HITP

数学奥林匹克系列

2006·第一辑

数学奥林匹克与数学文化

Mathematical Olympiads and Mathematical Culture

刘培杰　主编

《数学奥林匹克与数学文化》是数学竞赛与数学文化方面的系列专业文集。本文集旨在为从事数学竞赛的师生与从事数学文化研究与传播的专业人员提供深度阅读，搭建表达平台，促进海内外华人同业人士的学术交流与合作，推动数学的普及与进步。

哈尔滨工业大学出版社

HITP

数学奥林匹克系列

2006·第一辑

数学奥林匹克与数学文化

Mathematical Olympiads and Mathematical Culture

刘培杰　主编

哈尔滨工业大学出版社

图书在版编目(CIP)数据

数学奥林匹克与数学文化.第1卷/刘培杰主编.
哈尔滨:哈尔滨工业大学出版社,2006.6
ISBN 7-5603-2341-3

Ⅰ.数…　Ⅱ.刘…　Ⅲ.数学课-中学-竞赛题-研究
Ⅳ.G634.505

中国版本图书馆CIP数据核字(2006)第026851号

策划编辑　刘培杰　责任编辑　康云霞　杨明蕾
封面设计　卞秉利
出版发行　哈尔滨工业大学出版社
社　　址　哈尔滨市南岗区复华四道街10号　邮编150006
传　　真　0451-86414749
网　　址　http://hitpress.hit.edu.cn
印　　刷　黑龙江省教育厅印刷厂
开　　本　787mm×1092mm　1/16　印张31.5　字数565千字
版　　次　2006年6月第1版　2006年6月第1次印刷
印　　数　1~3 000册
定　　价　48.00元

数学奥林匹克与数学文化

Mathematical Olympiads and Mathematical Culture

2006·第一辑

目 录 CONTENTS

莎士比亚新诗真伪之鉴定

黄文璋[①]

1985年11月14日,英国及美国的报纸及杂志争相报导研究莎士比亚(Shakespeare,1564—1616,英国剧作家及诗人)的美国学者Gary Taylor,在英国牛津大学(Oxford University)的Bodleian图书馆(是英国历史最悠久、最重要,且藏书只供参考概不外借的图书馆之一,现有馆藏420余万册印刷本及5万余册手稿本)的一本自1755年起收藏的书中,找到一首很可能是莎士比亚的抒情诗(见Lelyveld(1985)及Taylor(1985)).该诗的第二抄本稍后在美国耶鲁大学(Yale University)的图书馆出现.我们姑且先称这首诗为泰勒诗(Taylor poem).如果能证实泰勒诗确为莎士比亚所作,则这将是17世纪以来,莎士比亚作品最重要的一次发现.英美学者为了这首诗之真伪争论不休,大打笔战.只是不少专家认为这首泰勒诗不论在用字遣词与韵味风格上,都迥异于莎士比亚其他的作品.有趣的是,统计学者也介入这场纷争.1986年1月24日出版的Science杂志,刊登了一篇"莎士比亚的新诗——向统计学礼赞"(Shakespeare's new poem: an ode to statistics),介绍著名的统计学者,美国斯坦福大学(Stanford University)的Efom教授及芝加哥大学(University of Chicago)的Thisted教授,如何以统计的方法鉴定这首新出土的泰勒诗,是否为莎士比亚所作.

早在1976年,Efom和Thisted便把莎士比亚作品中所用的字作了一番统计分析.他们想回答若发现了一新的作品,如何经由统计分析(statistical analysis)其中所用的字,以判断此作品是否为莎士比亚所作这一问题.他们当初作此分析只是为了好玩,没想到10年后真派上了用场.

这并非统计学家第一次协助解决文学上的问题.而由于统计分析是如此的具说服力,因此往往能使一些文学上长期的争论迅速地平息.例如,哈佛大学(Harvard University)的Mosteller及Wallace利用统计方法,判定The Federalist Pa-

① 中国台湾高雄大学,应用数学教授兼副校长,美国普渡大学统计博士.

pers 的作者为 James Madison，而非 Alexander Hamiliton. 在 19 世纪 50 年代，著名的英国统计学者 David Cox 及文学家 L. Brandwood 也曾利用统计方法，以解决争论长达 1 000 年的关于柏拉图(Plato，约公元前 428—347 年)众多作品的先后顺序问题. 柏拉图为古希腊三大哲学家之一，和苏格拉底(Socrates，约公元前 470—399 年)及亚里士多德(Aristotle，公元前 384—322 年)，共同奠定西方文化的哲学基础. 柏拉图本来对民主抱着希望，但在苏格拉底被判死刑后，他终于认清一个有良知的人，在活跃的政治中是无容身之处的(There is no place for a man of conscience in active politics)，颇令人深省.

虽然利用统计方法来解答文学上的问题之想法并非创新，Efom 与 Thisted 所采用的特别方法却不曾在这方面用过. 此方法可追溯至 19 世纪 40 年代，那时生物学家 Williams 向英国统计学家费雪(R. A. Fisher，1890—1962，现代统计学的奠基者，其生平事迹可参考 Gower(1990/91))，提出一个似乎不可能回答的问题. Williams 曾前往马来西亚采集蝴蝶，他把自己共见过几种(species)蝴蝶，以及每种各见过的次数(有些见过几十次，有些见过几次，有些只见过一次)，都告诉费雪. Williams 想知道在马来西亚的蝴蝶中，他没见过的究竟有多少?

一般人会觉得此问题毫无头绪，不可能解答. 但统计学家却有办法估计尚有几种蝴蝶还未被捕捉到. 只要假设蝴蝶是依照每一种之只数的比例，随机地(randomly)被捕捉. 而这只要假设每一件事(包括蝴蝶的分布，捕捉的技术等)随时都很均匀，不会先是把某一种蝴蝶捕捉殆尽，之后再大量捕捉另一种. 若有某一种蝴蝶尚未被捕捉到，那纯粹只是运气的关系，而非该种蝴蝶特别会躲藏. 费雪所用方法之细节，由于蛮技术性的，此处无法细说，可参考 Fisher et al. (1943).

Efron 与 Thisted 介入文学是很偶然的. 有一次他们聆听加州大学 Santa Barbara 分校(University of California at Santa Barbara)的 Gani(1924—，著名的应用机率①学者，Journal of Applied Probability，Advances in Applied Probability 等重要期刊均为他所创. 他曾到过台湾)的演讲. Gani 的目的是要分析莎士比亚作品的结构. 在演讲中他提到德国 Münster 的 Westfälische Wilhelms Universität 的 Marvin Spevack，已将莎士比亚的所有作品输入计算机中，并已计算出莎士比亚所用过的全部字数，及每一个字使用过的次数，见 Spevack(1968). 听完演讲后，Efom 与 Thisted(那时为 Efron 的学生)决定把费雪所用的方法拿来分析莎士比亚的作品. Efron and Thisted(1976)一文就是他们的研究报告，发表在统计学中极权威的学术期刊 Biometrika.

这一篇文章的题目为：Estimating the number of unseen species：How many

① 机率即概率(编者注).

words did Shakespeare know? 如前所述,在生态学中往往会估计某生物尚未见到的种数.而在该文中,尚未见到的"种",却是莎士比亚知道但不曾用过的字.

莎士比亚全部作品(以下简称总作品)之总字数为 884 647,其中有 14 376 个相异字只出现一次,4 343 个相异字只出现两次.表 1(见 Efron and Thisted (1976))列出只出现 100 次的字数.令 n_x 表示出现 x 次的字数.当 $x=43$,即列 40+,行 3,其对应的 $n_x=30$,表出现 43 次的字共有 30 个,余类推.

在总作品中,莎士比亚共用了 31 534 个不同的字,即

$$\sum_{x=1}^{\infty} n_x = 31\ 534$$

必须注意的是像"girl"与"girls"乃视为不同的字.由表 1,出现次数不超过 100 的有 30 688 个字,因此有 846 个字出现的次数超过 100.

Efron 与 Thisted 分别依据 Fisher et al.(1943)所用的参数模式(parametric model)及 Good and Toulmin(1956)所用的非参数模式(nonparametric model),来估计莎士比亚尚认识多少字,所得到的估计值(estimate)都差不多是 11 460,且标准差(standard deviation)小于 150.

表 1　莎士比亚总作品中字出现之频率

x	1	2	3	4	5	6	7	8	9	10	小计
0+	14 376	4 343	2 292	1 463	1 043	837	638	519	430	364	26 305
10+	305	259	242	223	187	181	179	130	127	128	1 961
20+	104	105	99	112	93	74	83	76	72	63	881
30+	73	47	56	59	53	45	34	49	45	52	513
40+	49	41	30	35	37	21	41	30	28	19	331
50+	25	19	28	27	31	19	19	22	23	14	227
60+	30	19	21	18	15	10	15	14	11	16	169
70+	13	12	10	16	18	11	8	15	12	7	122
80+	13	12	11	8	10	11	7	12	9	8	101
90+	4	7	6	7	10	10	15	7	7	5	78

事实上在 Efron and Thisted(1976)一文中,他们想回答更一般的问题.他们写着"假设有另一很大量的莎士比亚的作品被发现,譬如说共有 884 647 t 个字,则除了那 31 534 个不同的字外,预期可找到几个新字?"莎士比亚共认识几个字就对应 $t=\infty$ 的情况.由于自 17 世纪之后,便没有莎士比亚的新作品出现,所以 Efron 与 Thisted 从未想过会有机会真正地以莎士比亚的作品来检验他

们的理论.他们做此研究的动机纯粹是觉得有趣.据 Efron 说"It never possibly occurred to me that we'd have a chance to use it."甚至当他从新闻上获知这首泰勒诗出现,且有可能是莎士比亚所作,他一时并没想到他其实曾与 Thisted 做过这方面的研究.直到 Thisted 提醒他,他才想到 10 年前所做的工作.尘封已久的莎士比亚又浮现在眼前.

这首泰勒诗与总作品相比,相当短,共只有 429 个字.Efron 与 Thisted 基于罕用字(unusual words)出现的频率(frequency of occurrence),发展出一在统计上算是简单的检定法.有一些字是大家常用的,如"a, is, the",在每一篇文章中可能都出现不少次.但对罕用字,每个作者使用的习惯可能便不同了.Efron 与 Thisted 指出,在总作品中,罕用字的使用非常普遍,在全部使用的 31 534 个相异字中,有接近 2/3 的比例(共 21 011 个),只用了不超过 3 次(表 1).

这首泰勒诗虽只有 429 个字,但其中包含 258 个相异字.令 m_x 表示在泰勒诗中所使用的字,在总作品中出现 x 次的相异字数.表 2 列出 $x=0\sim99$ 的统计.例如,泰勒诗中有 7 个相异字在总作品中只出现过 1 次,即 $m_1=7$,有 5 个相异字在总作品中出现过 23 次,即 $m_{23}=5$,余类推.表 2 中包含 118 个相异字,泰勒诗中另有 140 个相异字在总作品中出现 100 次以上.其中特别值得注意的是 $m_0=9$,此为泰勒诗中不曾在总作品中出现的相异字数.这就是 Efron and Thisted(1976)一文中所要估计的量,在该文中采用的符号为 $\Delta(t)$.

Efron 与 Thisted 将他们的结果整理为 Thisted and Efron(1987)一文,题目就叫"Did Shakespeare write a newly—discovered poem?"表 3 即为他们求出的若泰勒诗真为莎士比亚所作,表 2 中 m_x 之期望值(expectation)之估计值.

表 2　泰勒诗在总作品中出现 x 次的相异字 m_x

x	1	2	3	4	5	6	7	8	9	10	小计
0+	9	7	5	4	4	2	4	0	2	3	40
10+	1	0	3	0	1	1	1	2	1	0	10
20+	2	2	1	5	3	1	0	2	2	3	21
30+	4	1	1	1	2	1	0	0	3	3	16
40+	1	2	0	0	2	1	1	2	1	1	11
50+	0	1	1	1	1	0	0	1	0	2	7
60+	0	1	0	0	1	1	0	0	1	0	4
70+	0	0	1	0	0	1	0	0	1	1	4
80+	0	0	1	1	0	0	0	0	0	0	2
90+	0	0	0	1	0	1	1	0	0	0	3

表 3　表 2 中 m_x 之期望值之估计值 $\hat{v}_x$

x	1	2	3	4	5	6	7	8	9	10
0+	6.97	4.21	3.33	2.84	2.53	2.43	2.16	2.01	1.87	1.76
10+	1.62	1.50	1.52	1.51	1.36	1.38	1.33	1.28	1.25	1.22
20+	1.18	1.16	1.13	1.11	1.09	1.06	1.04	1.02	1.00	0.98
30+	0.96	0.94	0.93	0.91	0.90	0.88	0.86	0.85	0.83	0.82
40+	0.80	0.79	0.77	0.76	0.75	0.74	0.73	0.72	0.70	0.69
50+	0.68	0.67	0.66	0.65	0.64	0.63	0.62	0.61	0.60	0.59
60+	0.58	0.57	0.56	0.55	0.54	0.53	0.52	0.51	0.50	0.50
70+	0.49	0.48	0.48	0.47	0.47	0.46	0.45	0.45	0.44	0.44
80+	0.43	0.42	0.42	0.41	0.41	0.40	0.39	0.39	0.38	0.38
90+	0.37	0.36	0.36	0.35	0.35	0.34	0.34	0.33	0.32	0.32

由表 3,若此泰勒诗确为莎士比亚所作,则 Efron 与 Thisted 估计其中含有 $\Delta(t)=6.97$ 个在总作品中未曾出现的相异字.若再考量标准差,则此首新诗约包含 6.97 ± 2.64 个新字,即新字的数目约介于 $4.33\sim9.61$ 间.实际的新字有 9 个(即表 1 中的 m_0),的确在 4.33 与 9.61 间.估计曾出现一次的字的 4.21 ± 2.05 个,实际则为 7,估计曾出现两次的字有 3.33 ± 1.83 个,实际则为 5.Efron 与 Thisted 一直分析到曾出现 100 次的字,其吻合程度皆相当惊人,可以说通过严格的统计检定(quite delicate statistical tests).看起来这首诗确是莎士比亚所写,或者用比较保守的统计术语来说:没有足够的证据可推翻此诗为莎士比亚所做之假设.

传奇统计学者 Persi Diaconis(当时在斯坦福大学任教,现已转至哈佛大学统计系.1945 年出生于纽约,14 岁时离开学校在街头变魔术为生,24 岁时到纽约市立学院(City College)的夜间部就读,1971 年毕业随即进入哈佛大学统计系,而于 1974 年完成博士学位.DeGroot(1986)为一篇对他的访问稿,戴贞德(1987)为节译稿),他熟悉 Efron 及 Thisted 的分析.据他讲,他第一次读这首泰勒诗时,觉得这一点也不像莎士比亚的作品,他认为只要做一些数值分析,就可以证明诗中字所摆的位置完全是错的.但读了 Efron 及 Thisted 的分析后,他相信这首诗很可能是莎士比亚所作.

Efron 强调他们的分析并无法证明这首泰勒诗真是莎士比亚写的.但他说这首泰勒诗在罕用字的使用情况,如此吻合莎士比亚的总作品,确实令人惊讶.

Efron 及 Thisted 也对 John Donne(1572—1631),Christophor Marlowe(1564—1593)及 Ben Jonson(1573—1637)等三位约略与莎士比亚同时代的诗人,各取一首诗,另取四首莎士比亚的诗,与这首泰勒诗作比较.此 8 首诗之资料列在表

4.表5是列出8首诗之用字在总作品中之出现频率.只是此处计算是分成11类.例如,在Jonson的诗中,有10个相异字在总作品中出现60~79次.至于Efron及Thisted所作之估计值则列在表6.

表4 8首待分析的诗

编写	说明	总字数	相异字数
1.JON	Ben Jonson;"An Elegy"	411	243
2.MAR	C.Marlowe;four poems	495	272
3.DON	J.Donne;"The Ecstacy"	487	252
4.CYM	Shakespeare:from "Cymbeline"	323	215
5.PUC	from "A Midsummer Night's Dream"	234	156
6.PHO	"The Phoenix and Turtle"	352	216
7.SON	"Sonnets, Nos.12~15"	448	264
8.TAY	Taylor poem	429	258

表5 8首诗之用字依在总作品中出现的频率而分类(如表2)

在总作品之出现次数

诗名	0	1	2	3~4	5~9	10~19	20~29	30~39	40~59	60~79	80~99
1.JON	8	2	1	6	9	9	12	12	13	10	13
2.MAR	10	8	8	16	22	20	13	9	14	9	5
3.DON	17	5	6	5	12	17	14	6	12	3	10
4.CYM	7	4	3	5	13	17	9	12	17	4	4
5.PUC	1	4	0	3	9	6	9	4	5	9	3
6.PHO	14	5	5	9	8	18	13	7	13	8	5
7.SON	7	8	1	5	16	14	12	13	12	13	8
8.TAY	9	7	5	8	11	10	21	16	18	8	5

经过3种严密的统计检定(其过程我们自然无法在此介绍性的短文中讲述),发现对前三首(非莎士比亚之作品),罕用字出现次数之实际值与预测值(假设其为莎士比亚所作)皆不吻合.而虽然挑选的4首莎士比亚的诗偶尔有不吻合处,总地来说是可接受的.本来仅是一师生的游戏之作,说不定当初还被"有识之士"视为纸上谈兵,10年后竟令那些向来可能不常接触统计的文学学者折服.不过与其将此视为无心插柳,不如相信统计的用途是无所不在的.

表 6　表 5 中之相异字之期望值的估计值在总作品之出现次数

诗名	0	1	2	3~4	5~9	10~19	20~29	30~39	40~59	60~79	80~99
1.JON	6.68	4.03	3.19	5.14	9.81	13.61	9.64	8.18	12.68	9.17	6.83
2.MAR	8.04	4.86	3.85	6.19	11.81	15.91	12.03	9.92	14.92	10.72	8.26
3.DON	7.91	4.78	3.78	6.09	11.62	15.59	11.77	9.68	14.99	10.83	8.06
4.CYM	5.25	3.17	2.51	4.04	7.71	10.35	7.82	6.44	9.99	7.23	5.39
5.PUC	3.79	2.29	1.81	2.91	5.57	7.47	5.65	4.66	7.22	5.23	3.91
6.PHO	5.72	3.46	2.73	4.40	8.40	11.28	8.52	7.02	10.87	7.87	5.87
7.SON	7.28	4.40	3.48	5.60	10.69	14.52	11.10	9.06	13.71	10.02	7.96
8.TAY	6.97	4.21	3.33	5.36	10.24	13.96	10.77	8.87	13.77	9.99	7.48

关于估计种类的方法尚有不少,可参考赵莲菊一文[2].

最后我们以 Efron 及 Thisted 特地挑出来咏怀他们心情的一首莎士比亚的诗作为本文之结束.

Why is my verse so barren of new pride,
So far from variation or quick change?
Why with the time do I not glance aside
To new - found methods and to compounds strange?
Why write I still all one, ever the same,
And keep invention in a noted weed,
That every word doth almost tell my name,
Showing their birth and where they did proceed?
O, know, sweet love, I always write of you,
And you and love are still my argument;
So all my best is dressing old words new,
Spending again what is already spent;
For as the sun is daily new and old,
So is my love still telling what is told.

William Shakespeare
Sonnet LXXVI

参考文献

[1] 戴贞德.与统计魔术大师对话录[J].科学月刊,1987,18(5):353-355.
[2] 赵莲菊.种类知多少[J].数学传播季刊,1995,19(2):3-7.

[3] DEGROOT M H. A conversation with Persi Diaconis[J]. Statistical Science, 1995, 19(2): 319-334.
[4] EFRON B, THISTED R. Estimating the number of unseen species: How many words did shakespeare know[J]. Biometrika, 1976(63): 435-447.
[5] FISHER R A, CORBET A S, WILLIAMS C B. The relation between the number of species and the number of individuals in a random sample of an animal population[J]. Journal of Animal Ecology, 1943(12): 42-58.
[6] GOOD I J, TOULMIN G H. The number of new species, and the increase in population converage, when a sample is increased[J]. Biometrika, 1956(43): 45-63.
[7] GOWER J C. Sir Ronald Aylmer Fisher, 1890—1962[J]. Mathematical Spectrum, 1990, 91(23): 76-86.
[8] KOLATA G. Shakespeare's new poem: an ode to statistics[J]. Science, 1986(231): 335-336.
[9] LELYVELD J. A scholar's find: Shakespearean lyru[J]. New York Times, 1985, 1(12).
[10] SPEVACK M. A Complete and Systematic Concordance to the Words of Shakespeare[M]. George Olms, Hildesheim, 1968.
[11] TAYLOR G. Shakespeare's new poem. A scholar's clues and conclusions[M]. New York: New York Times Book Review, 1985: 11-14.
[12] THISTED R, EFRON B. Did Shakespeare write a newlydiscovered poem[J]. Biometrika, 1987(74): 445-455.

(原文载于《数学欣赏》,中国统计出版社,经作者黄文璋同意转载)

编后语 伴随着人们对理性的渴望,数学攀升到了科学宝塔的顶点,于是开始了对社会科学的大举"入侵",先是经济学全面沦陷(数学家康托洛维奇和纳什成为诺贝尔经济学奖得主即为证明).随后社会学与心理学也是半壁江山不保(出现了拓扑心理学等交叉学科).接着开始染指文学艺术这块纯而又纯的社会科学"净土".从用微分方程鉴别名画真伪到判断《红楼梦》后四十回的真正作者,从《静静的顿河》的文风到莎士比亚新诗真伪的鉴定,黄文璋教授既沐欧风美雨,又有深厚国学功底,文章独到,读之爱不释手,并承蒙2005年11月27日亲笔回函,慨然允准重登,为本文集增辉不少.

不知从何时开始,中国大陆的知名数学家已从数学科普领域悄然而退,可能一是专业数学家竞争加剧使他们力争上游,无法分身;二是功利心日重而将精力投向了更有收益的领域;三是志存高远为夺菲尔兹奖而不屑于从事数学科普.但更重要的原因恐怕一是缺乏大师那样通透的理解,"博大精深"之后无法浅出;二是由于历史的原因,新一代的数学家其人文素质及文章技巧远不及华罗庚先生那一辈,便知难而退,于是二、三流科普作品充斥于市,不仅如此,"伪数学家"及"民间数学爱好者"开始占领这一市场,着实可怕.黄文璋先生身兼台湾高雄大学应用数学系教授兼主任、中华概率统计学会理事长(台湾),亲力亲为为统计数学鼓与呼,以引人入胜之题材诱人于统计数学之腹地,令人向往,催人奋进.

几个多项式问题

林常[①]

§1 全 k 次方值蕴涵 k 次方式

给定正整数 $k \geqslant 2$ 和实系数多项式 $f(x)$,如果对每个充分大的正整数 n,$f(n)$ 都是整数的 k 次方,则必有整值多项式 $g(x)$,使得 $f(x) = g(x)^k$.

证明 不妨设 $\deg f = m \geqslant 1, f(x) = \sum_{i=0}^{m} a_i x^{m-i}$,记 $r = \left[\frac{m}{k}\right]$.

$$u(x) = \sqrt[k]{f(x)} = \sqrt[k]{a_0 x^m}\sqrt[k]{1 + \frac{a_1}{a_0 x} + \cdots} \quad (\text{当 } k \text{ 为偶数时取算术根}).$$

利用函数 $(1 + c_1 t + \cdots)^{1/k}$ 在 $t = 0$ 处的马克劳林(Maclaurin)展开,我们有

$$u(x) = \sum_{i=0}^{r} a_i^{(0)} x^{(m/k)-i} + \varepsilon$$

其中 ε 表示 $x \to \infty$ 时的无穷小量.

注意到对 $t > 0$,差分函数

$$\Delta x^t = (x+1)^t - x^t = x^t\left(\left(1 + \frac{1}{x}\right)^t - 1\right) = \sum_{i=1}^{[t]} c_i x^{t-i} + \varepsilon$$

故
$$\Delta u(x) = \sum_{i=1}^{r} a_i^{(1)} x^{(m/k)-i} + \varepsilon$$

归纳可证
$$\Delta^j u(x) = \sum_{i=j}^{r} a_i^{(j)} x^{(m/k)-i} + \varepsilon \quad (j \geqslant 1)$$

最后有
$$\Delta^{r+1} u(x) = \varepsilon$$

也就是
$$\lim_{x \to +\infty} \Delta^{r+1} u(x) = 0$$

现在根据已知条件,对充分大的正整数 n,$u(n)$ 都是整数,由差分的定义

① 林常,福建省教育学院.

$\Delta f(n)=f(n+1)-f(n)$ 可知，$\Delta^j u(n)$ 都是整数，特别，$\Delta^{r+1}u(n)$ 都是整数. 再由上面的极限式得到：存在 n_0，使得 $n\geqslant n_0$ 时 $\Delta^{r+1}u(n)=0$.

于是存在(至多) r 次的整值多项式 $g(x)$，使得 $n\geqslant n_0$ 时 $u(n)=g(n)$，从而

$$f(n)=u(n)^k=g(n)^k$$

最后，由于 f 与 g^k 都是多项式，根据多项式恒等定理得到 $f(x)=g(x)^k$.

单墫老师在《解题研究》(南京师范大学出版社，2002.6) 第五章第 8 节(平方数与平方式)一书中用初等方法证明了 $k=m=2$ 的情形. 对一般的 k,m 恐怕很难用初等方法证明.

§2　Чебышев 多项式引申出的几个问题

n 次 Чебышев 多项式 $T_n(x)$ 归纳定义为：$T_0(x)=1$，$T_1(x)=x$，而对于 $n\geqslant 2$，有

$$T_n(x)=2xT_{n-1}(x)-T_{n-2}(x)$$

容易归纳证明，$n\geqslant 1$ 时 $T_n(x)$ 是首项系数为 2^{n-1} 的 n 次整系数多项式.

归纳可证 $T_n(\cos\theta)=\cos(n\theta)$. 由此可知，$|x|\leqslant 1$ 时 $|T_n(x)|\leqslant 1$. $T_n(x)$ 有 n 个不同实根 $u_k=\cos\dfrac{(2k+1)\pi}{2n}(0\leqslant k\leqslant n-1)$ 和 $n-1$ 个极值点 $v_k=\cos\dfrac{k\pi}{n}(1\leqslant k\leqslant n-1)$，$T_n(v_k)=(-1)^k$，它们在 $[-1,1]$ 中交错分布：$v_0>u_0>v_1>u_1>\cdots$.

同样归纳可证 $T_n(\dfrac{u+u^{-1}}{2})=\dfrac{1}{2}(u^n+u^{-n})$，由此有 $|x|>1$ 时 $|T_n(x)|>1$.

利用特征方程可求出显式

$$T_n(x)=\frac{1}{2}((x+\sqrt{x^2-1})^n+(x-\sqrt{x^2-1})^n)=\sum_{k=0}^{[\frac{n}{2}]}C_n^{2k}(x^2-1)^k x^{n-2k}$$

Чебышев 定理　给定正整数 n，区间 $[\alpha,\beta]$ 和非零实数 a，首项系数为 a 的 n 次实系数多项式 $f(x)$ 在区间 $[\alpha,\beta]$ 上的振幅

$$M(f,[\alpha,\beta])=\max_{x\in[\alpha,\beta]}|f(x)|\geqslant\frac{|a|(\beta-\alpha)^n}{2^{2n-1}}$$

达到最小值的多项式只有一个

$$f(x)=\frac{a(\beta-\alpha)^n}{2^{2n-1}}T_n(\frac{2x-\alpha-\beta}{\beta-\alpha})=$$