.

  - 条件概率也是概率：

① $P(A\,|\,B)\geq 0$

② $P(\Omega\,|\,B)=1$

③ $P\left(\bigcup\limits_{i=1}^{\infty}A_i\Big|\,B\right)=\sum\limits_{i=1}^{\infty}P\left(A_i\big|\,B\right)$

A_i 两两互不相容

④ $P\left(\overline{A}\big|\,B\right)=1-\left(A\big|\,B\right)$

⑤ $P\left(A\cup B\big|\,C\right)=P\left(A\big|\,C\right)+P\left(B\big|\,C\right)-P\left(AB\big|\,C\right)$

已学知识点

- 第二章　条件概率与统计独立性

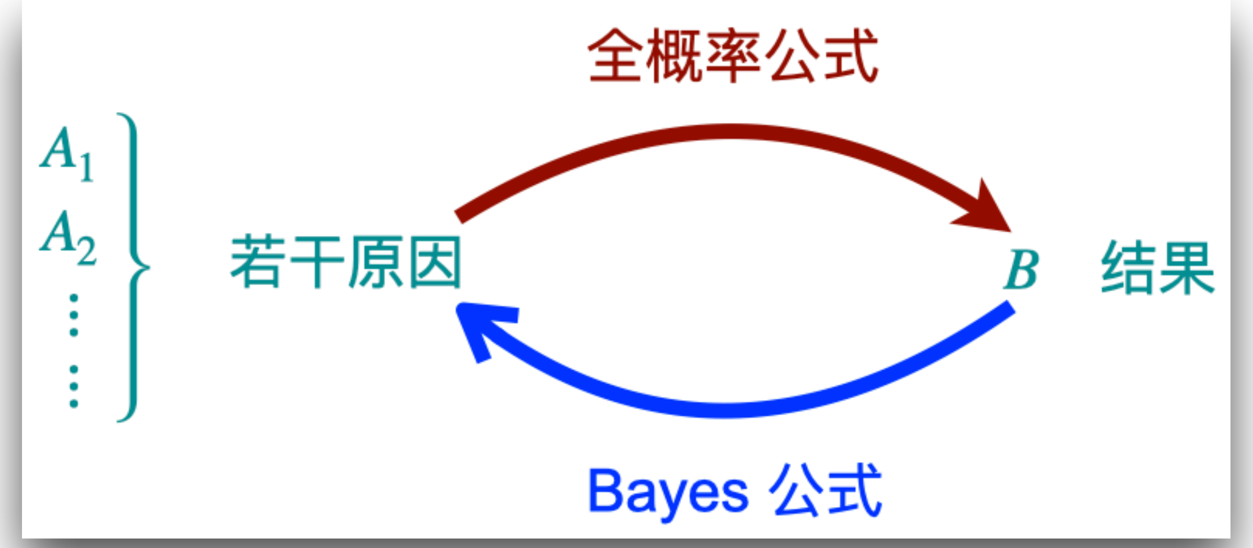
  - 乘法公式：$P(AB) = P(B)\cdot P(A\,|\,B) = P(A)\cdot P(B\,|\,A)$

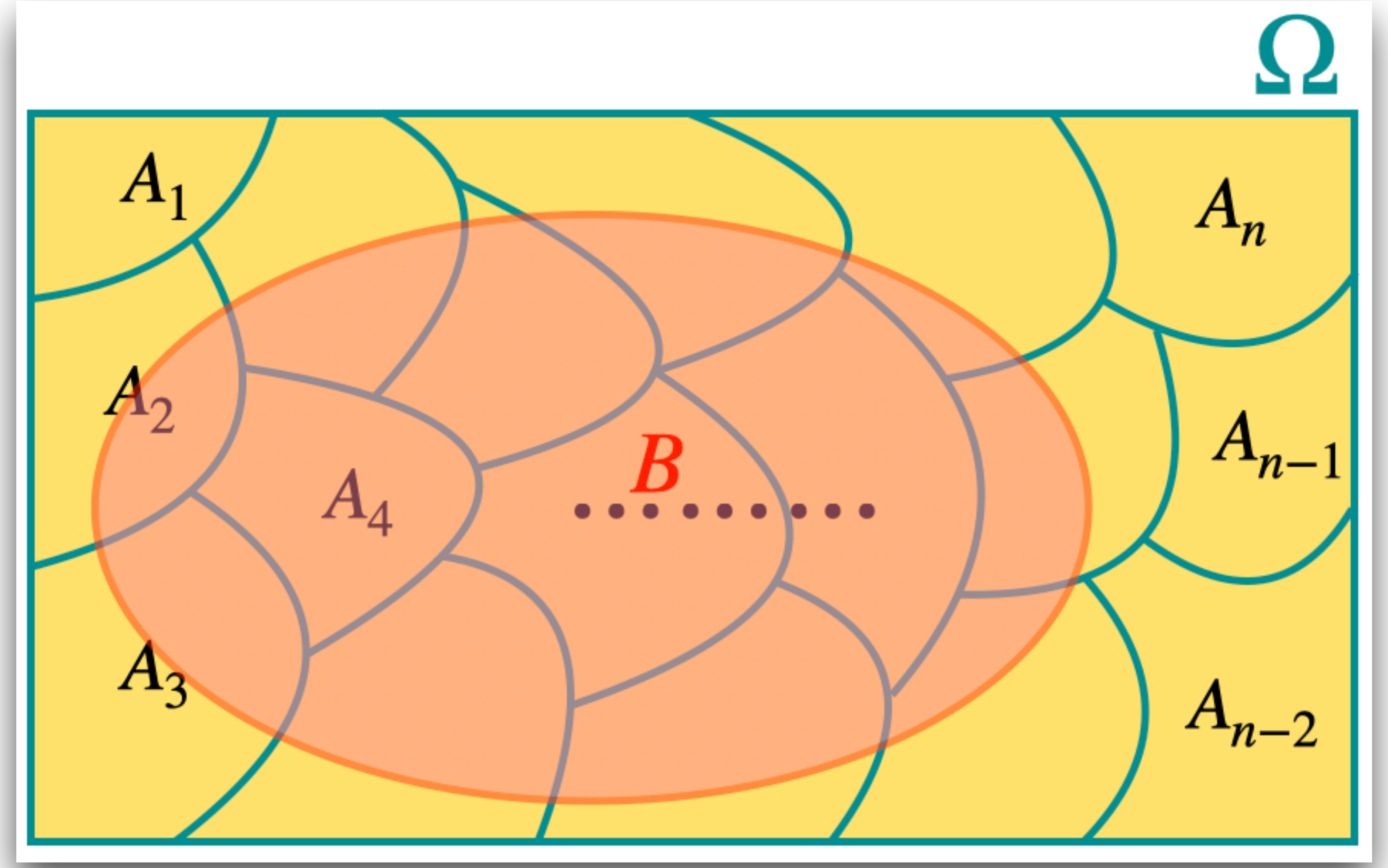
    $$P\left(A_1A_2\cdots A_n\right) = P\left(A_1\right)\cdot P\left(A_2\middle|\,A_1\right)\cdot P\left(A_3\middle|\,A_1A_2\right)\cdot\cdots\cdot P\left(A_n\middle|\,A_1A_2\cdots A_{n-1}\right)$$

    乘法公式主要用于计算若干个事件同时发生的概率.

  - 全概率公式：$P(B) = \sum_{i=1}^{\infty} P\left(A_i\right)\cdot P\left(B\,\middle|\,A_i\right)$

  - Bayes 公式：$P\left(A_k\,\middle|\,B\right) = \dfrac{P\left(A_k\right)\cdot P\left(B\,\middle|\,A_k\right)}{\sum_{i=1}^{\infty} P\left(A_i\right)\cdot P\left(B\,\middle|\,A_i\right)}$





已学知识点

- 第二章 条件概率与统计独立性
  - 两个事件独立：$P(AB) = P(A) \cdot P(B).$

$$\left.\begin{matrix} \overline{A}\ ,\ B \\ A\ ,\ \overline{B} \\ \overline{A}\ ,\ \overline{B} \end{matrix}\right\}\ \text{均相互独立}$$

  - 三个事件独立：$\begin{cases} P(AB) = P(A) \cdot P(B) \\ P(AC) = P(A) \cdot P(C) \\ P(BC) = P(B) \cdot P(C) \\ P(ABC) = P(A) \cdot P(B) \cdot P(C) \end{cases}$
  - $A_1, A_2, \ \ldots, A_n$ 相互独立，则

$$P\left(A_1 \cup A_2 \cup \cdots \cup A_n\right) = 1 - \left[1 - P\left(A_1\right)\right]\left[1 - P\left(A_2\right)\right] \cdot \cdots \cdot \left[1 - P\left(A_n\right)\right]$$

已学知识点

- 第二章 条件概率与统计独立性
  - 试验相互独立：一个试验的结果对其它各试验的可能结果的概率都无影响.
  - 试验独立性的定义：设 A_i 是试验 E_i 中的任一事件，$i=1,2,\ldots,n$，若
    $$P\left(A_1A_2\cdots A_n\right)=P\left(A_1\right)\cdot P\left(A_2\right)\cdot\cdots\cdot P\left(A_n\right)$$
    则称试验 $E_1, E_2, \ldots, E_n$ 是相互独立的.
  - 重复独立试验：研究在相同条件下重复进行的独立试验的数学模型.

已学知识点

- 第二章 条件概率与统计独立性

  - Bernoulli 试验 E：概率空间 $\left(\Omega, \mathscr{F}, P\right)$，其中

$$A \subset \Omega, \quad \mathscr{F} = \left\{\Phi, A, \bar{A}, \Omega\right\}, \quad P(A) = p, P\left(\bar{A}\right) = q, \quad (p > 0, q > 0, p + q = 1)$$

  - n 重 Bernoulli 试验 E_n：n 次独立重复的 Bernoulli 试验

    - Bernoulli 分布：$n = 1$ 的 Bernoulli 试验.

$$P_1(k) = p^k q^{1-k}, \quad k = 0, 1$$

样本点	H	T
概率	p	$1 - p$

    - 二项 (binomial) 分布：n 重 Bernoulli 试验中，成功 (事件 A) 出现的次数

$$b\left(k;\ n, p\right) \triangleq P_n(k) = C_n^k p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \ldots, n$$

已学知识点

- 第二章 条件概率与统计独立性
  - Bernoulli 试验 E：概率空间 $(\Omega, \mathscr{F}, P)$，其中

$$A \subset \Omega, \quad \mathscr{F} = \{\Phi, A, \bar{A}, \Omega\}, \quad P(A) = p, P(\bar{A}) = q, \quad (p > 0, q > 0, p + q = 1)$$

  - n 重 Bernoulli 试验 E_n：n 次独立重复的 Bernoulli 试验
    - 几何分布：n 重 Bernoulli 试验中，首次成功出现在第 k 次

$$g(k;\ p) = q^{k-1}p = (1-p)^{k-1}p\ , \quad k = 1, 2, \ \dots$$

    - Pascal 分布：n 重 Bernoulli 试验中，第 r 次成功出现在第 k 次

$$f(k;\ r, p) = C_{k-1}^{r-1}p^r q^{k-r} = \binom{k-1}{r-1} p^r (1-p)^{k-r}\ , \quad k = r, r+1, r+2, \ \dots$$

2.4 二项分布与 Poisson 分布

一、二项分布

二、二项分布的 Poisson 逼近

三、Poisson 分布

一、二项分布

- 二项分布：n 重 Bernoulli 试验中，事件 A 恰好发生 k 次的概率

$$b(k;\ n, p) = \binom{n}{k} p^k (1-p)^{n-k}\ , \quad k = 0,\ 1,\ 2,\ \ldots,\ n$$

- 二项分布的计算：使用 $\mathfrak{R}$ 计算.

```
?dbinom
```

Binomial {stats}

The Binomial Distribution

Description

Density, distribution function, quantile function and random generation for the binomial distribution with parameters `size` and `prob`.

This is conventionally interpreted as the number of 'successes' in `size` trials.

Usage

```
dbinom(x, size, prob, log = FALSE)
pbinom(q, size, prob, lower.tail = TRUE, log.p = FALSE)
qbinom(p, size, prob, lower.tail = TRUE, log.p = FALSE)
rbinom(n, size, prob)
```

Arguments

`x, q`	vector of quantiles.
`p`	vector of probabilities.
`n`	number of observations. If `length(n) > 1`, the length is taken to be the number required.
`size`	number of trials (zero or more).
`prob`	probability of success on each trial.
`log, log.p`	logical; if TRUE, probabilities p are given as log(p).
`lower.tail`	logical; if TRUE (default), probabilities are $P[X \le x]$, otherwise, $P[X > x]$.

一、二项分布

- 例：$n = 20,\ p_1 = 0.1,\ p_2 = 0.3,\ p_3 = 0.5$ 时的二项分布概率值.

```r
n = 20
p1 = 0.1
p2 = 0.3
p3 = 0.5
x1 = dbinom(0:20, n, p1)
x2 = dbinom(0:20, n, p2)
x3 = dbinom(0:20, n, p3)
P = data.frame(cbind(k = 0:20, p_1 = x1, p_2 = x2, p_3 = x3))
round(P, digits = 4)
```

```
> round(P, digits = 4)
       k    p_1    p_2    p_3
 [1,]  0 0.1216 0.0008 0.0000
 [2,]  1 0.2702 0.0068 0.0000
 [3,]  2 0.2852 0.0278 0.0002
 [4,]  3 0.1901 0.0716 0.0011
 [5,]  4 0.0898 0.1304 0.0046
 [6,]  5 0.0319 0.1789 0.0148
 [7,]  6 0.0089 0.1916 0.0370
 [8,]  7 0.0020 0.1643 0.0739
 [9,]  8 0.0004 0.1144 0.1201
[10,]  9 0.0001 0.0654 0.1602
[11,] 10 0.0000 0.0308 0.1762
[12,] 11 0.0000 0.0120 0.1602
[13,] 12 0.0000 0.0039 0.1201
[14,] 13 0.0000 0.0010 0.0739
[15,] 14 0.0000 0.0002 0.0370
[16,] 15 0.0000 0.0000 0.0148
[17,] 16 0.0000 0.0000 0.0046
[18,] 17 0.0000 0.0000 0.0011
[19,] 18 0.0000 0.0000 0.0002
[20,] 19 0.0000 0.0000 0.0000
[21,] 20 0.0000 0.0000 0.0000
```

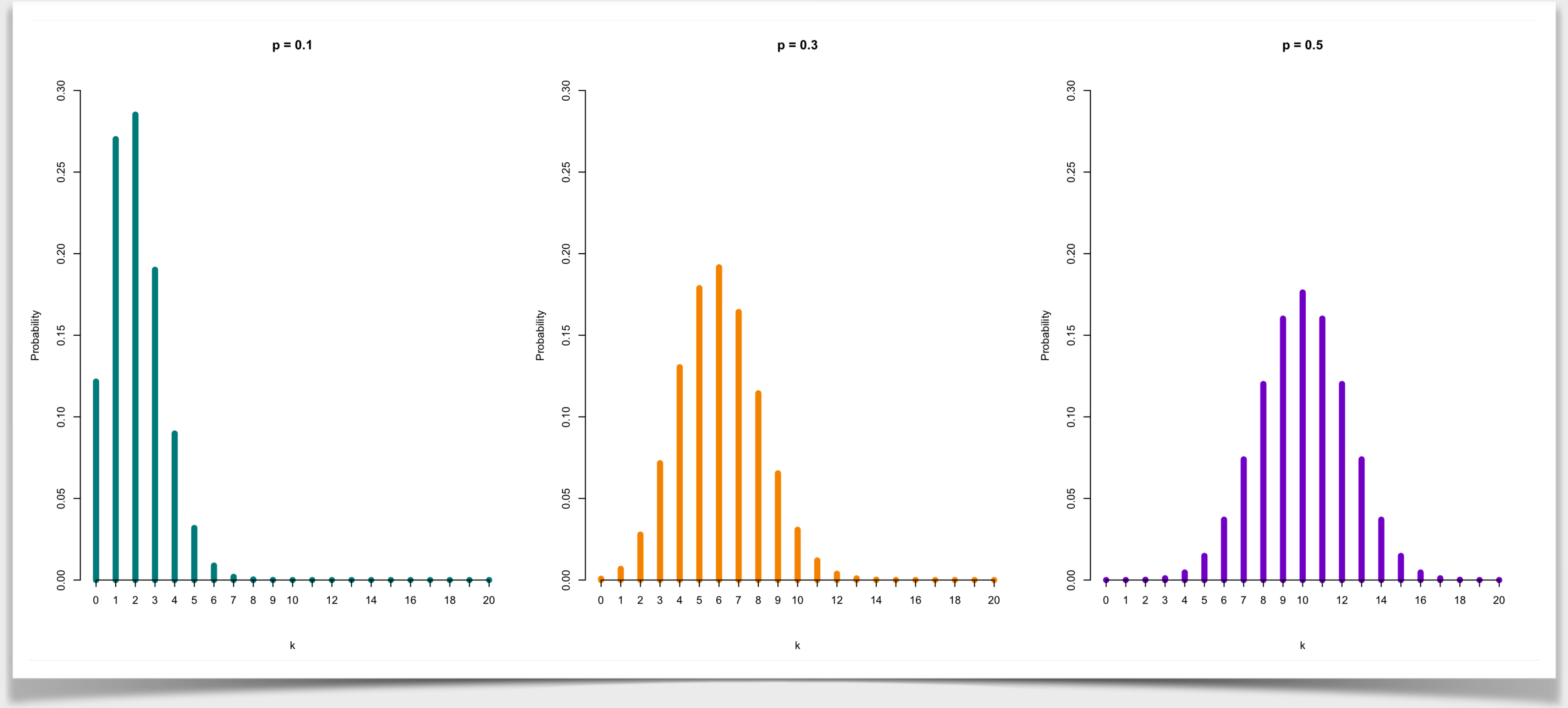
一、二项分布

- 例：$n = 20,\ p_1 = 0.1,\ p_2 = 0.3,\ p_3 = 0.5$ 时的二项分布概率值.

```
par(mfrow=c(1, 3))
plot(P$k, P$p_1, type = 'h', xlab = 'k', ylab = 'Probability', axes=F, ylim = c(0, 0.3),
    col = 'cyan4', lwd = 6, main = 'p = 0.1')
axis(1, at = 0:20, pos = 0)
axis(2)

plot(P$k, P$p_2, type = 'h', xlab = 'k', ylab = 'Probability', axes=F, ylim = c(0, 0.3),
    col = 'orange2', lwd = 6, main = 'p = 0.3')
axis(1, at = 0:20, pos = 0)
axis(2)

plot(P$k, P$p_3, type = 'h', xlab = 'k', ylab = 'Probability', axes=F, ylim = c(0, 0.3),
    col = 'purple3', lwd = 6, main = 'p = 0.5')
axis(1, at = 0:20, pos = 0)
axis(2)
```



一、二项分布

- 例：血清试验
  - 设在家畜中感染某种疾病的概率是 30%，新发现了一种血清可能对预防此病有效，为此对 20 只健康的动物注射了这种血清. 若注射后有一只动物受感染，我们应对此种血清的作用如何评价?
  - 如果此种血清毫无作用，则注射后动物受感染的概率还是 30%，那么 20 只这种动物中有 k 只受感染的概率为 $b(k;\ 20,\ 0.3)$.
  - 发生一只动物受感染或者更好的情况 (无动物受感染) 的概率为

$$b(0;\ 20,\ 0.3)+b(1;\ 20,\ 0.3)=0.007637$$

由于这个概率十分小，因此我们不能认为血清毫无价值.

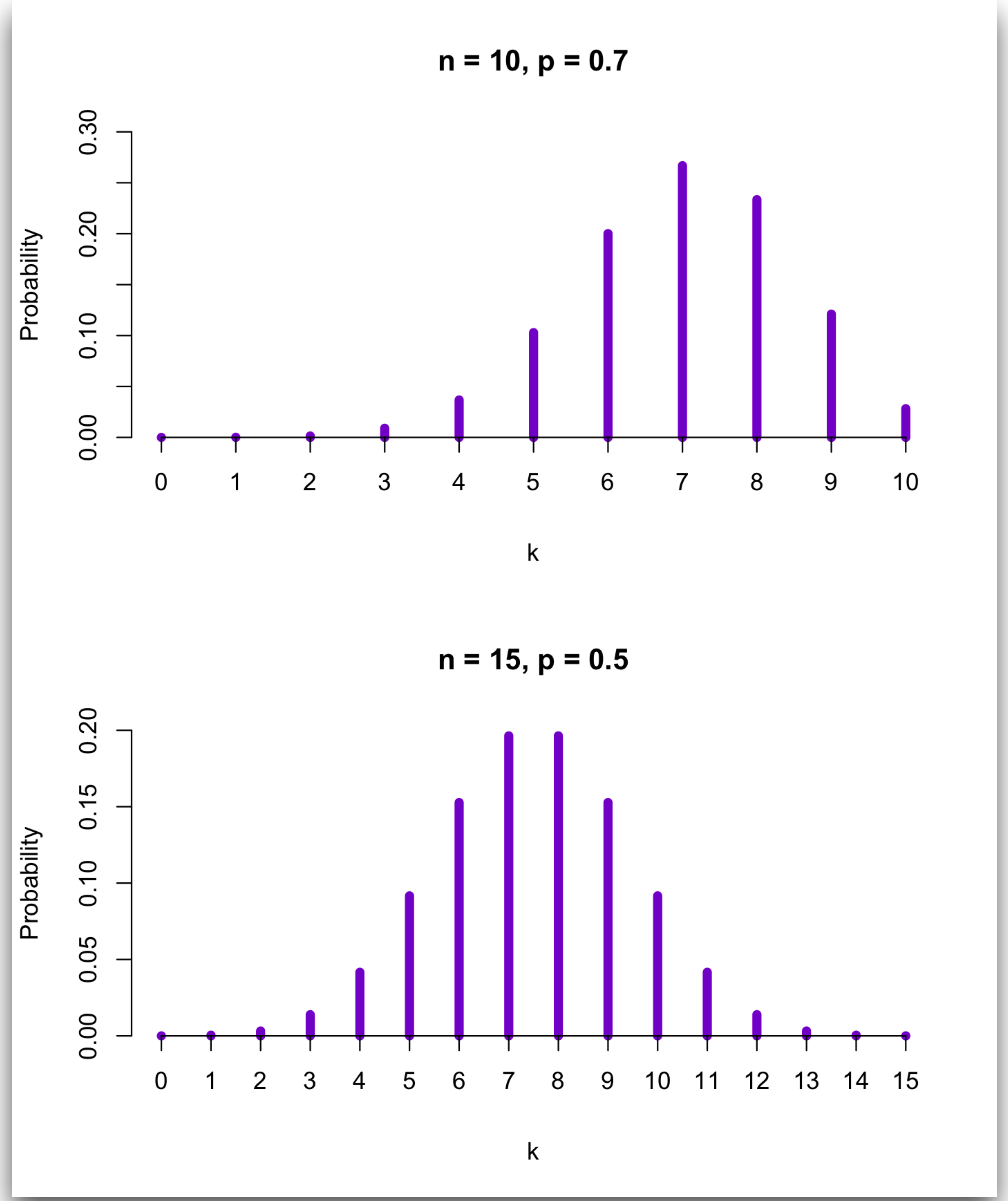
```
dbinom(0, 20, 0.3) + dbinom(1, 20, 0.3)
```

```
> dbinom(0, 20, 0.3) + dbinom(1, 20, 0.3)
[1] 0.007637
```

一、二项分布

二项分布的性质

- 固定 n 与 p，当 k 增加时，概率 $b(k;\ n, p)$ 先随之增加直至达到最大值，随后单调减少.
- 当 $(n+1)\cdot p$ 不为整数时，$b(k;\ n, p)$ 在 $\underbrace{\left[(n+1)\cdot p\right]}_{\text{取整}}$ 达到最大值.
- 当 $(n+1)\cdot p = m$ 为整数时，$b(k;\ n, p)$ 在 $k = m-1$ 和 $k = m$ 达到最大值.



一、二项分布

- 二项分布的性质

  - 固定 n 与 p，当 k 增加时，概率 $b(k;\ n, p)$ 先随之增加直至达到最大值，随后单调减少.
  - 当 $(n+1)\cdot p$ 不为整数时，$b(k;\ n, p)$ 在 $\left[(n+1)\cdot p\right]$ 达到最大值. 取整
  - 当 $(n+1)\cdot p = m$ 为整数时，$b(k;\ n, p)$ 在 $k = m-1$ 和 $k = m$ 达到最大值.

$p+q=1$

$$\frac{b(k;\ n, p)}{b(k-1;\ n, p)} = \frac{\binom{n}{k} p^k q^{n-k}}{\binom{n}{k-1} p^{k-1} q^{n-k+1}} = \frac{(n-k+1)p}{kq} = \frac{(n+1)p-kp}{kq}$$

$$= 1 + \frac{(n+1)p-k}{kq}, \quad (0<p<1)$$

$$\begin{cases} k < (n+1)p \Longrightarrow b(k;\ n, p) > b(k-1;\ n, p) \\ k = (n+1)p \Longrightarrow b(k;\ n, p) = b(k-1;\ n, p) \\ k > (n+1)p \Longrightarrow b(k;\ n, p) < b(k-1;\ n, p) \end{cases}$$

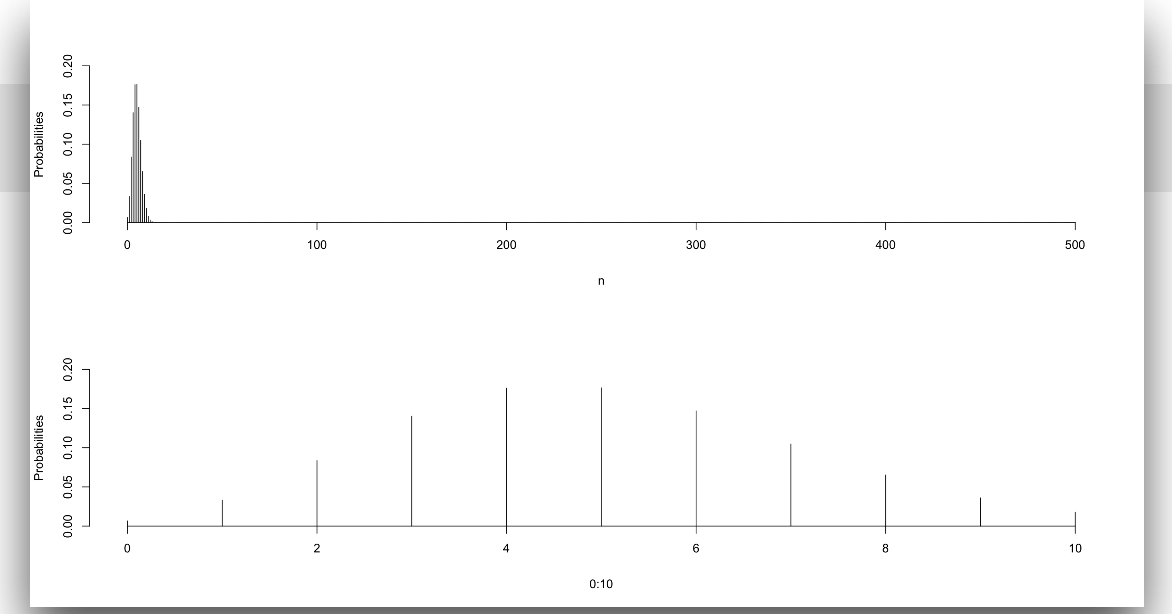
一、二项分布

- **例**：设某种疾病的发病率为 0.01，问在 500 人的社区中进行普查，最可能的发病人数是多少？并求其相应的概率.
  - 检测每一个人是 Bernoulli 概型，发病人数服从二项分布：$n = 500,\ p = 0.01.$
  - 因为 $(n+1)p = (500+1) \times 0.01 = 5.01$ 非整数，该二项分布的概率在 $[\ 5.01\ ] = 5$ 处取最大值.
  - 最可能的发病人数为 5 人. 相应的概率为：

$$b(5;\ 500,\ 0.01) = \binom{500}{5} \times 0.01^{5} \times 0.99^{495} = 0.17635105$$

dbinom(5, 500, 0.01)

```
> dbinom(5, 500, 0.01)
[1] 0.17635105
```



一、二项分布

- 例：(人寿保险) 某年龄段参加保险者当中，假设一年中每个人死亡的概率为 0.005，有 10000 个这类人参加人寿保险，求未来一年中参加保险者里面

  ① 有 40 个人死亡的概率.

  ② 死亡人数不超过 70 人的概率.

▶ Bernoulli 概型，$n = 10000$, $p = 0.005$ 的二项分布.

① 有 40 个人死亡的概率.

$$b(40;\ 10000,\ 0.005) = \binom{10000}{40} \times 0.005^{40} \times 0.995^{9960} = 0.021434812$$

```
dbinom(40, 10000, 0.005)
```

```
> dbinom(40, 10000, 0.005)
[1] 0.021434812
```

一、二项分布

- 例：(人寿保险) 某年龄段参加保险者当中，假设一年中每个人死亡的概率为 0.005，有 10000 个这类人参加人寿保险，求未来一年中参加保险者里面

  ① 有 40 个人死亡的概率.

  ② 死亡人数不超过 70 人的概率.

▸ Bernoulli 概型，$n = 10000,\ p = 0.005$ 的二项分布.

② 死亡人数不超过 70 人的概率.

$$\sum_{k=0}^{70} b(k;\ 10000,\ 0.005) = \sum_{k=0}^{70} \binom{10000}{k} \times 0.005^k \times 0.995^{10000-k} = 0.99709705$$

直接计算相当困难，
需要更有效的方法.

```
sum(dbinom(0:70, 10000, 0.005))
```

```
> sum(dbinom(0:70, 10000, 0.005))
[1] 0.99709705
```