全ての天皇の誕生日を祝日にすると何日休みになるか v1.0

清水 信宏、南條 創、山中 卓
大阪大学大学院理学研究科 物理学専攻

2017 年 12 月 28 日

1 はじめに

日本には、歴代 125 人の天皇がいるとされているが、そのうちその誕生日が祝日となっているのは現在、明治天皇 (11 月 3 日 文化の日)、昭和天皇 (4 月 29 日 昭和の日)、今上天皇 (12 月 23 日 天皇誕生日) の 3 日間だけである。そもそも歴代 125 人全員の誕生日を祝日にすれば 1 年のうち何日が休みになるのか、という疑問が昼食の会話でもち上がった。誕生日が重なる場合もあるので、125 日よりは少なくなるはずである[1]。そこで、様々な方法でこの「天皇誕生日」の数の期待値と確率分布を求めてみた。

以下、$N = 365$ を 1 年の日数、$M = 125$ を歴代の天皇の人数とする。また、$P(M, N|n)$ を n 日が天皇誕生日で祝日となる確率とする。

2 天皇をグループ分けするゴリ押し法

ここでは、天皇を何日かにグループ分けする方法を考える。天皇が 7 人の場合を例にとって説明する。この場合、誕生日の分け方には次の場合がある。

- (1)：7 人が 1 日に集まっている
- (1+6)：1 人が 1 日、6 人が別の 1 日に集まっている
- (2+5)：2 人が 1 日、5 人が別の 1 日に集まっている
- (1+2+2+2)：1 人が 1 日、2 人が別の 1 日、2 人がまた別の 1 日、2 人がさらに別の 1 日に集まっている
- $\cdots$

など、15 種類。この分け方の種類の数は「分割数」と呼ばれている。

例えば、(1+2+2+2) の分け方の場合、祝日は 4 日あり、天皇の分け方が何通りあるかは次の値の積を求めればよい。

[1] これは、山中が高校の時に計算しようとしてできなかった問題でもある。

- ${}_N\mathrm{C}_1$ ：1 人の入る 1 日を選ぶ
- ${}_7\mathrm{C}_1$ ：その 1 日に入れる 1 人を選ぶ
- ${}_{N-1}\mathrm{C}_3$ ：残りの日から (2+2+2) の 3 日を選ぶ
- ${}_6\mathrm{C}_{2\times 3}$ ：残りの 6 人から 2×3 人を選ぶ
- $(2\times 3)!/(2!)^3$：6 人を 1 列に並べ、各祝日内の並べ方の数で割る

これらの積は 1,833,153,121,800 となる。図 1 のように祝日の日数が 1 日から 7 日までの表を作っておき、7 人の分割方法全てに対して、上のような計算を行ってその積を表に足していけば、各祝日の日数に対して、何通りの天皇の分け方があるかが決まる。それを N^M (この例の場合 365^7) で割れば、確率分布が求まる。

#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

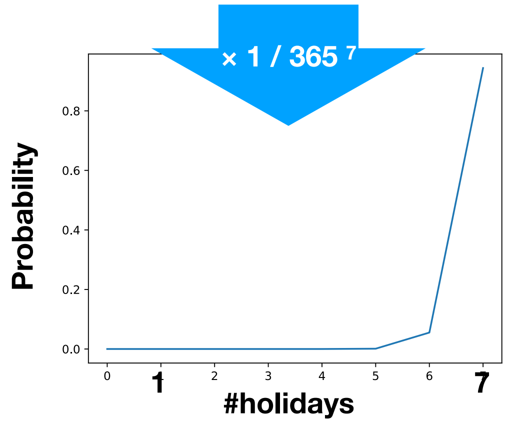

図 1: 天皇が 7 人の場合の例。7 の分割方法全てについて、場合の数を足し上げたのち、365^7 で割れば、確率分布となる。

この方法の問題は、図 2 に示すように、分割数が天皇の人数に対して指数関数的に増えることである。外挿すると $M=125$ の分割数は 10^9 台である。$M=90$ の場合でも分割数は 5.7×10^7 あり、計算に約 100 分かかった。$M=125$ の場合は約 80 時間かかると予測する。

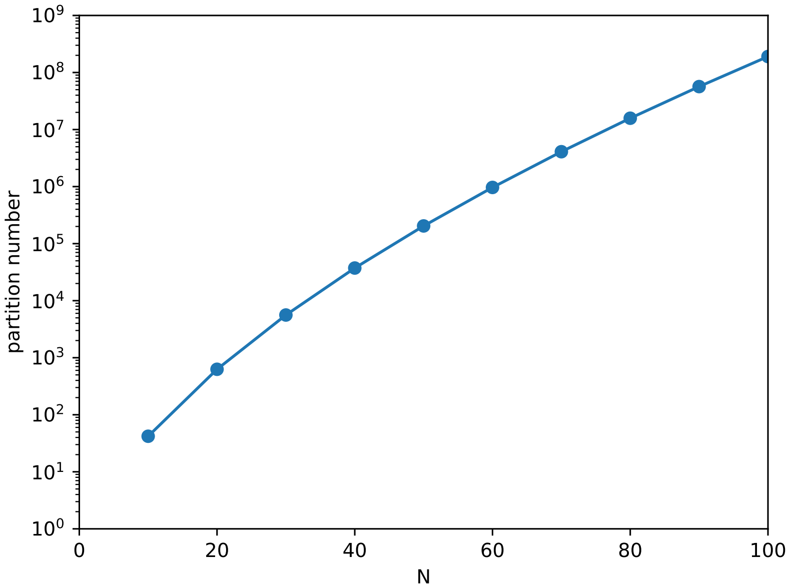

図 2: N の分割数 (partition number)

3　モンテカルロ法を用いた方法

モンテカルロ (Monte Carlo:MC) 法を用いて祝日の日数を計算する方法を考える。各天皇を一様分布に従って誕生させ，それを M 人分繰り返すことで仮想的に日本の歴史を一回生成する。図 3 は，10^5 回日本の歴史を生成したシミュレーション結果を示す。黒 ($M=125$) が現在の日本の歴代天皇の数に相当する。この計算によれば，一年あたり，平均 106 回祝日が制定されるということになる。驚くべきことに，その二乗平均 (Root Mean Square: RMS) は約 3.5 日と 106 日に比べて非常に小さく，安定して一年の三分の一休みが取れることが読み取れる。

ちなみに，厚生労働省は誕生月の割合の統計値を公表している [1]。その報告を、整理したものを表 1 に示す。表からわかるように，九月に誕生している子供が最も多く，平成 12 年の統計では平均からの乖離は 2.2 RMS にもおよぶ。畏れ多くも，一般国民の誕生月の分布を用いて天皇家のそれを近似することにした。上で述べた一様分布を，割合に応じて重み付けした分布に変えるだけでよい。この資料には，日付ごとの割合は示されていないため，ある月の中の 30 日 (ないし 31 日) に対しては一様に分布させることとする。図 4 にその結果を示す。一様分布では 105.98 ± 0.01 日であった祝日の期待値が，105.95 ± 0.01 日へと変化している。現実の誕生月には偏りがあるため，重複する日数が増すことによって祝日の日数が減っていることを示している。それでもなお，その変化の大きさは軽微であり、誕生月の分布が極端に変化しない限り安定して休みが確保されるので安心である。

以上で述べたように，モンテカルロ法を利用すれば，期待される日数を簡単に評価することが可能である。特に，現実の分布を考慮して祝日の数の期待値を求めることは，以下に示す解析的手法では難しく，モンテカルロ法の単純さ，そしてそれ故の強力さを改めて

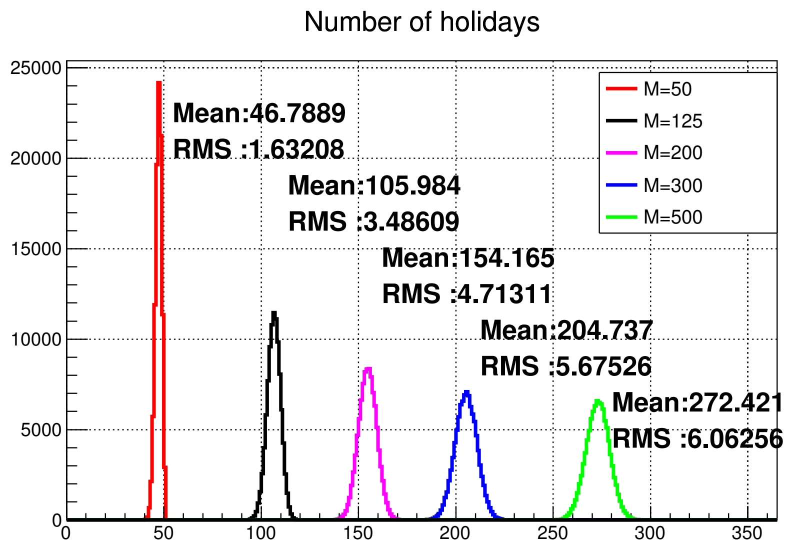

図 3: モンテカルロ法を用いてシミュレーションした祝日の数の分布。赤 ($M = 50$), 黒 ($M = 125$)，マゼンタ ($M = 200$), 青 ($M = 300$), 緑 ($M = 500$)

表 1: 月ごとの誕生割合 (平成 12 年厚生労働省統計) [1]

月	割合 %	平均からの乖離 (RMS)
1	8.35	0.09
2	8.26	-0.36
3	8.17	-0.81
4	8.08	-1.26
5	8.26	-0.36
6	8.17	-0.81
7	8.52	0.94
8	8.61	1.39
9	8.79	2.29
10	8.35	0.09
11	8.26	-0.36
12	8.17	-0.81
平均	8.33 ± 0.20 (RMS)	

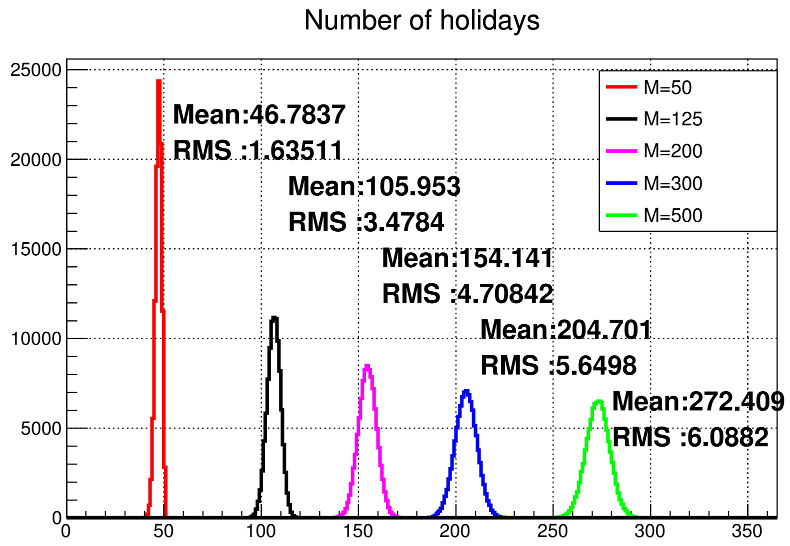

図 4: 実際の誕生月の割合を考慮して，シミュレーションした祝日の数の分布。赤 ($M = 50$)，黒 ($M = 125$)，マゼンタ ($M = 200$)，青 ($M = 300$)，緑 ($M = 500$)

認識できる。

4 漸化式を用いた方法

4.1 漸化式の導出

次に、漸化式を用いた方法を考える。$Q(M|n)$ を、ある特定の n 日に M 人を分配する場合の数とする。M 人それぞれが n 通り選ぶ選び方は n^M 通りあるが、これは M 人全員が 1 日に集中してしまう場合や、2 日間に集中してしまう場合などを含む。したがってそれらの寄与を差し引くと、

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

となる。ただし、差し引く日数の最大値は $n-1$ と M のうち小さい方であるため、

$$Q(M|n) \quad = \quad n^M - \sum_{i=1}^{\min(M,n-1)} {}_n\mathrm{C}_i\ \ Q(M|i) \tag{2}$$

と表せる。

この $Q(M|n)$ を用いると、1 年 N 日のうち、n 日に M 人が入る確率は

$$P(M,N|n) \quad = \quad {}_N\mathrm{C}_n \frac{Q(M|n)}{N^M} \tag{3}$$

である。したがって、漸化式の式 (2) を用いて $Q(M|i)$ を $i = 1$ から M まで順に求めておけば、祝日の日数の確率分布 $P(M,N|n)$ を計算できる。

4.2　漸化式の計算

$Q(M|i)$ は整数であるので正確に計算できるが、$N^M = 365^{125} \simeq 10^{320}$ という大きな数を扱う必要がある。そこで、整数の桁数に制限のない Python を用いて計算した。$P(M,N|n)$ の計算では実数で割り算するために、割る数と割られる数 (整数) を、Python で表せる実数の範囲 $O(10^{308})$ に収まるまで 10^{100}(整数) で約分した。

図 5 に、漸化式で求めた天皇誕生日の日数の確率分布 $P(125,365|n)$ を示す。期待値は 106.0 日であり、分布の RMS は 3.5 日である。この結果は、3 節で述べたモンテカルロ法の結果と一致する。

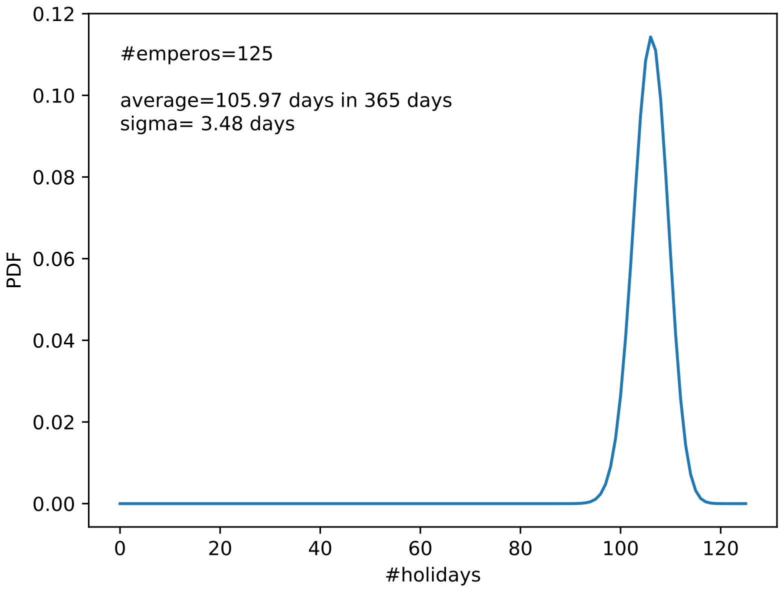

図 5: 漸化式を用いた、天皇誕生日の日数の確率分布。

5　初等的な手法

次に、初等的な方法を考える。とある 1 日が祝日となる確率 $P_{holiday}$ が解れば、祝日数の期待値は $N \times P_{holiday}$ となる。とある 1 日が祝日となる確率 $P_{holiday}$ を求めればよい。

5.1　誕生日表

状況を整理する上で、以下のような M 人 $\times N$ 日の表に、誕生日 (✓) か否か (✗) を埋めていくと考えると分かり易い。

この誕生日表の作りかたは以下である。

1. とあるセルについて、✓の確率は $\alpha \equiv 1/N$、✗の確率は $1-\alpha$.

表 2: Table of birthdays. ✓: a brithday, ✗: not a birthday.

		Day			
		0	1	⋯	N-1
Person	0	✗	✓	⋯	✗
	1	✗	✗	⋯	✗
	⋯				
	M-1	✗	✗	⋯	✗

2. 各列について、✓が必ず 1 個のみあること。(誕生日は必ず 1 回。)

次に平日、祝日については以下の条件となる。

- 日付の列で 1 個でも ✓があれば祝日。$\rightarrow P_{holiday}$
- 日付の列で全て ✗である場合平日。$\rightarrow P_{workday} \equiv 1 - P_{holiday}$

5.2 方針

$M \times N$ の全セルを ✓か ✗で埋めると全確率は、

$$[\alpha + (1-\alpha)]^{M+N} = 1 \tag{4}$$

一方で、このなかには人の行あたり、✓(誕生日) が 0 であったり、2 回以上のケースも含む。これを除いた全パターンを分母とし、とある 1 日が平日となるケースを分子として、$P_{workday}$ を求め、$P_{holiday} = P_{workday} - 1$ を求めるという方針をとる。

5.3 分母

各行毎に必ず 1 個 ✓それ以外 ✗より、

$$\text{各行} \rightarrow {}_N\mathrm{C}_1\ \alpha^1(1-\alpha)^{N-1} \tag{5}$$

$$\text{M 行} \rightarrow \left[{}_N\mathrm{C}_1\ \alpha^1(1-\alpha)^{N-1}\right]^M \tag{6}$$

5.4 分子

とある 1 日は ✗で、それ以外のどこか 1 日は ✓、他 ✗

$$\text{各行} \rightarrow (1-\alpha) \times {}_{N-1}\mathrm{C}_1\ \alpha^1(1-\alpha)^{N-2} \tag{7}$$

$$\text{M 行} \rightarrow \left[(1-\alpha) \times {}_{N-1}\mathrm{C}_1\ \alpha^1(1-\alpha)^{N-2}\right]^M \tag{8}$$

5.5　ある 1 日が休日の確率

$$
\begin{aligned}
P_{workday} =& \frac{\left[(1-\alpha)\times {}_{N-1}\mathrm{C}_1\ \alpha^1(1-\alpha)^{N-2}\right]^M}{\left[{}_N\mathrm{C}_1\ \alpha^1(1-\alpha)^{N-1}\right]^M} && (9)\\
=& \left[\frac{(1-\alpha)\times(N-1)\alpha^1(1-\alpha)^{N-2}}{N\alpha^1(1-\alpha)^{N-1}}\right]^M && (10)\\
=& \left[1-1/N\right]^M && (11)\\
=& (1-\alpha)^M && (12)\\
P_{holiday} =& 1-P_{workday} = 1-(1-\alpha)^M = 1-(1-1/N)^M && (13)
\end{aligned}
$$

ところで、$P_{workday} = (1-\alpha)^M$ の式は誕生日が 0 や 2 回以上などのあってはならない場合を除くことを考えず、M 人分誕生日でない確率を掛け合せているのと完全に一致している。分母分子で、あってはならない場合を除くファクターがキャンセルしているのであろうが、これが自明であるかはわからない。

5.6　祝日数の期待値

$M = 125, N = 365$ とすると、

$$
P_{holiday} = 1-(1-1/365)^{125} = 0.29 \qquad (14)
$$

祝日数の期待値 μ は

$$
\mu = N\times P_{holiday} = 106(\text{日}). \qquad (15)
$$

5.7　祝日数の分布

とある日が祝日の確率 $P_{holiday}$ で、N 日分の 2 項分布とすると、RMS は、

$$
RMS = \sqrt{N\times P_{holiday}\times(1-P_{holiday})} = 8.7(\text{日}). \qquad (16)
$$

となり、MC 法や漸化式の手法の結果と一致せず、2 項分布の仮定がよくないことを意味する。各日付の $P_{holiday}$ が独立でなく、誕生日が必ず 1 個という条件から、ある日が平日なら、他のどこかが祝日になりやすいというような、相関があるからであろう。

5.8　誕生日表 2

思いかえして、誕生日表の作り方を再考する。

1. 各列について、✓が必ず 1 個のみあること。(誕生日は必ず 1 回。)

すると、各行に対して、✓の場所で N 通り。全パターンは、N^M という型式のほうが簡単。とある 1 日が平日である確率は、

$$
P_{workday} = (1-1/N)^M \qquad (17)
$$

となることはこれでよさそう。

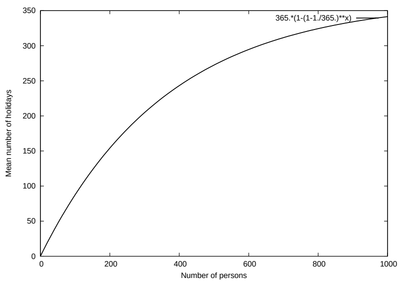

図 6: 2 項分布での祝日数の期待値。

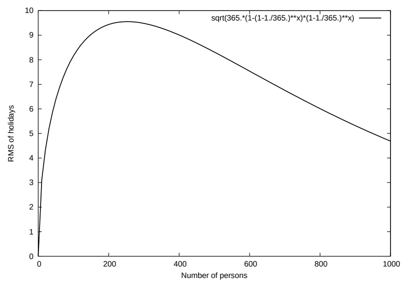

図 7: 2 項分布での祝日数の RMS。

表 3: Table of birthdays. ✓: a brithday.

		Day			
		0	1	$\cdots$	N-1
Person	0		✓	$\cdots$	
	1			$\cdots$	
	$\cdots$				
	M-1			$\cdots$	

6　計算結果の比較

3 節、4 節、5 節で説明したモンテカルロ法、漸化式による方法、初等的な方法の結果を比較した。

まず、Python 2.7.13 (64bit 版) を用い、漸化式による方法と初等的な方法によって求めた祝日の日数の期待値は、$1 \leq M \leq 1000$ の範囲で、少なくとも 12 桁一致した。しかし、なぜ二つの方法による期待値が

$$\sum_n n \cdot {}_N\mathrm{C}_n \; Q(M|n)/N^M \quad = \quad N\left[1-(1-1/N)^M\right] \tag{18}$$

のように一致するのかは不明である。

また、3 節のモンテカルロ法による期待値は図 8 に示すように $1 \leq M \leq 1000$ の範囲でモンテカルロの統計誤差の範囲内で漸化式で求めた方法と一致した。祝日日数の確率分布の RMS は、図 9 に示すように漸化式による方法とモンテカルロ法は一致するが、初等的な方法 (二項分布) による方法は大きくずれ、一致しない。

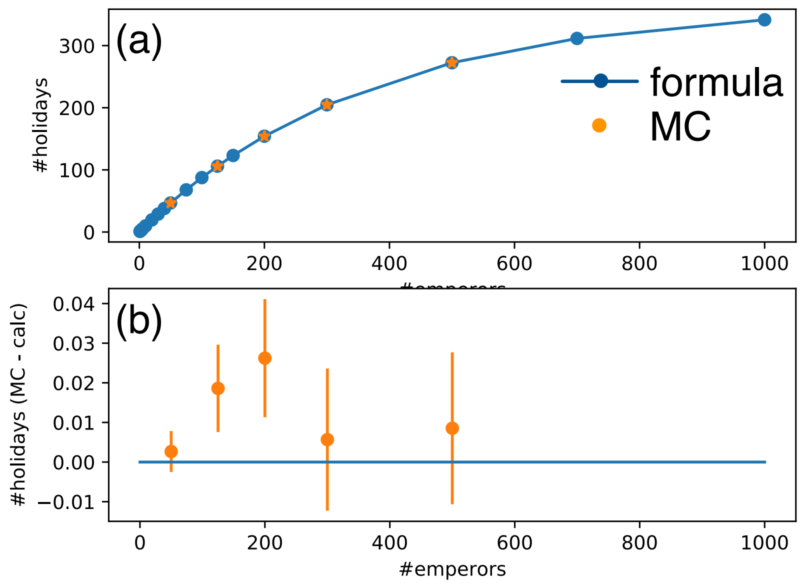

図 8: (a): 祝日日数の期待値の、モンテカルロ法 (オレンジ色) と漸化式 (紺色) による結果と M (天皇の数) の関係。(b): 祝日日数の期待値の、モンテカルロ法と漸化式の結果の差。誤差棒は、モンテカルロ法の統計誤差 (RMS/$\sqrt{10^5}$) を表す。

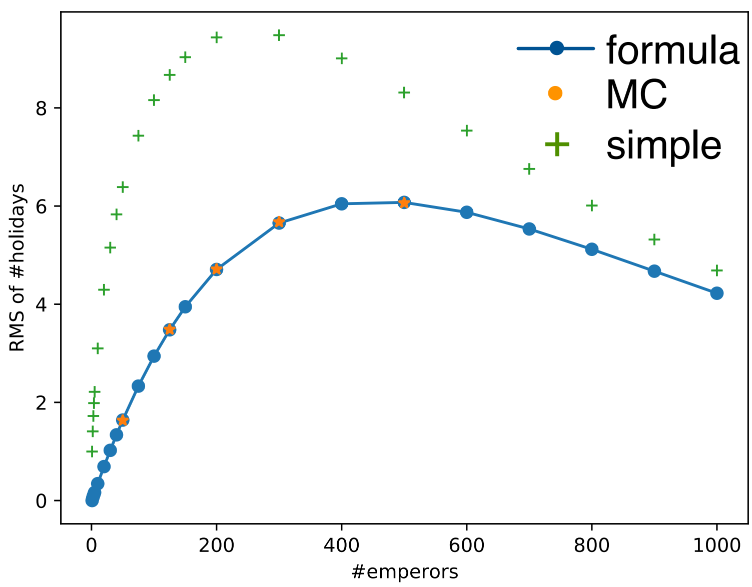

図 9: 祝日日数の確率分布の RMS と M (天皇の数) の関係。モンテカルロ法 (オレンジ色) と漸化式 (紺色) は一致するが、二項分布による結果 (緑) は他より大きい。

7　データとの比較

次に理論的な予測とデータを比較する。歴代 125 人全ての天皇の誕生日はわからないため [2]、誕生日のわかっている他のサンプルで予測を検証した。

例えば誕生日の公表されている [3, 4] 歴代の AKB48 のメンバー 138 人の誕生日は 114 日ある。式 (3) の漸化式を用いた期待値は 115.0 日で、その RMS は 3.7 日であり、データは計算の予測と一致する。

その他にも、NMB48、SKE48、HKT48 などのアイドルグループ [4]、宝塚歌劇団 [5]、俳優、政治家、野球選手、映画監督、競馬の騎手など様々な職種の人 [6] について、異なる誕生日の数を調べた。その結果を図 10 に示す。また、図 11 に測定値から漸化式による期待値を引いた値を示す。乃木坂 46 の 43 人から俳優 12,102 人の幅広い母集団の大きさに対して、漸化式で求めた値はよく一致している。ただし、俳優 12,102 人の誕生日の数は 366 であり、予測の 365.0 ± 0.0 より多い。これは、予測の計算がうるう年を考慮していないためである。

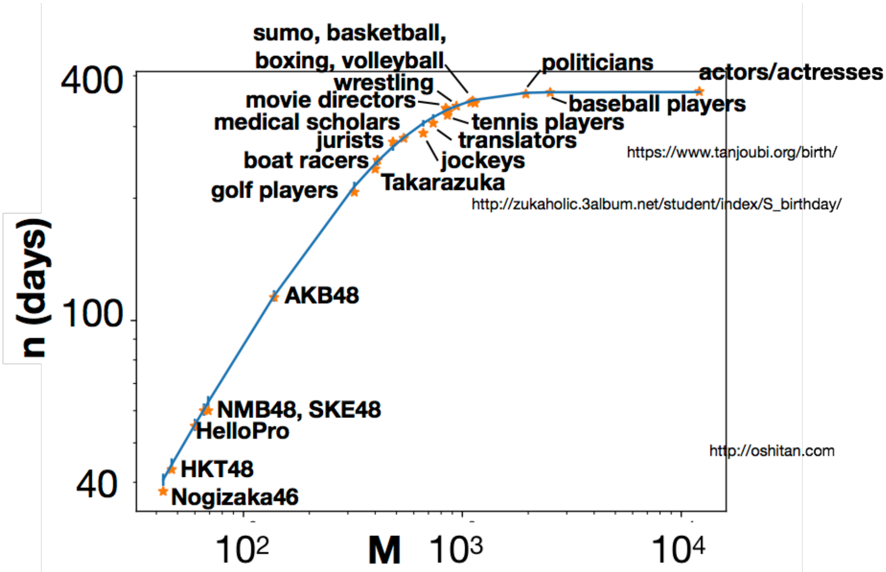

図 10: 各種のグループに対して、星印はその誕生日の数、直線は理論的な期待値。

8　結論

歴代の天皇 125 人全員の誕生日を祝日とすると、期待される祝日の日数は 106.0 ± 3.5 日である。また、祝日の日数の期待値は

$$\langle n \rangle \quad = \quad N\left[1-\left(1-\frac{1}{N}\right)^{M}\right] \tag{19}$$

(20)

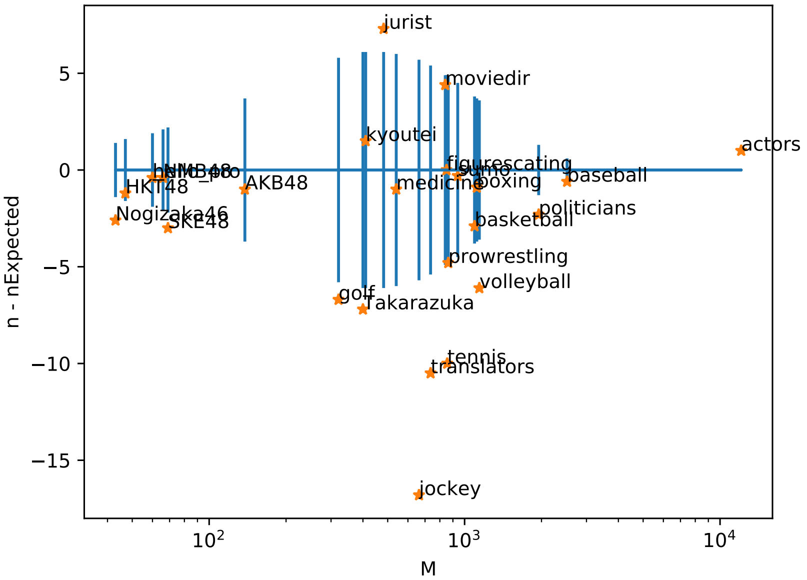

図 11: 各種のグループに対して、その誕生日の数から理論値を引いた値。誤差棒は漸化式を用いて求めた確率分布の標準偏差を表す。

であり、これは漸化式で求めた

$$\langle n \rangle \quad = \quad \sum_n n \cdot {}_N\mathrm{C}_n \; Q(M|n)/N^M \tag{21}$$

と一致する。しかし、これらの二つの式がなぜ等価なのかはまだ証明していない。

確率分布の幅 (RMS) については、初等的な方法では天皇の総数が M であることを保証できないために二項分布では求まらず、漸化式やモンテカルロ法が必要である。

この問題は、考え方によって、計算がいかに困難にも簡単にもなり得るかを示すいい例でもある。また、モンテカルロ法は推論の過ちに陥ることがないため、信頼度が高く、理論計算の検証にも使える。半導体光検出器の MPPC の非線形性も、本質的には同じ問題である。

参考文献

[1] 「出生に関する統計」の概況、人口動態統計特殊報告、厚生労働省 [accessed 2017 Dec. 26]. `http://www.mhlw.go.jp/toukei/saikin/hw/jinkou/tokusyu/syussyo-4/syussyo1-2.html`.

[2] 天皇の一覧、Wikipedia [accessed 2017 Dec. 26]. `https://ja.wikipedia.org/wiki/天皇の一覧`.

[3] AKB 公式サイト、メンバー [accessed 2017 Dec. 26]. `https://www.akb48.co.jp`.

[4] 推しメン誕生日相性チェッカー☆彡 [accessed 2017 Dec. 26]. http://oshitan.com.

[5] づかホリック [accessed 2017 Dec. 26]. http://zukaholic.3album.net/student/index/S_birthday/.

[6] 誕生日 (生年月日) [accessed 2017 Dec. 26]. https://www.tanjoubi.org/birth/.