数学建模

浙江大学数学科学学院　谈之奕

tanzy@zju.edu.cn

期望名次

- 第三个**Secretary Problem**
  - 招聘方录用一位应聘者，采用何种策略可使录用者绝对名次期望值尽可能小
    - 招聘方依据 $y_1, y_2, \cdots, y_i$ 值决定是否录用应聘者 A_i
    - 招聘方可以录用 $y_i > 1$ 的非备选者
- 在录用相对名次为 j 的应聘者 A_i 情况下，被录用者绝对名次的期望值为

$$\mathrm{E}(A_i \mid y_i = j) = \sum_{k=1}^{n} kP(A_i = k \mid y_i = j) = \sum_{k=1}^{n} \frac{kP(A_i = k, y_i = j)}{P(y_i = j)} = i\sum_{k=1}^{n} kP(A_i = k, y_i = j)$$

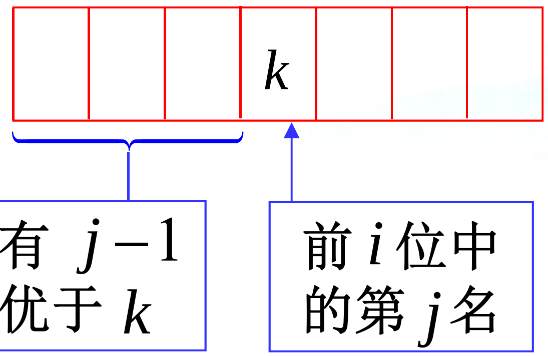
- $$P(y_i = 1) = P(y_i = 2) = \cdots = P(y_i = i) = \frac{1}{i}$$

条件概率

数学建模

$$P(A_i=k,y_i=j)=\frac{\binom{k-1}{j-1}\binom{n-k}{i-j}(i-1)!(n-i)!}{n!}=\frac{\binom{k-1}{j-1}\binom{n-k}{i-j}}{i\cdot\binom{n}{i}}$$

$$A_i=k,y_i=j$$



$$(k\ge j,n-k\ge i-j)$$

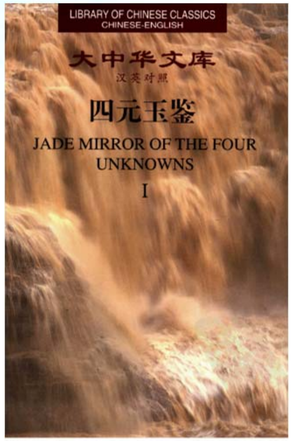
$$\mathrm{E}(A_i\mid y_i=j)=i\sum_{k=1}^{n}kP(A_i=k,y_i=j)$$

$$=\sum_{k=1}^{n}\frac{k\binom{k-1}{j-1}\binom{n-k}{i-j}}{\binom{n}{i}}=\sum_{k=1}^{n}\frac{j\binom{k}{j}\binom{n-k}{i-j}}{\binom{n}{i}}=\frac{j}{\binom{n}{i}}\sum_{k=1}^{n}\binom{k}{j}\binom{n-k}{i-j}$$

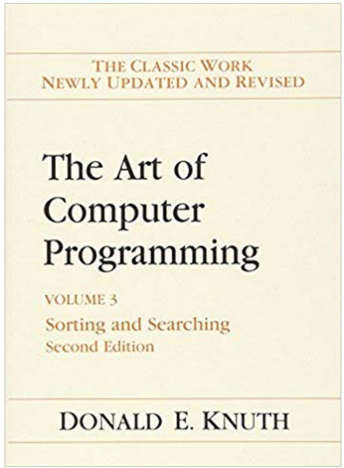
浙江大学 ZheJiang University

数学建模

Chu–Vandermonde Identity



《四元玉鉴》成书于1303年。全书分为28门，288题，是中国宋元数学高潮中水平最高的著作。



I. Sums of products. To complete our set of binomial-coefficient manipulations, we present the following very general identities, which are proved in the exercises at the end of this section. These formulas show how to sum over a product of two binomial coefficients, considering various places where the running variable k might appear:

$$\sum_k \binom{r}{k}\binom{s}{n-k} = \binom{r+s}{n}, \quad \text{integer } n. \tag{21}$$

$$\sum_k \binom{r}{m+k}\binom{s}{n+k} = \binom{r+s}{r-m+n}, \quad \text{integer } m, \text{ integer } n, \text{ integer } r \ge 0. \tag{22}$$

$$\sum_k \binom{r}{k}\binom{s+k}{n}(-1)^{r-k} = \binom{s}{n-r}, \quad \text{integer } n, \text{ integer } r \ge 0. \tag{23}$$

$$\sum_{k=0}^{r} \binom{r-k}{m}\binom{s}{k-t}(-1)^{k-t} = \binom{r-t-s}{r-t-m}, \quad \text{integer } t \ge 0, \text{ integer } r \ge 0, \text{ integer } m \ge 0. \tag{24}$$

$$\sum_{k=0}^{r} \binom{r-k}{m}\binom{s+k}{n} = \binom{r+s+1}{m+n+1}, \quad \text{integer } n \ge \text{integer } s \ge 0, \text{ integer } m \ge 0, \text{ integer } r \ge 0. \tag{25}$$

$$\sum_{k\ge 0} \binom{r-tk}{k}\binom{s-t(n-k)}{n-k}\frac{r}{r-tk} = \binom{r+s-tn}{n}, \quad \text{integer } n. \tag{26}$$

Of these identities, Eq. (21) is by far the most important, and it should be memorized. One way to remember it is to interpret the right-hand side as the number of ways to select n people from among r men and s women; each term on the left is the number of ways to choose k of the men and $n-k$ of the women. Equation (21) is commonly called Vandermonde's convolution, since A. Vandermonde published it in *Mém. Acad. Roy. Sciences Paris* (1772), 489–498. However, it had appeared already in Shih-Chieh Chu's 1303 treatise mentioned earlier [see J. Needham, *Science and Civilization in China* **3** (Cambridge University Press, 1959), 138–139].

朱世杰(约1249-1314)，元代数学家，燕（今北京或其附近）人。主要著作有《算学启蒙》和《四元玉鉴》。朱世杰在“天元术”、“二元术”、“三元术”基础上，创造了“四元术”，即四元高次方程组解法。

Knuth DE. *The Art of Computer Programming*. Vol. 1. Addison-Wesley, 1997.

三角垛

- 三角垛
  - 茭草垛
  - 三角垛(落一形垛)
  - 撒星形垛(三角落一形垛)
  - 三角撒星形垛(撒星更落一形垛)
  - 三角撒星更落一形垛

$$\sum_{i=1}^{n} i=\frac{1}{2!}n(n+1)$$

$$\sum_{i=1}^{n}\frac{1}{2!}i(i+1)=\frac{1}{3!}n(n+1)(n+2)$$

$$\sum_{i=1}^{n}\frac{1}{3!}i(i+1)(i+2)=\frac{1}{4!}n(n+1)(n+2)(n+3)$$

$$\sum_{i=1}^{n}\frac{1}{4!}i(i+1)(i+2)(i+3)=\frac{1}{5!}n(n+1)(n+2)(n+3)(n+4)$$

$$\sum_{i=1}^{n}\frac{1}{5!}i(i+1)(i+2)(i+3)(i+4)=\frac{1}{6!}n(n+1)(n+2)(n+3)(n+4)(n+5)$$

$$\frac{1}{5!}n(n+1)(n+2)(n+3)(n+4)=8568 \qquad n^5+10n^4+35n^3+50n^2+24n=1028160$$

“茭草形段” 之四:

今有茭草八千五百六十八束，欲令撒星更落一形垛之，問底子幾何?

答曰：一十四束。

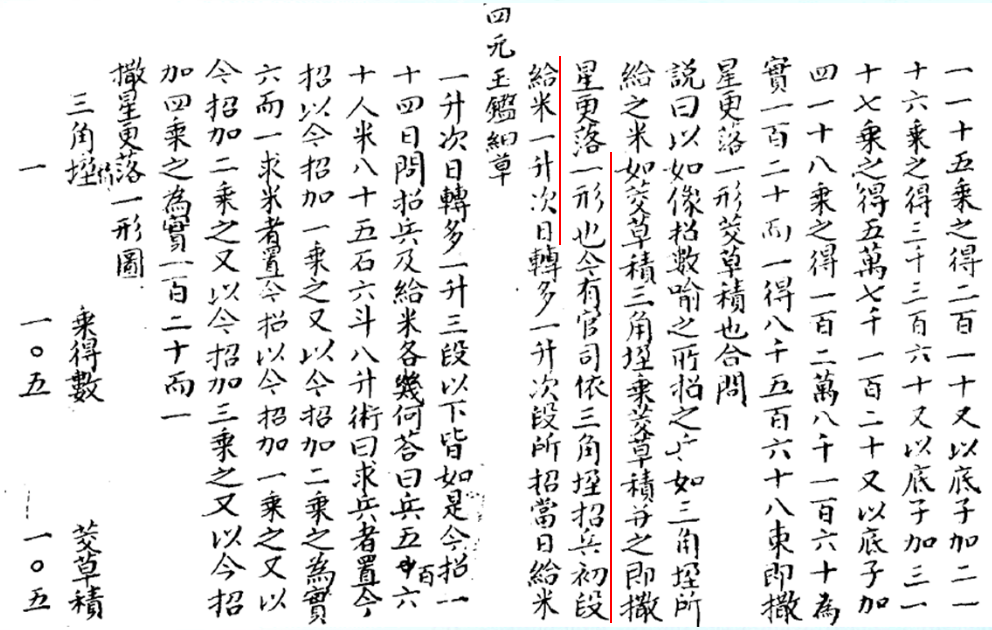
術曰：立天元一為撒星更落一底子，如積求之，得一百二萬八千一百六十為益實，二十四為從方，五十為從上廉，三十五為從二廉，一十為從三廉，一為正隅。四乘方開之，合問。

朱世杰—范德蒙恒等式

数学建模

钱宝琮
(1892-1974)
浙江嘉兴人
数学史家、天文史家、数学教育家
浙江大学数学系首任系主任

钱宝琮，朱世杰垛积术广义. 学艺. 第4卷第7号, 1923
杜石然，朱世杰研究，《宋元数学史论文集》（钱宝琮编）, 科学出版社, 1966，166-209页
罗见今，朱世杰—范德蒙公式的发展简介，数学传播，32(4), 66-71, 2008



一一十五乘之得二百一十又以底子加二一
十六乘之得三千三百六十又以底子加三一
十七乘之得五萬七千一百二十又以底子加
四一十八乘之得一百二萬八千一百六十為
實一百二十而一得八千五百六十八束即撒
星更落一形茭草積也合問
說曰以如像招數喻之所招之亦如三角垛所
給之米如茭草積三角垛乘茭草積爲之即撒
星更落一形也今有官司依三角垛招兵初段
給米一升次日轉多一升次段所招當日給米
四元玉鑑細草
一升次日轉多一升三段以下皆如是今招一
十四日問招兵及給米各幾何答曰兵五百六
十人米八十五石六斗八升術曰求兵者置今
招以今招加一乘之又以今招加二乘之為實
六而一求米者置今招以今招加一乘之又以
今招加二乘之又以今招加三乘之又以今招
加四乘之為實一百二十而一
撒星更落一形圖
三角垛 乘得數 茭草積
一 一〇五 一〇五

[清]罗士琳撰《四元玉鉴细草》

$$\sum_{k=1}^{n}\binom{k+2}{3}\binom{n+2-k}{2}=\sum_{k=1}^{n}\binom{k+4}{5}$$

组合恒等式

- 生成函数（generating function）法
  - $\dfrac{1}{1-x}=1+x+x^2+\cdots+x^n+\cdots=\sum_{l=0}^{\infty}x^l$
  - $\dfrac{1}{(1-x)^{p+1}}=\sum_{l=0}^{\infty}\dfrac{(l+p)(l+p-1)\cdots(l+1)}{p\cdot(p-1)\cdots 1}x^l=\sum_{l=0}^{\infty}\binom{l+p}{p}x^l$
  - $\sum_{l=0}^{\infty}\binom{l+i+1}{i+1}x^l=\dfrac{1}{(1-x)^{i+2}}=\dfrac{1}{(1-x)^{j+1}}\dfrac{1}{(1-x)^{i-j+1}}=\sum_{l=0}^{\infty}\binom{l+j}{j}x^l\cdot\sum_{l=0}^{\infty}\binom{l+i-j}{i-j}x^l$
  - 比较两端 x^{n-i} 项系数
    - $\binom{n+1}{i+1}=\sum_{k=0}^{n-i}\binom{k+j}{j}\binom{n-k-j}{i-j}=\sum_{k=j}^{n}\binom{k}{j}\binom{n-k}{i-j}$

从期望到决策

数学建模

- **期望**
  - 在录用相对名次为y_i的应聘者A_i情况下，录用者绝对名次的期望值为

$$E(A_i \mid y_i = j) = \frac{j}{\binom{n}{i}} \sum_{k=1}^{n} \binom{k}{j}\binom{n-k}{i-j} = \frac{j}{\binom{n}{i}}\binom{n+1}{i+1} = \frac{n+1}{i+1} j$$

- **决策**
  - 面试 A_i 后招聘方可作录用和不录用**两种**决策
    - 若录用 A_i，招聘结束
    - 若不录用A_i，继续面试 A_{i+1}

$$\boxed{\binom{n+1}{i+1} = \sum_{k=1}^{n} \binom{k}{j}\binom{n-k}{i-j}}$$

递推

- 令 $U(j,i)$ 为面试相对名次 $y_i = j$ 的 A_i 时可能取得的**最优**绝对名次期望值
  - 录用 A_i：$U(j,i)=\dfrac{n+1}{i+1}j$
  - 若不录用 A_i，$y_{i+1}=k$ 的概率为 $\dfrac{1}{i+1}$：$U(j,i)=\dfrac{1}{i+1}\sum_{k=1}^{i+1}U(k,i+1)$
  - $U(j,i)=\min\left\{\dfrac{n+1}{i+1}j,\dfrac{1}{i+1}\sum_{k=1}^{i+1}U(k,i+1)\right\}$ 录用与否由两值大小比较决定
  - 面试 A_n 时，相对名次即为绝对名次，招聘方必定录用，$U(j,n)=j,\ j=1,\cdots,n$
  - 自首位应聘者面试起，招聘方采用正确决策所能得到的最优绝对名次期望值为 $U(1,1)$

递推

- 从 $U(j,i)$ 到 C_{i-1}
  - 令 $C_{i-1}=\frac{1}{i}\sum_{k=1}^{i}U(k,i)$，则 $U(j,i)=\min\left\{\frac{n+1}{i+1}j, C_i\right\}$
  - 令 $s_i=\left\lfloor\frac{C_i(i+1)}{n+1}\right\rfloor, i=n-1,\cdots,1$，$\frac{n+1}{i+1}j\le C_i$ 当且仅当 $j\le s_i$
  - $C_{i-1}=\frac{1}{i}\sum_{j=1}^{i}U(j,i)=\frac{1}{i}\sum_{j=1}^{i}\min\left\{\frac{n+1}{i+1}j, C_i\right\}$

    $=\frac{1}{i}\left(\frac{n+1}{i+1}(1+\cdots+s_i)+(i-s_i)C_i\right)=\frac{1}{i}\left(\frac{n+1}{i+1}\cdot\frac{s_i(1+s_i)}{2}+(i-s_i)C_i\right)$
  - $C_{n-1}=\frac{1}{n}\sum_{k=1}^{n}U(k,n)=\frac{1}{n}\sum_{k=1}^{n}j=\frac{n+1}{2}$，$U(1,1)=C_0$

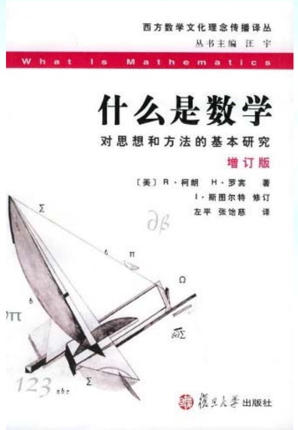
录用与否由两值大小比较决定

$$\boxed{U(j,i)=\min\left\{\frac{n+1}{i+1}j, \frac{1}{i+1}\sum_{k=1}^{i+1}U(k,i+1)\right\} \qquad U(j,n)=j, j=1,\cdots,n}$$

最优决策

数学建模

- 最优策略 $(s_n = n)$
  - 面试 A_i 时，若 $y_i \le s_i$，则录用 A_i，否则继续面试 A_{i+1}
- 数值结果
  - $n=4$ 时，$C_3=\dfrac{5}{2}, C_2=\dfrac{25}{12}, C_1=\dfrac{15}{8}, C_0=\dfrac{15}{8}$，$s_4=4, s_3=2, s_2=1, s_1=0$
  - 记 τ_n 为应聘者数量为 n 时的最优名次期望
    - $\tau_{10}=2.56, \tau_{100}=3.60, \tau_{1000}=3.83$
    - $$\lim_{n\to\infty}\tau_n=\prod_{j=1}^{\infty}\left(\frac{j+2}{j}\right)^{\frac{1}{j}+1}\approx 3.8695$$



Herbert Ellis Robbins
(1915-2001)
美国数学家

Courant R, Robbins H, *What Is Mathematics?An Elementary Approach to Ideas and Methods*, 1941年初版(中译本：左平、张饴慈译，复旦大学出版社2005年)

Chow YS, Moriguti S, Robbins H, Samuels SM, Optimal selection based on relative rank (the “secretary problem”), *Israel Journal of Mathematics*, 2, 81-90, 1964

问题推广

数学建模

- 第四个**Secretary Problem**（**Robbins**问题）
  - 若 A_i 的分布已知，招聘方录用一位应聘者，采用何种策略可使录用者绝对名次期望值尽可能小
- 其他推广
  - 招聘方拟录用某应聘者时，应聘者以 $p(0<p<1)$的概率接受聘用
  - 招聘方可以录用在当前面试者之前第 r个接受面试的应聘者，应聘者仍然接受聘用的概率为$q(r)$，$q(r)$为 r 的非增函数
  - 应聘者数目为一随机变量

Freeman PR. The secretary problem and its extensions: A review. ***International Statistical Review*****, 51: 189-206, 1983.**


浙江大学
ZheJiang University
谢 谢