How many days would be holidays if we respect the birthdays of all the past emperors?

Hajime Nanjo, Nobuhiro Shimizu, Taku Yamanaka

2017-12-28

Kuno and Yamanaka Groups

End-of-the Year Presentation

Introduction

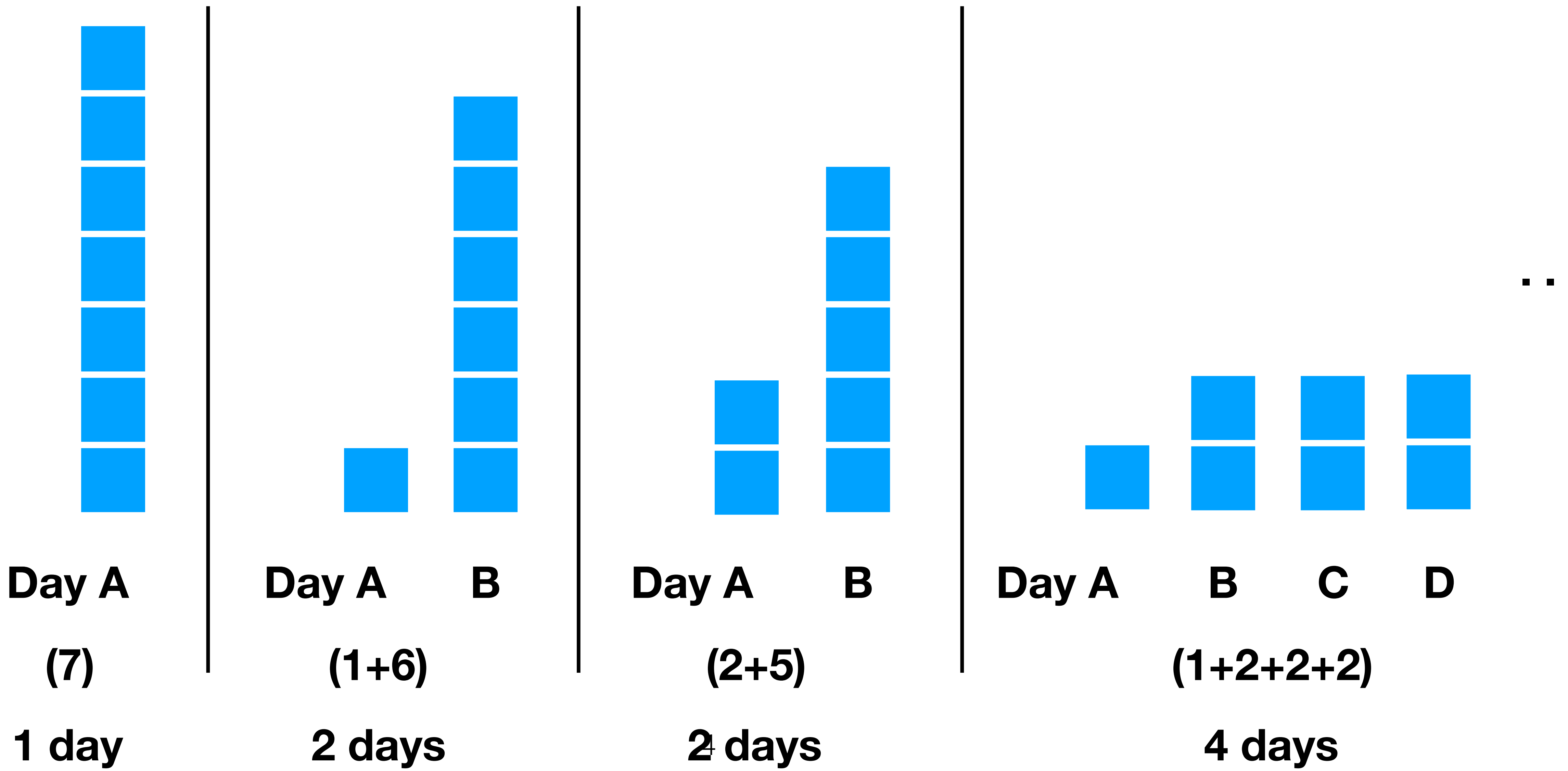
- 125 emperors in Japan in the past
- Birthdays of 3 emperors are holidays
  - Meiji Emperor : Nov. 3, Culture Day
  - Showa Emperor : Apr. 29, Showa Day
  - Current Emperor : Dec. 23, The Emperor’s Birthday

Question

- **If we make the birthdays of all the past 125 emperors as holidays, how many “Emperor’s birthday” holidays will we have?**
- N = 365 (#days/year)
- M = 125 (#emperors)
- n : #holidays

1. Brute-force method

- Split emperors into multiple days
- Example: 7 emperors



Example: #cases for the splitting pattern

Day A B C D

(1+2+2+2)

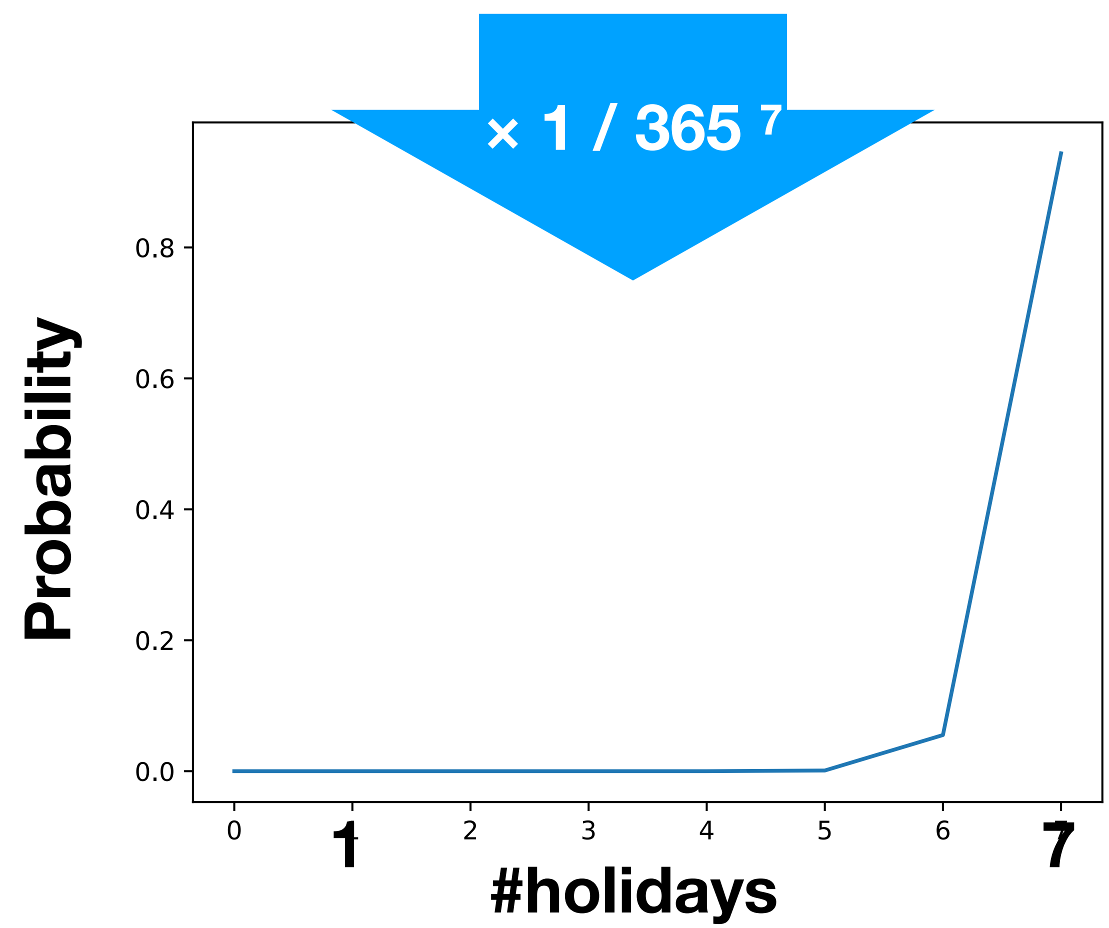
4 days

- ${}_{365}C_1$: pick 1 day
- $\times\ {}_{7}C_1$: pick 1 emperor
- $\times\ {}_{365-1}C_3$: pick 3 days
- $\times\ {}_{7-1}C_{2x3}$: pick 2x3 emperors
- $\times(2x3)!$: Line up emperors
- $\times\ 1/(2!)^3$: Remove double-counts

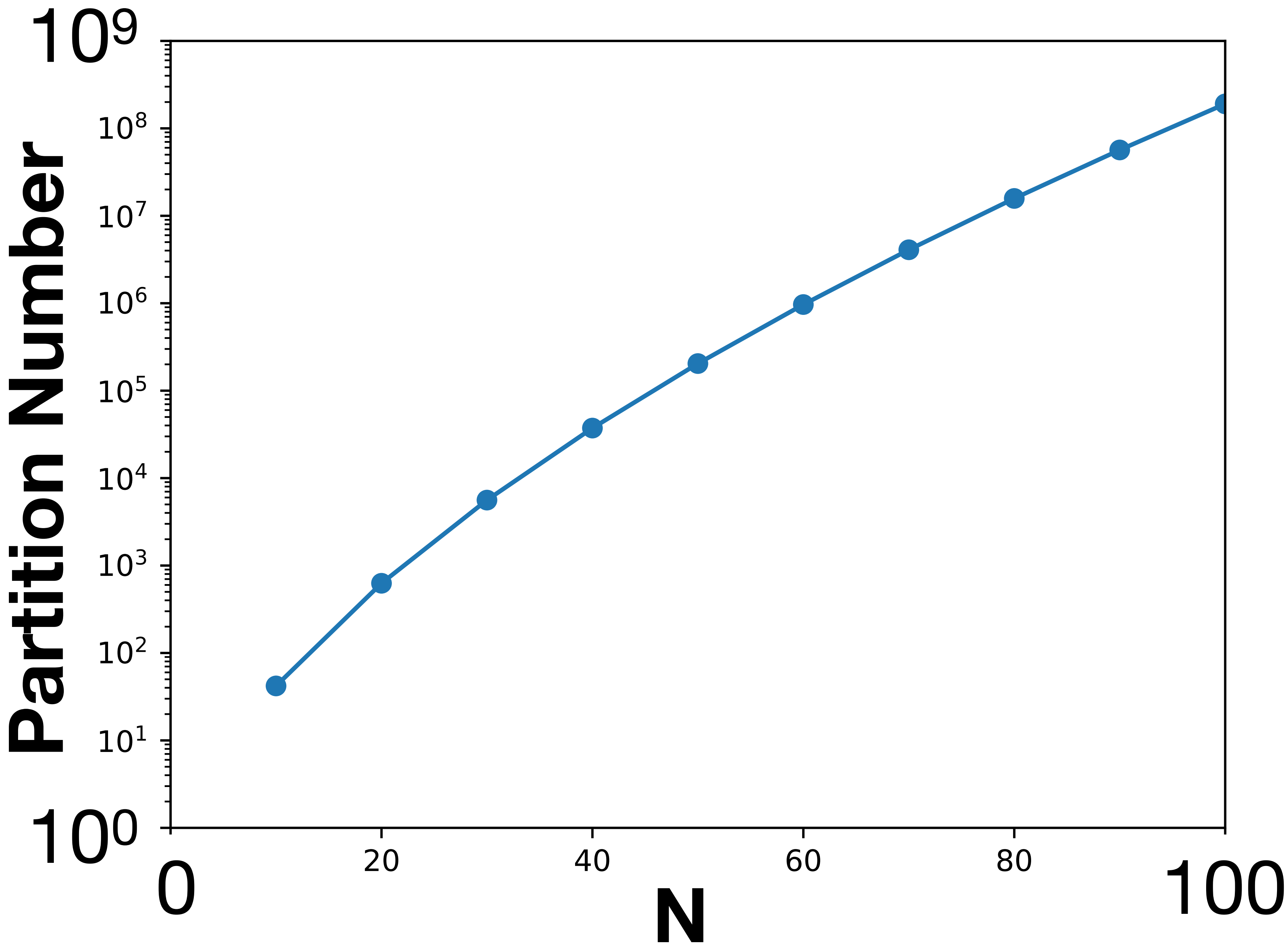
= 1,833,153,121,800

#holidays	1	2	3	4	5	6	7
partitions	(7)	(1+6)	(1+1+5)	(1+1+1+4)	(1+1+1+1+3)	(1+1+1+1+1+2)	(1+1+1+1+1+1+1)
		(2+5)	(1+2+4)	(1+1+2+3)	(1+1+1+2+2)		
		(3+4)	(1+3+3)	(1+2+2+2)			
			(2+2+3)				
#cases	365	837018	1.45E+10	6.11E+12	8.82E+14	4.76E+16	8.15E+17

Partition Number (分割数) = 15

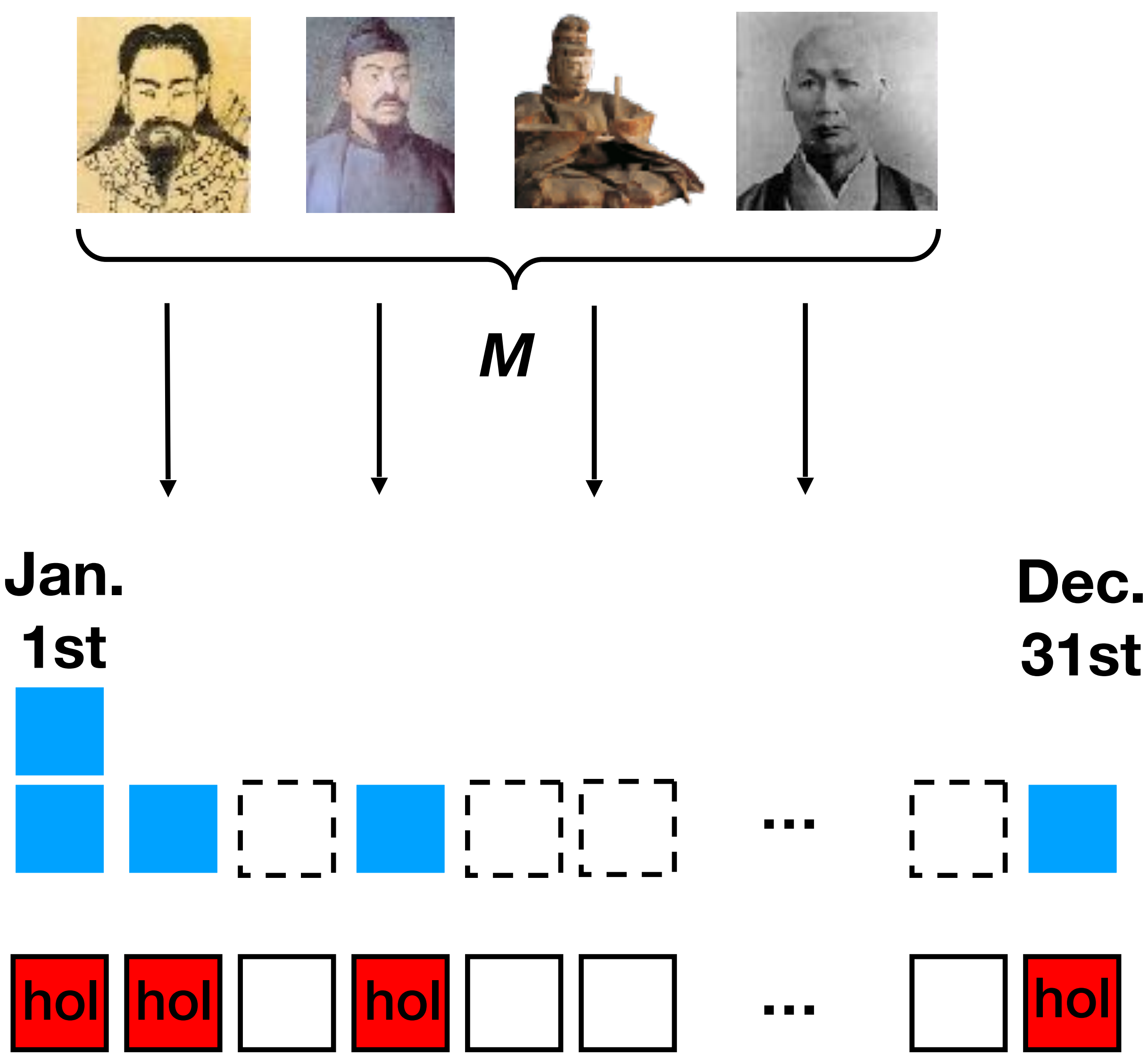


Problem of Partition No.



- It is LARGE!
- Took 100 min for M=90.
- Expect 80 hours for M=125

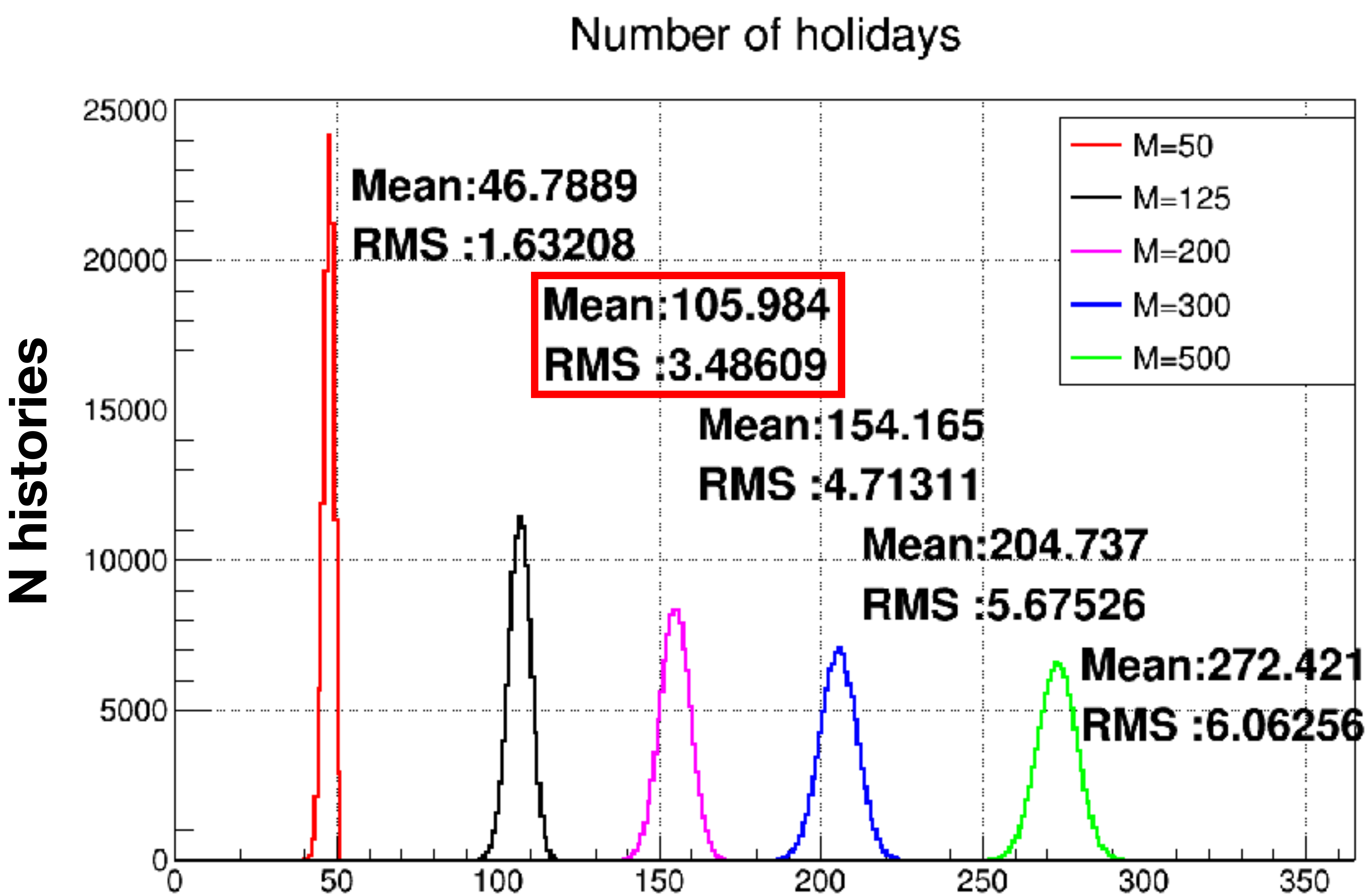
2. Monte Carlo Method



- *Monte Carlo* method can be used to calculate the expected number of holidays
- *M* emperors are uniformly assigned to 365 boxes.

← **Number of holidays is immediately obtained.**

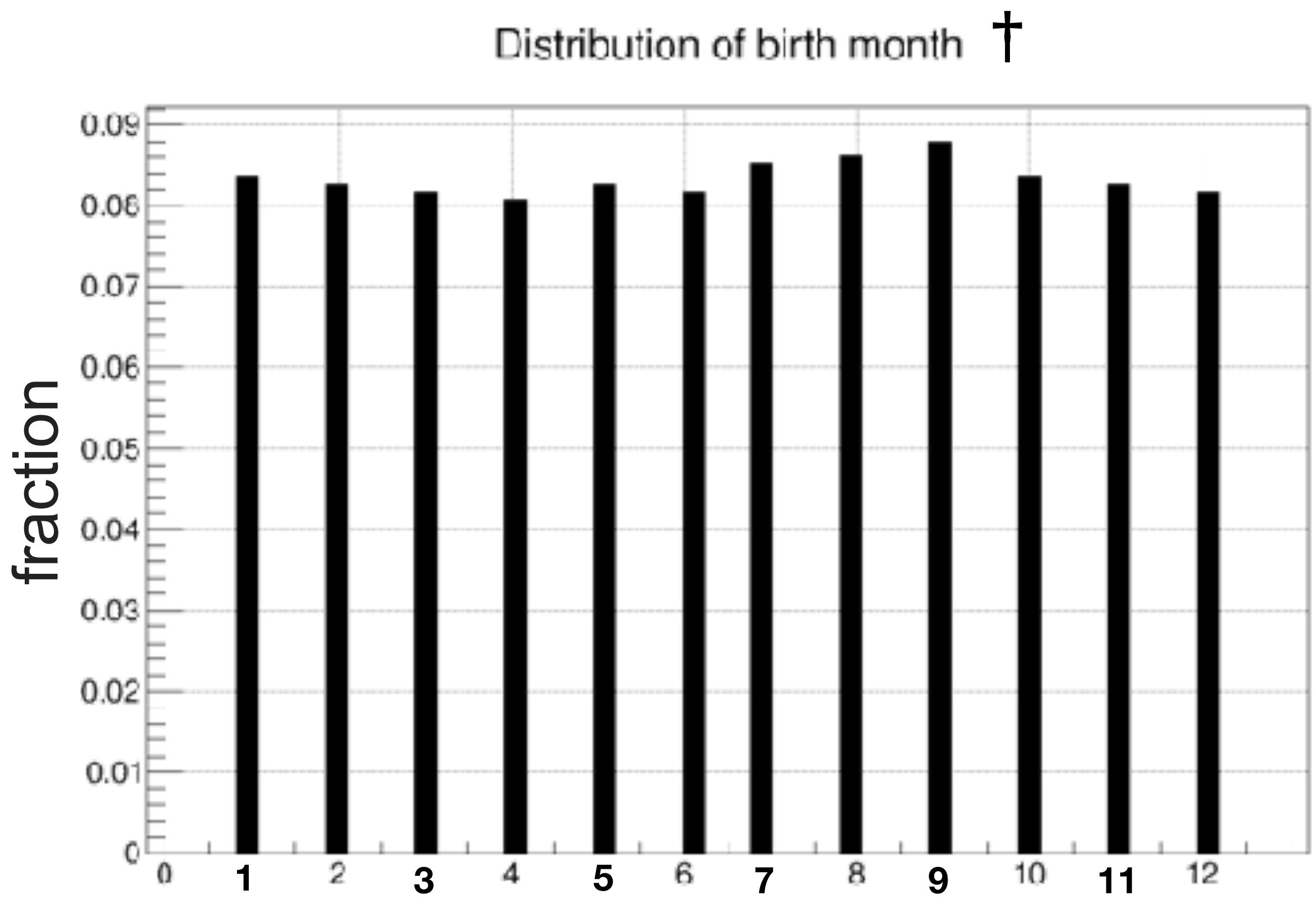
We virtually *created* 10^5 histories of Japan.



Number of holidays
N histories
Mean:46.7889
RMS :1.63208
Mean:105.984
RMS :3.48609
Mean:154.165
RMS :4.71311
Mean:204.737
RMS :5.67526
Mean:272.421
RMS :6.06256
M=50
M=125
M=200
M=300
M=500
25000
20000
15000
10000
5000
0
0
50
100
150
200
250
300
350

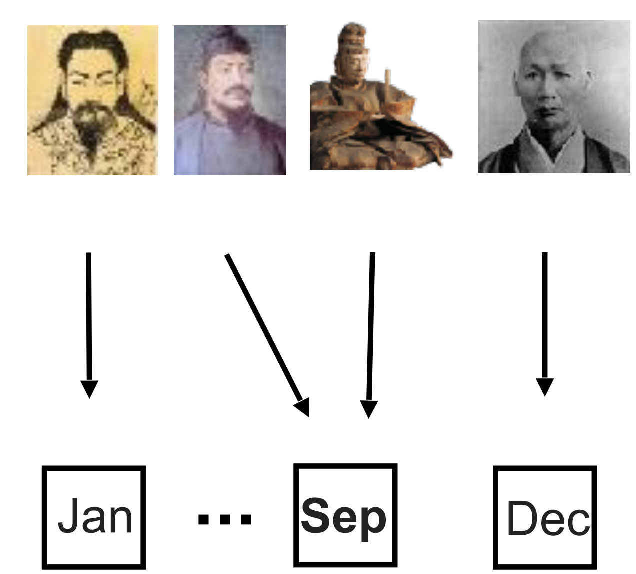
More realistic distribution

Japanese government provides the statistic information of fraction of birth-month.

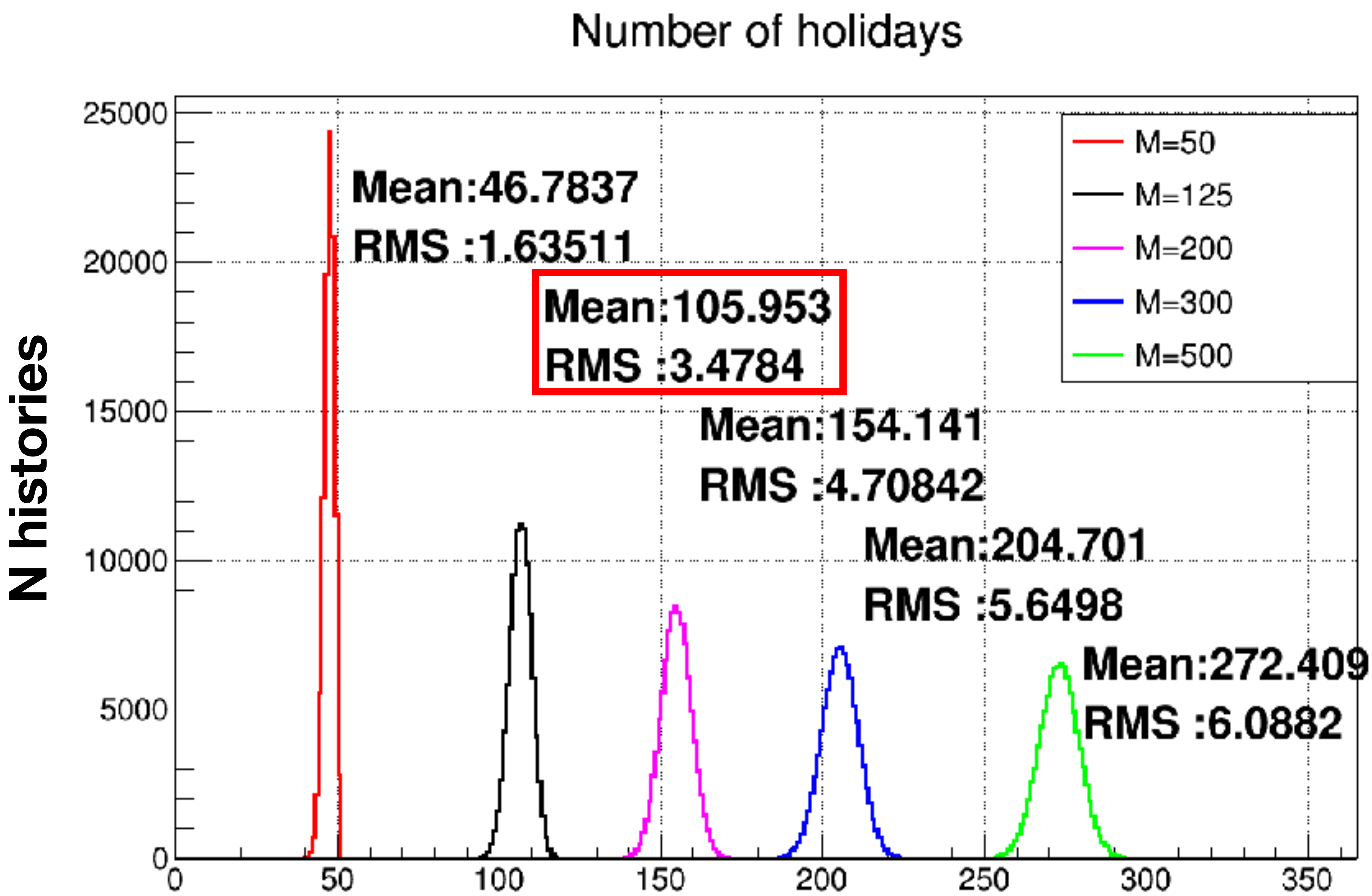


The real distribution is NOT uniform at all!

The assignment of birth days is weighted according to this fraction.



† http://www.mhlw.go.jp/toukei/saikin/hw/jinkou/tokusyu/syussyo-4/syussyo1-2.html



Expected number of holidays decreases by only 0.03 days (43 minutes).

→ Relax! Still we can take sufficient number of holidays even in this realistic situation.

3. Recurrence Formula

- **Q(M | n)** : The number of cases for distributing *M* emperors to *n* specific days

all in 1 day **all in 2 days**

$$Q(M|n) = n^M - {}_n\mathrm{C}_1\ Q(M|1) - {}_n\mathrm{C}_2\ Q(M|2) \cdots$$

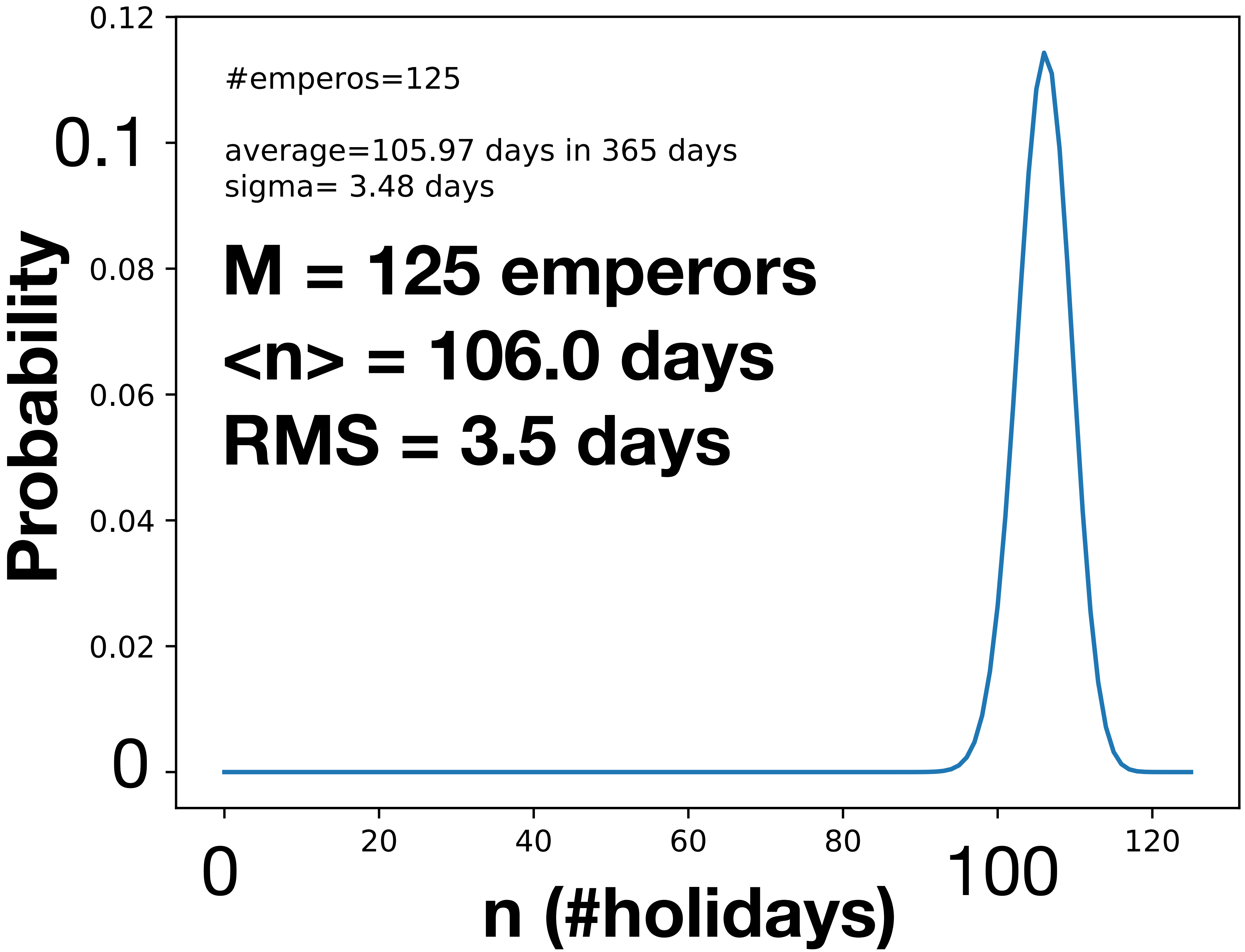
$$= n^M - \sum_{i=1}^{\min(M,n-1)} {}_n\mathrm{C}_i\ Q(M|i)$$

- Probability that *M* emperors fit in *n* days out of *N* days in a year

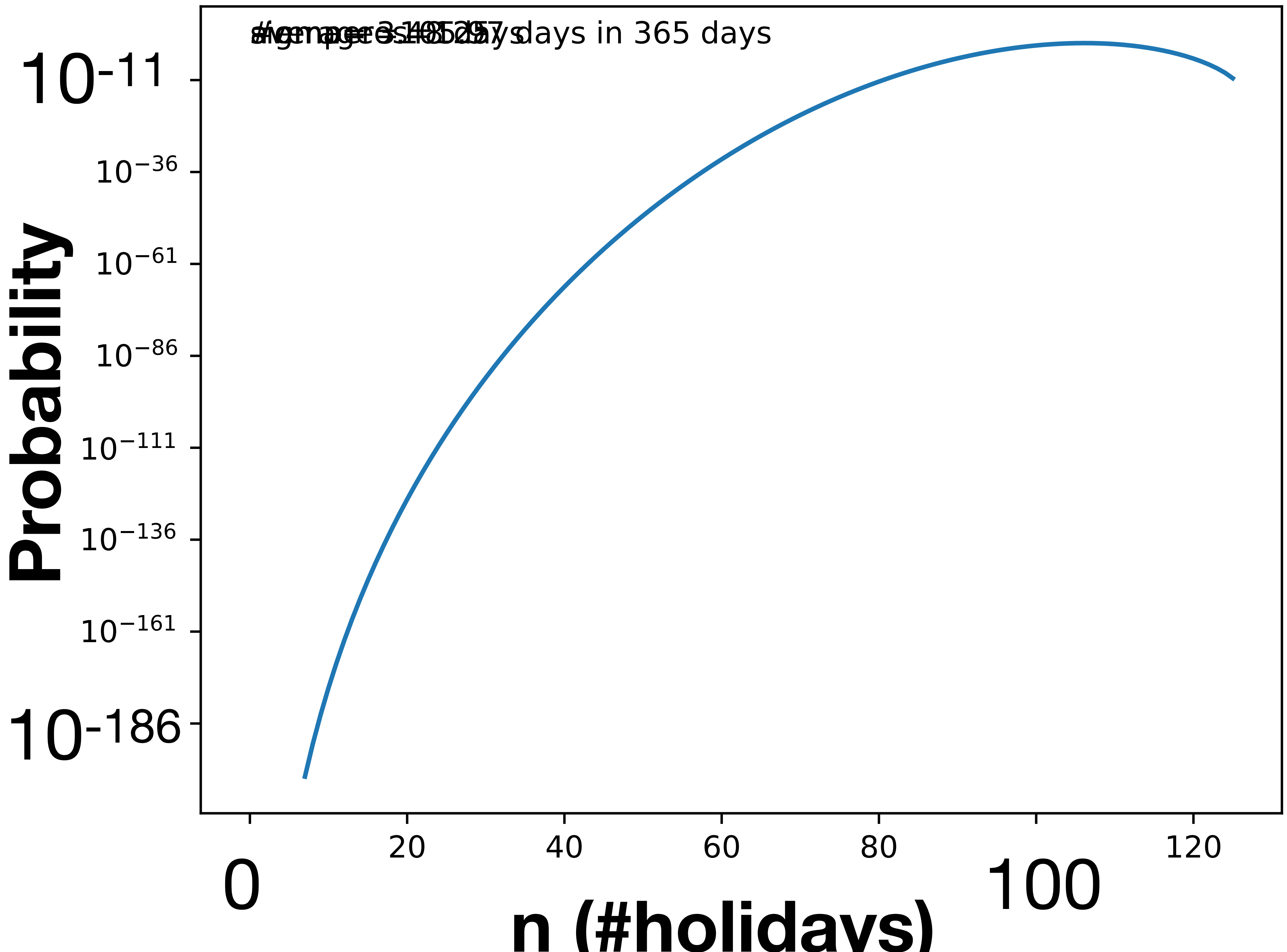
$$P(M,N|n) \quad = \quad {}_N\mathrm{C}_n\ \frac{Q(M|n)}{N^M}$$

- Used Python (infinite #digits for integers!)

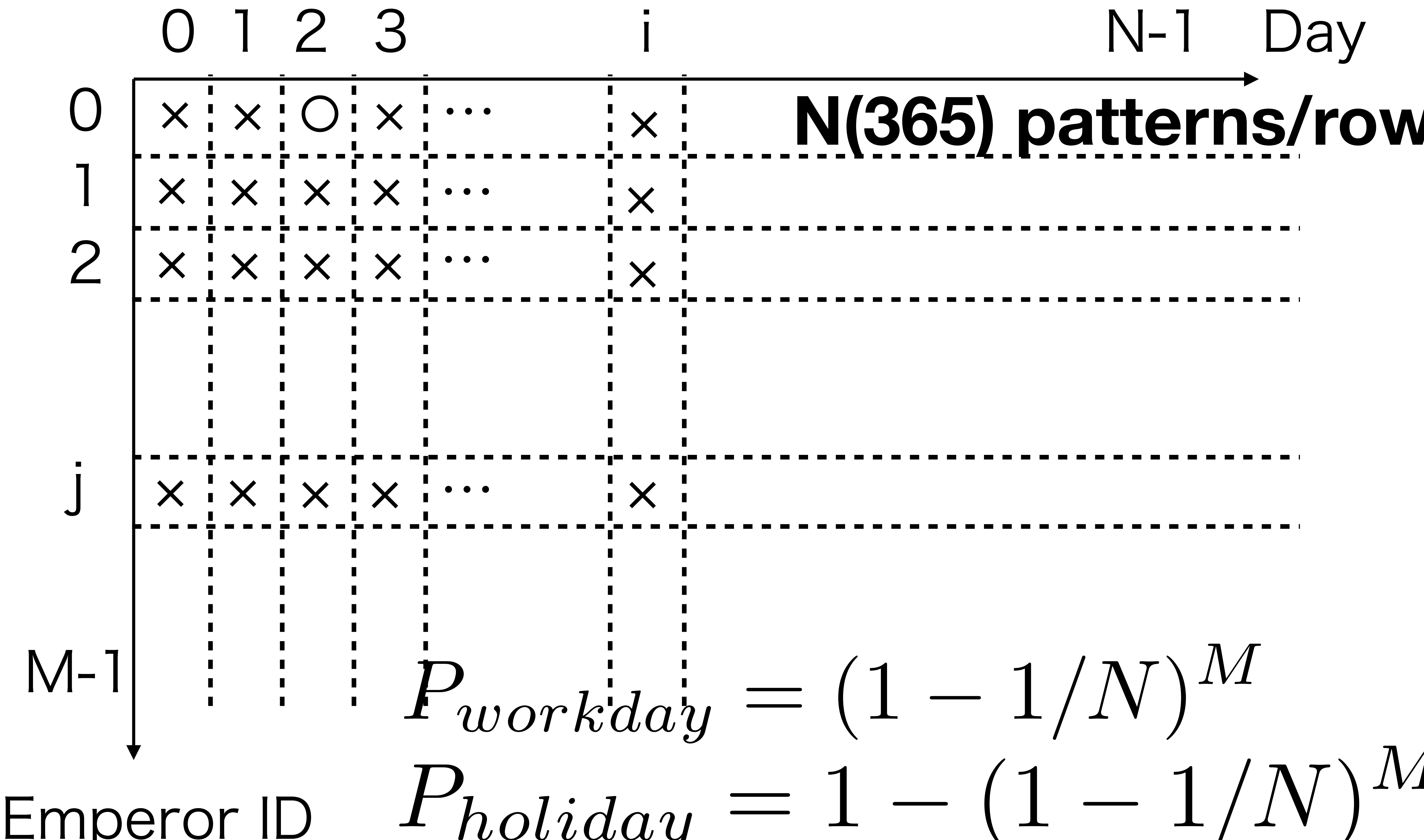
Result of recurrence formula



Result of recurrence formula



4. Simple calculation



$$P_{workday} = (1 - 1/N)^M$$

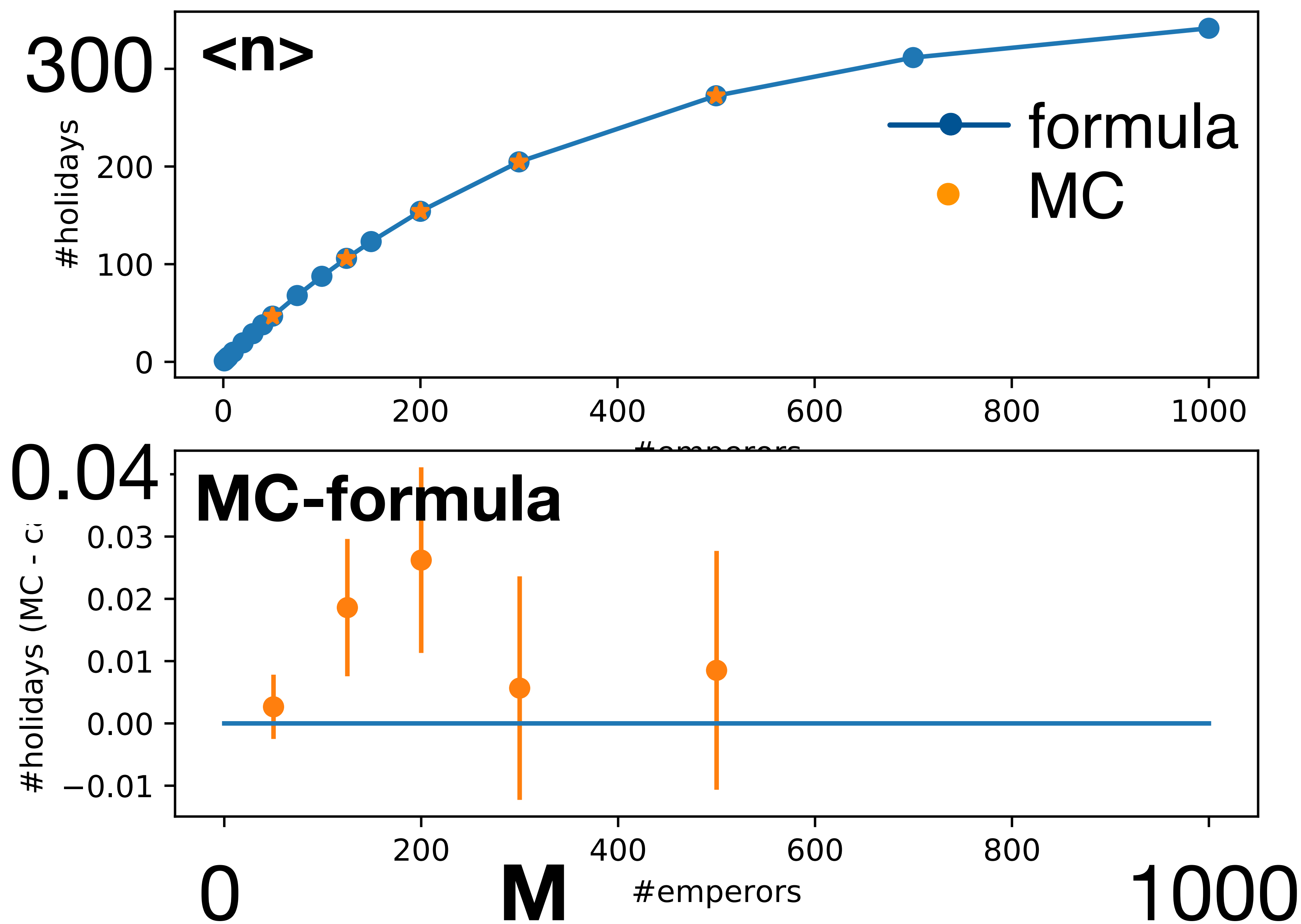
$$P_{holiday} = 1 - (1 - 1/N)^M$$

Result of simple calculation

- Binomial distribution : $N, P_{holiday} \rightarrow P(n)$
- Mean = $N \times P_{holiday} = 365 \times (1-(1-1/365)^{125}) = 106$ days
- RMS = $\sqrt{NP_{holiday}(1-P_{holiday})} = 8.7$ days

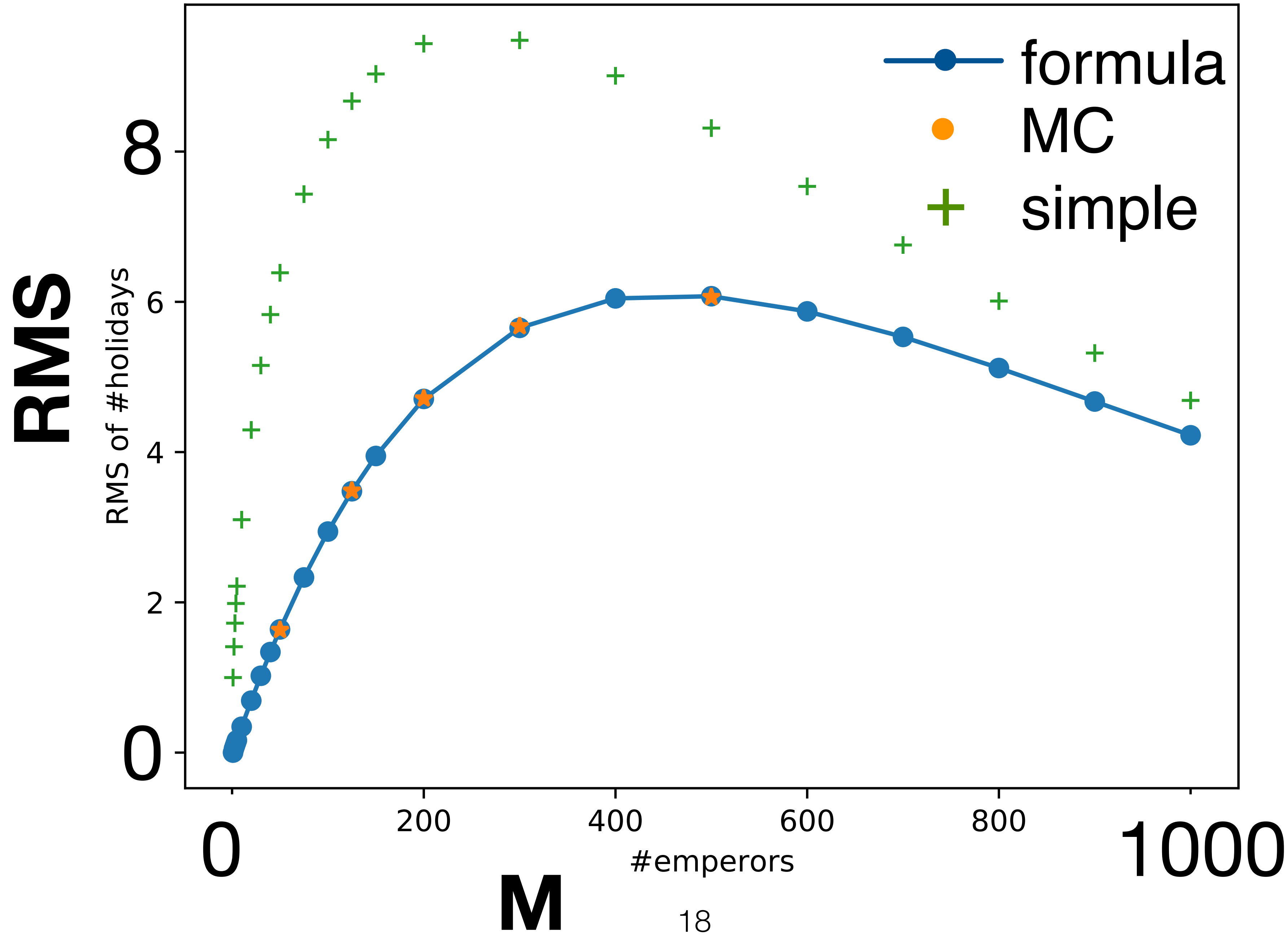
<n> Comparisons

- Recurrence formula and the simple formula agree up to 12 digits. But why $\sum_n n \cdot {}_N\mathrm{C}_n Q(M|n)/N^M = N\left[1-(1-1/N)^M\right]$?
- MC is consistent with the recurrence formula



RMS Comparisons

- RMS of the simple formula does not agree



Conclusion

- #holidays = $N\left[1-(1-1/N)^M\right]$

  106.0 ± 3.5 days (for 125 emperors)

- Probability distribution function =

$$
\begin{aligned}
P(M,N|n) &= {}_N\mathrm{C}_n \, \frac{Q(M|n)}{N^M} \\
Q(M|n) &= n^M - \sum_{i=1}^{\min(M,n-1)} {}_n\mathrm{C}_i \; Q(M|i)
\end{aligned}
$$

- RMS needs the recurrence formula or MC

Conclusion

- Depending on how you approach/think, the problem can be
  - extremely difficult and time consuming, or
  - extremely simple and quick
- Monte Carlo is easy and robust
- Writeup 「全ての天皇の誕生日を祝日にすると何日休みになるか」is available on the program page

Happy Holidays!

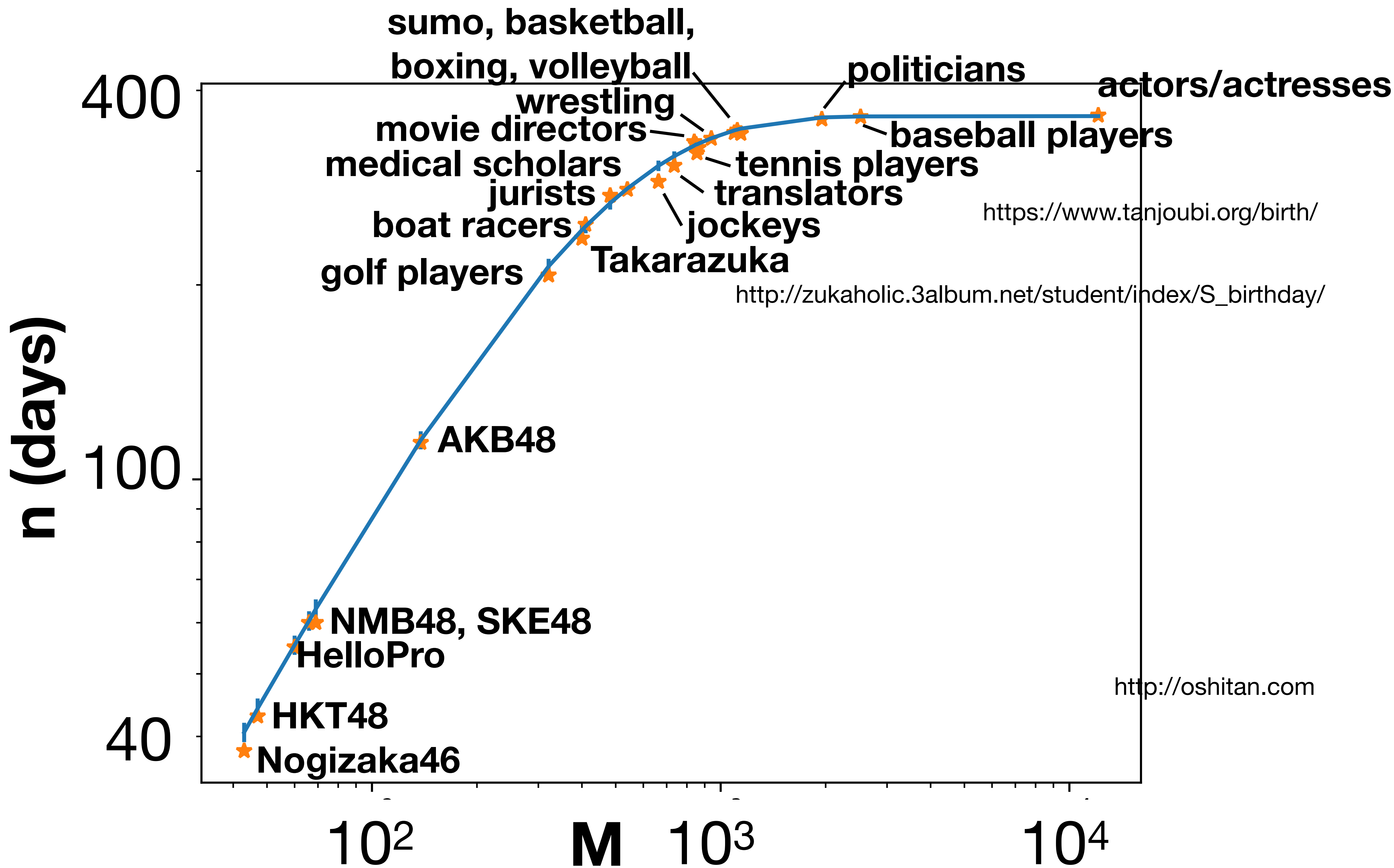
Backup

Check with data

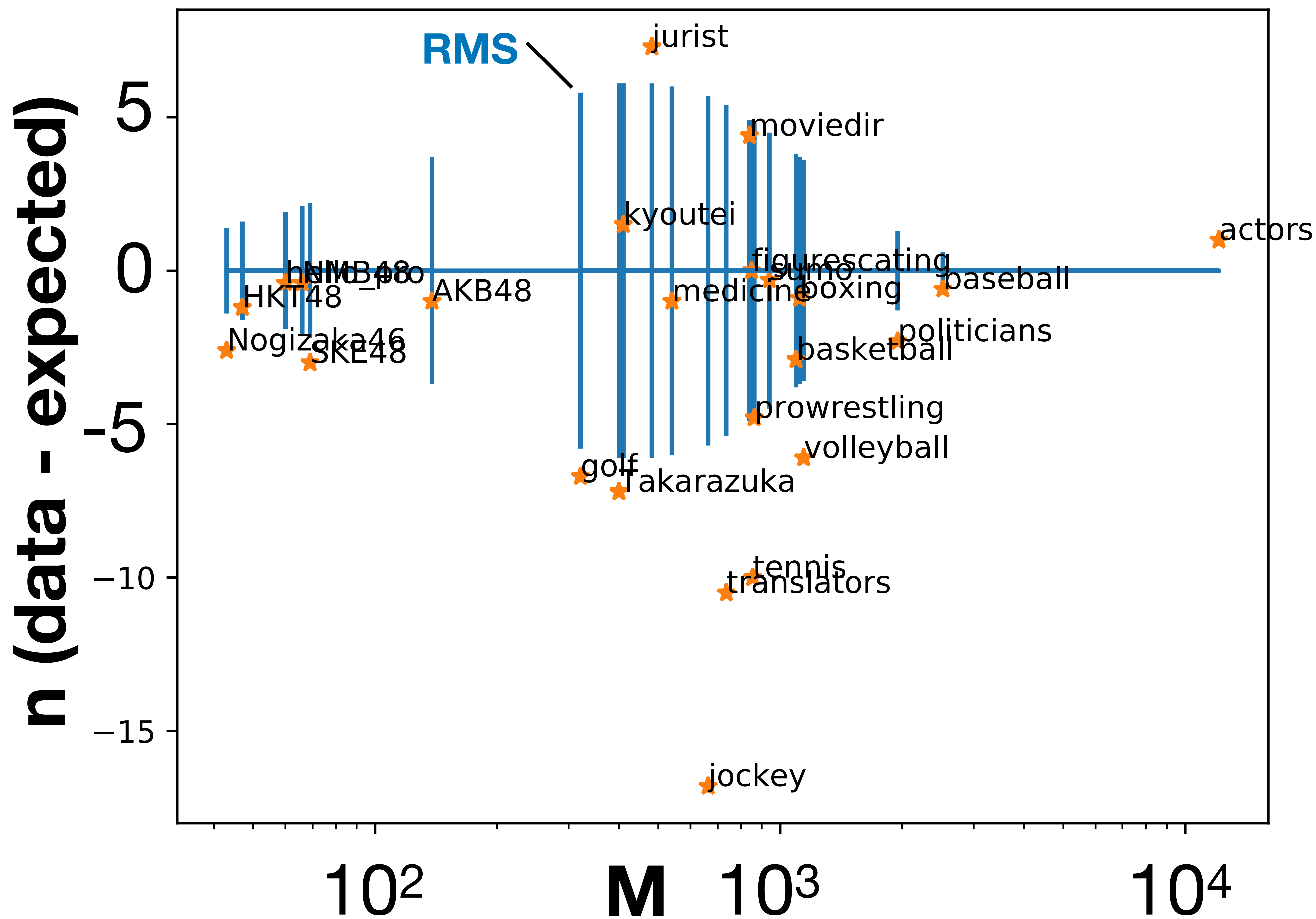
https://www.akb48.co.jp/about/members/

- M = 138 members http://oshitan.com/akb48/
- n = 114 different birthdays
- expected n = 115.0 ± 3.7

Check with data



Differences between data and calculations



5. Discussion

- Expected number of holidays

$$\begin{aligned}\langle n \rangle &= N\left[1 - \left(1 - \frac{1}{N}\right)^{M}\right] \\ &= N\left[1 - \left(1 - \frac{1}{N}\right)^{N\frac{M}{N}}\right] \\ &\simeq N\left[1 - \exp\left(-\frac{M}{N}\right)\right]\end{aligned}$$

(if N>>1)

- Why exponential??

Actually, …

- **For 1 specific day**, the expected number of emperors whose birthday is that day

$$\mu = M/N (= 125/365)$$

- Probability that some emperor's birthday is **on that day**, with Poisson distribution

$$\begin{aligned} 1 - P(\mu, 0) &= 1 - \frac{e^{-\mu}\mu^0}{0!} \\ &= 1 - e^{-\mu} \end{aligned}$$

- Expected number of birthdays **in a year with *N* days**

$$N\left(1 - e^{-M/N}\right)$$

How to count birthdays

- Example: Download https://www.tanjoubi.org/birth/bunrui_076.html to 076_1.txt, 076_2,txt, …

・1902年01月22日 田畑忍 (たばた しのぶ)
・1902年08月21日 石本雅男 (いしもと まさお) <img src="../images/
death.gif" title="この方は亡くなられています。"/>

- cat 076_*.txt | tr '・' '\n' (split to lines)

1902年01月22日 田畑忍 (たばた しのぶ)
1902年08月21日 石本雅男 (いしもと まさお)

- cat 076_*.txt | tr '・' '\n' | grep 14px | sed 's/\<a href.*//' | sed ’s/.*\>//' | grep 日 | sed 's/.*年//'

01月22日
08月21日

- cat 076_*.txt | tr ' ・ ' '\n' | grep 14px | sed 's/\<a href.*//' | sed ’s/.*\>//' | grep 日 | sed ’s/.*年//' | **wc -l**

  Count the number of lines
  482

- cat 076_*.txt | tr ' ・ ' '\n' | grep 14px | sed 's/\<a href.*//' | sed ’s/.*\>//' | grep 日 | sed ’s/.*年//' | **sort -u | wc -l**

  Count the number of lines after removing duplicate lines
  275