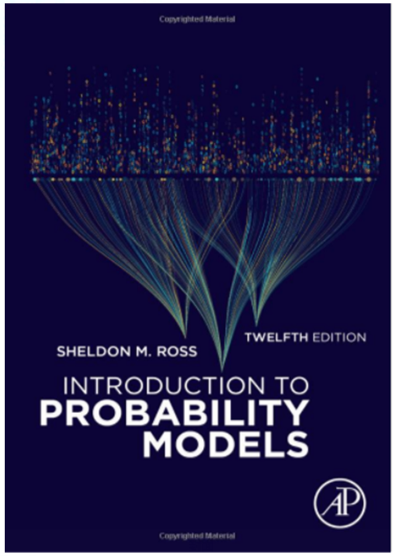
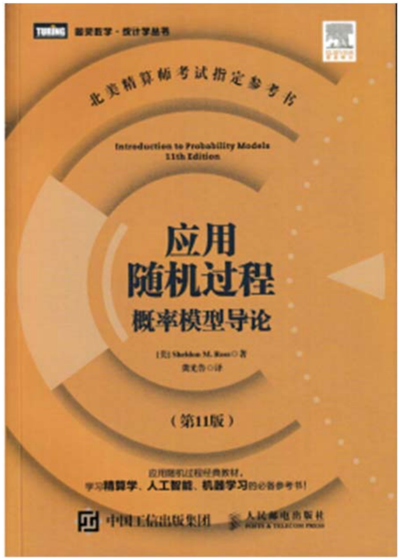
# 数学建模

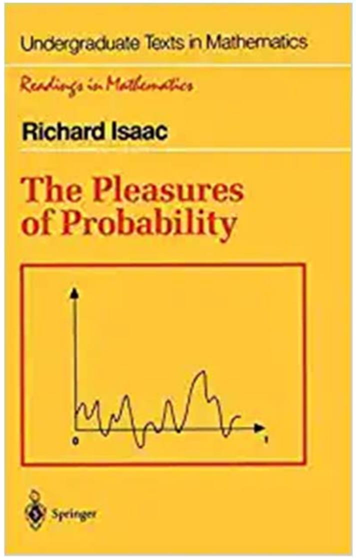
浙江大学数学科学学院　谈之奕

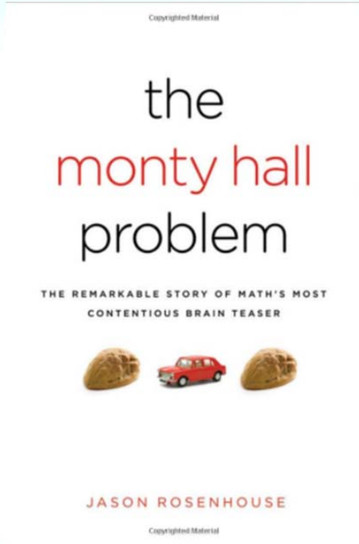
*tanzy@zju.edu.cn*

# 随机模型




Ross SM, *Introduction to Probability Models* (12th), Academic Press, 2019. (中译本：应用随机过程——概率模型导论（第11版），龚光鲁译，人民邮电出版社，2016年)

Isaac R, *The Pleasures of Probability*, Springer, 2016

Rosenhouse J, *The Monty Hall Problem: The Remarkable Story of Math's Most Contentious Brain Teaser*, Oxford University Press, 2009.

# 幸运之门

数学建模

- 舞台上有三扇道具门，其中一扇门后置有一辆汽车，另两扇门后各置有一套汽车杂志
- 节目主持人知道汽车所在位置，门后物品在节目进行过程中不可移动
- 竞猜者可任选其中一扇门并获赠门后物品
- 竞猜者选择了其中一扇门后，主持人打开了另两扇门中的一扇，后面是一套杂志
- 主持人允许竞猜者改变之前的选择，竞猜者为增加获得汽车的可能性，是否应该改变当前的选择

浙江大学
ZheJiang University
数学建模

# 改变 or 不变

- 假设汽车位于1号门后，竞猜者以相同的概率选择1、2、3号门
- 主持人打开的门既不是竞猜者选择的，也不是后置汽车的。若有两扇门符合以上要求，主持人以相同概率选择其中一扇
- 若竞猜者改变选择，他将选择既不是初次选择，也不是主持人打开过的门

| 汽车 | 杂志 | 杂志 |
|---|---|---|
| 1 | 2 | 3 |
| 竞猜者 | →?→ | 主持人 |

# 选择不变

- 若竞猜者不改变选择，则获得汽车的概率为 $\frac{1}{3}$

| 汽车位置 | 竞猜者初次选择 | | 主持人打开 | | 竞猜者再次选择的门 | 获赠物品 |
|---|---|---|---|---|---|---|
| | 门 | 概率 | 门 | 概率 | | |
| 1 | 1 | 1/3 | 2 | 1/6 | 1 | 汽车 |
| | | | 3 | 1/6 | | |
| | 2 | 1/3 | 3 | 1/3 | 2 | 杂志 |
| | 3 | 1/3 | 2 | 1/3 | 3 | 杂志 |

# 选择改变

- 若竞猜者改变选择，则获得汽车的概率为 $\frac{2}{3}$

| 汽车位置 | 竞猜者初次选择 | | 主持人打开 | | 竞猜者再次选择的门 | 获赠物品 |
|---|---|---|---|---|---|---|
| | 门 | 概率 | 门 | 概率 | | |
| 1 | 1 | 1/3 | 2 | 1/6 | 3 | 杂志 |
| | | | 3 | 1/6 | 2 | 杂志 |
| | 2 | 1/3 | 3 | 1/3 | 1 | 汽车 |
| | 3 | 1/3 | 2 | 1/3 | 1 | 汽车 |

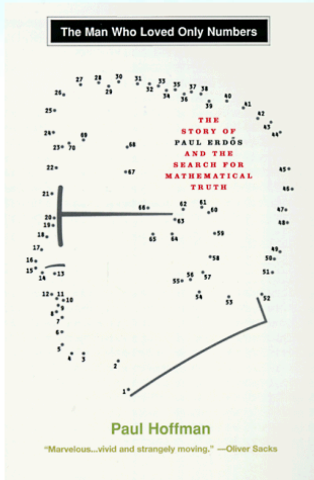
# 改变 or 不变

浙江大学 ZheJiang University

数学建模

Vázsonyi told Erdős about the Monty Hall dilemma. "I told Erdős that the answer was to switch," said Vázsonyi, "and fully expected to move to the next subject. But Erdős, to my surprise, said, 'No, that is impossible. It should make no difference.' At this point I was sorry I brought up the problem, because it was my experience that people get excited and emotional about the answer, and I end up with an unpleasant situation. But there was no way to bow out, of course." Vázsonyi wrote out a "decision tree," not unlike the table of possible outcomes that vos Savant had written out, but this did not convince him. "It was hopeless," Váz-

On his PC Vázsonyi ran a Monte Carlo simulation of the Monty Hall dilemma. Erdős, who never had much use for computers, watched the PC randomly choose whether to switch or stick. The outcome of hundreds of trials favored switching two to one, and Erdős conceded that he was wrong. But the simulation was no more satisfying than




**Paul Hoffman, *The Man Who Loved Only Numbers: The Story of Paul Erdos and the Search for Mathematical Truth*, Hyperion,1999 (中译本：数字情种——埃尔德什传，章晓燕等译，上海世纪出版集团，2009年版)**

浙江大学
ZheJiang University

# 基本数学模型

## 招聘问题

# 招聘问题

- $n$ 位求职者应聘某一职位，招聘方通过逐个面试予以考察
  - 应聘者的综合能力各不相同，通过面试可给出已面试的应聘者的综合能力大小顺序
- 应聘者以某一顺序接受面试，某个应聘者是否被录用必须在他面试结束后立即决定
  - 若录用，招聘即告结束，不再面试其他应聘者
  - 若不录用，招聘方继续面试下一位应聘者
  - 招聘方不得录用曾作出过不录用决定的应聘者
- 招聘方采用何种策略可使招聘到综合能力最强（第一名）的应聘者的概率最大

# 可行策略

- 备选者
  - 若目标为招聘到第一名的概率尽可能大，招聘方只会录用比之前所有应聘者均优的那位应聘者
- 策略 $s$
  - 从第 $s$ 位应聘者开始，录用首次出现的一名备选者
    - 若至最后一名应聘者面试时仍未有备选者出现，是否录用最后一名应聘者对结果没有影响

招聘方如何根据当前应聘者在所有已面试应聘者中的相对名次推测他是否为所有应聘者中的第一名

录用——后来者更好怎么办？不录用——后来者更差怎么办？

# 数学描述



- 应聘者、绝对名次、相对名次
  - 应聘者中综合能力居于第 $i$ 名，$i=1,\cdots,n$ 的应聘者记为 $i$，$i$ 称为其绝对名次
  - 第 $i$ 位，$i=1,\cdots,n$ 接受面试的应聘者记为 $A_i$
    - $1\le A_i\le n$，$A_i=j$ 表示第 $i$ 位接受面试的应聘者绝对名次为 $j$
  - $A_i$ 在前 $i$ 位接受面试的应聘者在 $A_1,\cdots,A_i$ 中综合能力名次排名记为 $y_i$，称为 $A_i$ 的相对名次
    - $1\le y_i\le \min\{i,A_i\}$，若 $y_i=1$，则 $A_i$ 为一名备选者
    - 招聘方只能根据每位应聘者的相对名次进行决策
- 假设对 $1,2,\cdots,n$ 的任一排列，应聘者以该顺序面试的概率均为 $\frac{1}{n!}$

# 实例演示

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1234 | 1 | | | |
| 1243 | 1 | | | |
| 1324 | 1 | | | |
| 1342 | 1 | | | |
| 1423 | 1 | | | |
| 1432 | 1 | | | |

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 2134 | 2 | 1 | | |
| 2143 | 2 | 1 | | |
| 2314 | 2 | 1 | 1 | |
| 2341 | 2 | 1 | 1 | 1 |
| 2413 | 2 | 1 | 1 | |
| 2431 | 2 | 1 | 1 | 1 |

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 3214 | 3 | 2 | 1 | |
| 3241 | 3 | 2 | 1 | 1 |
| 3124 | 3 | 1 | | |
| 3142 | 3 | 1 | | |
| 3421 | 3 | 2 | 2 | 1 |
| 3412 | 3 | 1 | 1 | |

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 4231 | 4 | 2 | 1 | 1 |
| 4213 | 4 | 2 | 1 | |
| 4321 | 4 | 3 | 2 | 1 |
| 4312 | 4 | 3 | 1 | |
| 4123 | 4 | 1 | | |
| 4132 | 4 | 1 | | |

策略 2 在24种可能顺序中有11次可录用到第一名，为最优策略

# 古典概型

数学建模

- $p_n(s)$：采用策略 $s$ 录用到第一名的概率
  - 采用策略1，必录用 $A_1$，$p_n(1)=P(A_1=1)=\dfrac{1}{n}$
- $p_n^k(s)$：采用策略 $s$ 录用 $A_k(k\geq s)$，且为第一名 $(A_k=1)$ 的概率
  - $A_s,\cdots,A_{k-1}$ 均不是备选者
  - 前 $k-1$ 位应聘者中的最佳者出现在前 $s-1$ 位应聘者中，否则他将先于 $A_k$ 被录用

# 古典概型

$$p_n^k(s)=\frac{1}{n}\cdot\frac{s-1}{k-1}$$

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | $s$ | | | | $k$ | | |

第一名出现在第 $k$ 位

前 $k-1$ 位中的最佳者只能选择在前 $s-1$ 位

前 $k-1$ 位中的最佳者

第一名

选 $k-1$ 位应聘者

前 $k-1$ 位中最佳者的可能位置数

其余 $k-2$ 位的可能排列数

其余 $n-k$ 位的可能排列数

$$p_n^k(s)=\frac{\binom{n-1}{k-1}(s-1)(k-2)!(n-k)!}{n!}=\frac{(n-1)!(s-1)(k-2)!(n-k)!}{n!(k-1)!(n-k)!}=\frac{s-1}{n(k-1)}$$

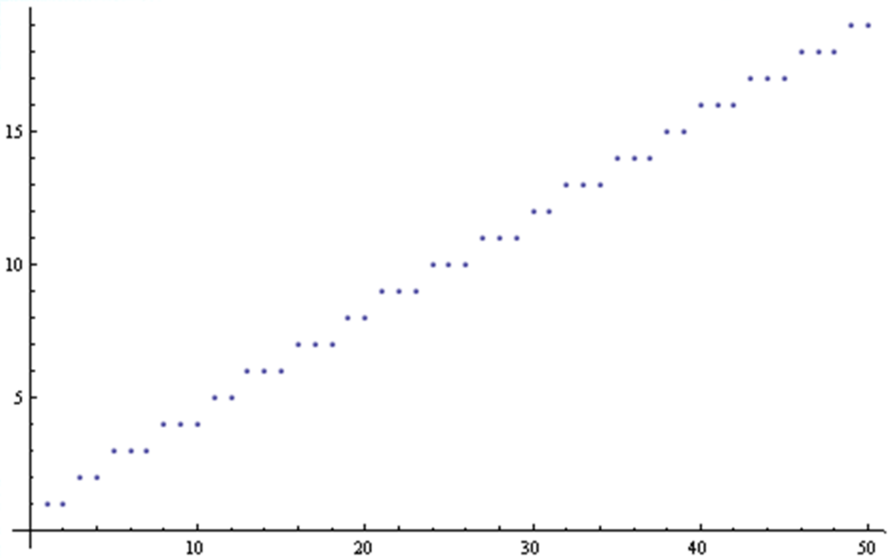
# 概率计算

数学建模

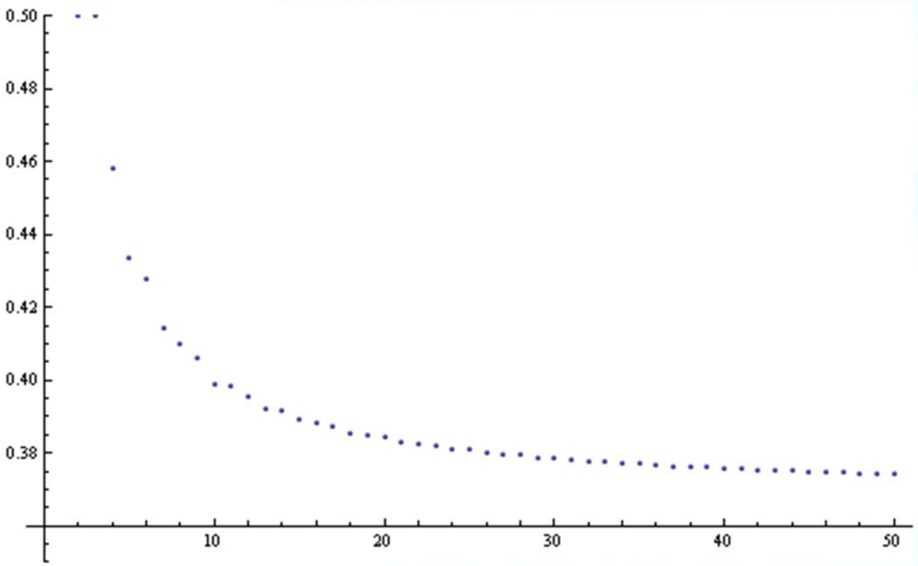
- $p_n(s)=\sum_{k=s}^{n} p_n^k(s)=\sum_{k=s}^{n}\frac{s-1}{n(k-1)}=\frac{s-1}{n}\sum_{k=s}^{n}\frac{1}{k-1}=\frac{s-1}{n}\sum_{k=s-1}^{n-1}\frac{1}{k}$
  - $p_n(s)-p_n(s-1)=\frac{s-1}{n}\sum_{k=s-1}^{n-1}\frac{1}{k}-\frac{s-2}{n}\sum_{k=s-2}^{n-1}\frac{1}{k}=\frac{1}{n}\sum_{k=s-1}^{n-1}\frac{1}{k}-\frac{1}{n}$
  - $p_n(s)-p_n(s+1)=\frac{s-1}{n}\sum_{k=s-1}^{n-1}\frac{1}{k}-\frac{s}{n}\sum_{k=s}^{n-1}\frac{1}{k}=\frac{1}{n}-\frac{1}{n}\sum_{k=s}^{n-1}\frac{1}{k}$
- 令 $s^*=\min\left\{s\,|\,\sum_{k=s}^{n-1}\frac{1}{k}<1\right\}$
  - $s\geq s^*$时，$\sum_{k=s}^{n-1}\frac{1}{k}<1, p_n(s)>p_n(s+1)$，$s\leq s^*$ 时 $\sum_{k=s-1}^{n-1}\frac{1}{k}\geq 1, p_n(s)\geq p_n(s-1)$
- 当 $s=s^*$时，$p_n(s)$ 达到最大值 $p_n(s^*)$

$$\boxed{p_n^k(s)=\frac{s-1}{n(k-1)}}$$

# 数值计算




$s^*$ 图像




$p_n(s^*)$ 图像

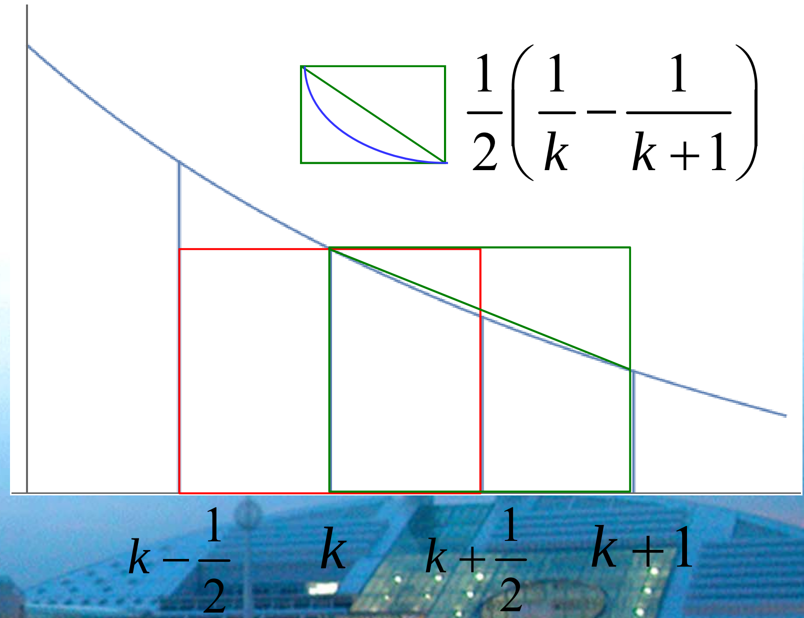
# $s^*$ 的估计

数学建模

- $$\int_k^{k+1}\frac{1}{x}dx+\frac{1}{2}\left(\frac{1}{k}-\frac{1}{k+1}\right)<\frac{1}{k}<\int_{k-\frac{1}{2}}^{k+\frac{1}{2}}\frac{1}{x}dx$$
  - $f(x)=\frac{1}{x}$ 为下凸函数
  - 令 $F(k)=\frac{1}{k}-\int_{k-\frac{1}{2}}^{k+\frac{1}{2}}\frac{1}{x}dx=\frac{1}{k}-\ln\frac{2k+1}{2k-1}$
    - $F(1)=1-\ln 3<0$
    - $$F'(k)=-\frac{1}{k^2}-\frac{1}{2k+1}+\frac{1}{2k-1}$$
      $$=\frac{1-2k^2}{k^2(4k^2-1)}<0$$

$$\boxed{s^*=\min\left\{s \,\middle|\, \sum_{k=s}^{n-1}\frac{1}{k}<1\right\}}$$

变求和为积分



# $s^*$ 的估计

- $$1\le \sum_{k=s^*-1}^{n-1}\frac{1}{k} < \int_{s^*-\frac{3}{2}}^{n-\frac{1}{2}}\frac{1}{x}dx = \ln\frac{2n-1}{2s^*-3} \Leftrightarrow e\le \frac{2n-1}{2s^*-3} \Leftrightarrow s^*\le \frac{1}{e}\left(n-\frac{1}{2}\right)+\frac{3}{2}$$
- $$1>\sum_{k=s^*}^{n-1}\frac{1}{k} > \int_{s^*}^{n}\frac{1}{x}dx + \frac{1}{2}\left(\frac{1}{s^*}-\frac{1}{n}\right) = \ln\frac{n}{s^*}+\frac{1}{2}\left(\frac{1}{s^*}-\frac{1}{n}\right)$$

$$\Leftrightarrow e > \frac{n}{s^*} e^{\frac{1}{2}\left(\frac{1}{s^*}-\frac{1}{n}\right)} > \frac{n}{s^*}\left(1+\frac{1}{2}\left(\frac{1}{s^*}-\frac{1}{n}\right)\right) \ge \frac{n}{s^*}\left(1+\frac{1}{2}\left(\frac{1}{\frac{1}{e}\left(n-\frac{1}{2}\right)+\frac{3}{2}}-\frac{1}{n}\right)\right)$$

$$\Leftrightarrow s^* \ge \frac{n}{e}\left(1-\frac{1}{2n}\right)+\frac{n}{2n+3e-1} = \frac{1}{e}\left(n-\frac{1}{2}\right)+\frac{1}{2}-\frac{3e-1}{2(2n+3e-1)}$$

$$\boxed{s^* = \min\left\{s \,\middle|\, \sum_{k=s}^{n-1}\frac{1}{k}<1\right\}}$$

$$\boxed{\int_{k}^{k+1}\frac{1}{x}dx+\frac{1}{2}\left(\frac{1}{k}-\frac{1}{k+1}\right)<\frac{1}{k}<\int_{k-\frac{1}{2}}^{k+\frac{1}{2}}\frac{1}{x}dx}$$

# Secretary Problem

浙江大学 ZheJiang University

数学建模

- $s^*$ 的值
  - $\frac{1}{e}\left(n-\frac{1}{2}\right)+\frac{1}{2}-\frac{3e-1}{2(2n+3e-1)} \le s^* \le \frac{1}{e}\left(n-\frac{1}{2}\right)+\frac{3}{2}$
  - $s^*$ 的上下界差距不超过 $1+\frac{3e-1}{2(2n+3e-1)} \approx 1+\frac{1.79}{n+1.79}$
  - $\lim_{n\to\infty}\frac{s^*}{n}=\frac{1}{e}$
- $p_n(s^*)$ 的值
  - $p_n(s^*)=\frac{s^*-1}{n}\sum_{k=s^*-1}^{n-1}\frac{1}{k}$
  - $\lim_{n\to\infty} p_n(s^*)=\frac{1}{e}$

$$\boxed{p_n(s)=\frac{s-1}{n}\sum_{k=s-1}^{n-1}\frac{1}{k}}$$

$$\boxed{\ln\frac{n}{s^*-1}=\int_{s^*-1}^{n}\frac{1}{x}dx \le \sum_{k=s^*-1}^{n-1}\frac{1}{k} \le \int_{s^*-\frac{1}{2}}^{n-\frac{1}{2}}\frac{1}{x}dx=\ln\frac{2n-1}{2s^*-1}}$$

# 两次选择

浙江大学 ZheJiang University 数学建模

- 招聘方可录用两名应聘者，但对每位应聘者聘用与否的决定仍需在该应聘者面试结束时给出
- 招聘方采用何种策略可使录用的两位应聘者中其中一位为第一名的概率尽可能大
- 策略 $(r,s)(s>r)$：
  - 录用自 $A_r$ 起首次出现的一名备选者，称为第一次录用
  - 若已录用一人，录用不早于 $A_s$ 的一名备选者，称为第二次录用

# 实例演示

| | (1,3) | (2,3) |
|---|---|---|
| **1234** | 1 | |
| **1243** | 1 | |
| **1324** | 1 | |
| **1342** | 1 | |
| **1423** | 1 | |
| **1432** | 1 | |

| | (1,3) | (2,3) |
|---|---|---|
| **2134** | 2 | 1 |
| **2143** | 2 | 1 |
| **2314** | 2 1 | 1 |
| **2341** | 2 1 | 1 |
| **2413** | 2 1 | 1 |
| **2431** | 2 1 | 1 |

| | (1,3) | (2,3) |
|---|---|---|
| **3124** | 3 | 1 |
| **3142** | 3 | 1 |
| **3214** | 3 1 | 2 1 |
| **3241** | 3 1 | 2 1 |
| **3412** | 3 1 | 1 |
| **3421** | 3 2 | 2 1 |

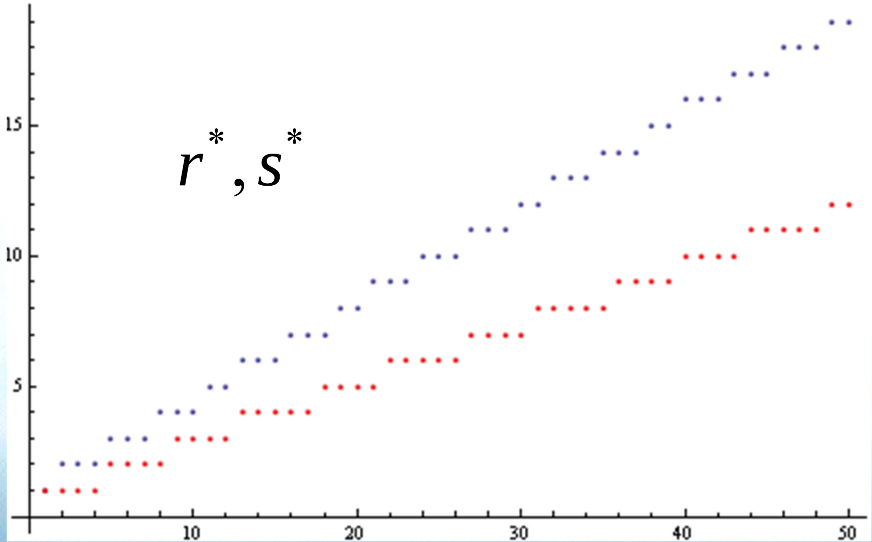
| | (1,3) | (2,3) |
|---|---|---|
| **4123** | 4 | 1 |
| **4132** | 4 | 1 |
| **4213** | 4 1 | 2 1 |
| **4231** | 4 1 | 2 1 |
| **4312** | 4 1 | 3 1 |
| **4321** | 4 2 | 3 2 |

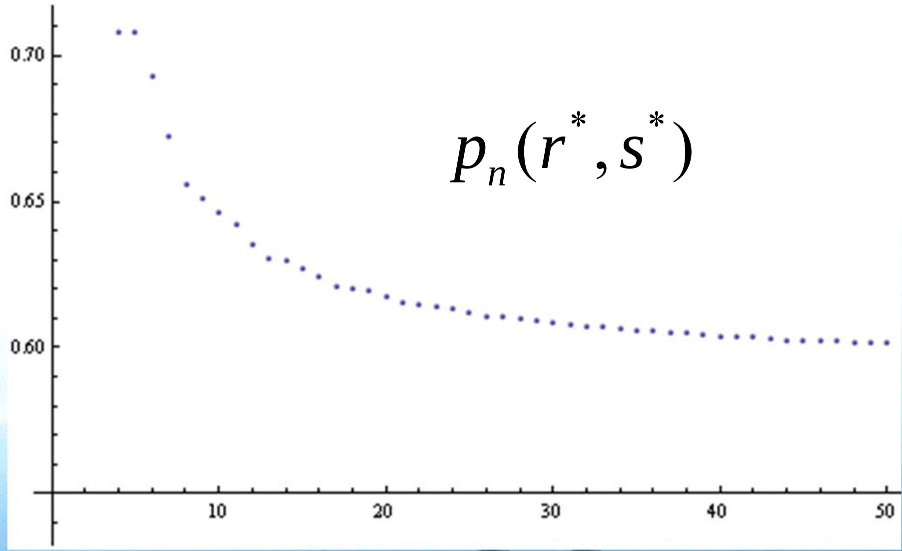
(1,3) 共**16**次　　　(2,3) 共**17**次

# 渐近估计

数学建模

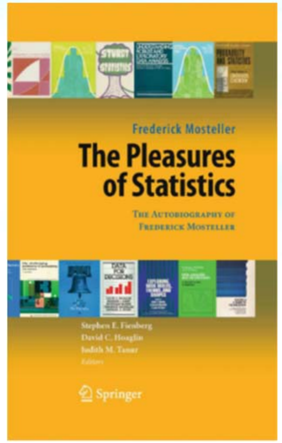
- 两次选择
  - 当 $n$ 充分大时，最优策略为 $(e^{-\frac{3}{2}}n, e^{-1}n)$，录用到第一名的概率约为 0.5910



# 问题推广

- 问题推广
  - 招聘方录用 $k$ 位应聘者，采用何种策略可使其中包含第一名的概率尽可能大
  - 招聘方录用一位应聘者，采用何种策略可使他为前 $k$ 名的概率尽可能大
- 第二个**Secretary Problem**
  - 若 $A_i$ 的分布已知，招聘方录用一位应聘者，采用何种策略可使他为第一名的概率尽可能大




**Frederick Mosteller**
**(1916-2006)**
美国统计学家

**Gilbert J P, Mosteller F. Recognizing the maximum of a sequence, *Journal of the American Statistical Association*, 61, 35-73, 1966.**

**Mosteller F. *The pleasures of statistics: The autobiography of Frederick Mosteller*. Springer, 2010.**