

研究員 の眼

誕生日が一致する確率 — 多くの人が集まる場では、誕生日の 話題で盛り上がりませんか —

取締役 保険研究部 研究理事

年金総合リサーチセンター長

TEL: (03)3512-1777

中村 亮一

E-mail: nryoichi@nli-research.co.jp

はじめに

以前の研究員の眼「[人間の直感の不確実性—数学的な正しさと乖離している場合があることを知っていますか—](#)」で、「誕生日のパラドックス」について紹介したところ、複数の照会があり、いくつかの質問を受けた。今回は、いろいろなケースで「誕生日が一致する確率」について、どのようになるのか紹介したい。

具体的には、①「(少なくとも2人1組ではなくて)2組以上のペアの誕生日が一致している確率」はどうなるのか、②「3人以上の誕生日が一致している確率」はどうなるのか、という問題である。

これからの年末年始の時期等に、多くの人が集まる場で話題にでもしてもらえればと思って、このトピックを採り上げることにした。

問題の設定

ある部屋に50人のグループがいると仮定する。以下のように、いろいろなケースで誕生日が一致する確率を考えてみる(なお、今回の計算では、1年365日とした)。

問題(その1)

まずは、個人を特定せずに、グループ全体での発生確率の問題を考える。

問題1 部屋の中の全員の誕生日が異なる確率 P_0

問題2 (部屋の中の誰でもよいので) 最低1組2人のペアの誕生日が一致している確率 Q_1

(ここでは、3人以上の誕生日が一致している場合は、含まれないとする。以下、同様)

(※)これに対して、少なくとも誰かと誰かの誕生日が一致する確率 $R_1 (= 1 - P_0)$ も考えられる。

この場合には、3人以上の誕生日が一致している場合も含まれる。以前の「研究員の眼」で述べたのは、このケースであり、この確率が人間の直感に比べて予想外に高いことが「誕生日のパラドックス」と呼ばれる由縁となっている。

問題3 丁度1組2人のペアの誕生日が一致している確率 P_1

問題 4 最低 2 組のペアの誕生日が一致している確率 Q2

問題 5 丁度 2 組 4 人のペアの誕生日が一致している確率 P2

問題 6 最低 3 組のペアの誕生日が一致している確率 Q3

問題 7 丁度 3 組のペアの誕生日が一致している確率 P3

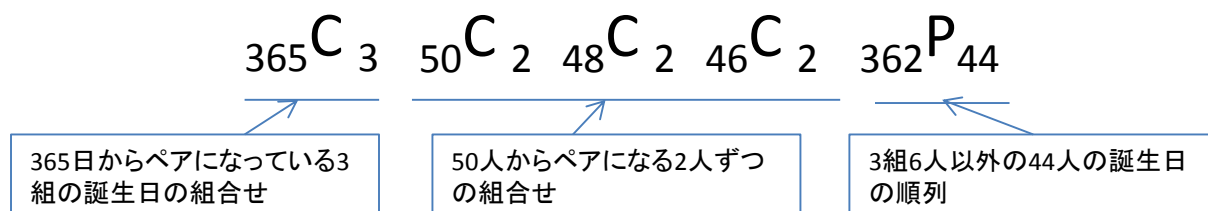
以下、同様に、丁度 k 組の 2 人ずつのペアの誕生日が一致している確率 Pk や最低 k 組のペアの誕生日が一致している確率 Qk が考えられる。この場合、

$$Q_k = P_1 + P_2 + \dots + P_k, \text{あるいは、} P_k = Q_k - Q_{(k-1)}$$

となる。

問題(その1)の解答

例えば、問題 7 を考えると、6 人 (2 人×3 組) が 3 種類の誕生日で、44 人は残りの 362 日の全て異なる誕生日となるので、そのような誕生日が発生するパターン数は、



となる。よって、求める確率は、これを全ての発生パターンの「365 の 50 乗」で割ることによって、

$$P_3 = (365 C_3 \quad 50 C_2 \quad 48 C_2 \quad 46 C_2 \quad 362 P_{44}) / 365^{50}$$

で、結果として P3=22.2% となる。

以上の考え方に基づいて計算した結果をまとめると、次表の通りとなる。

誕生日が一致する確率(50名の場合)

誕生日が一致するペアの数(k)	0組	1組	2組	3組	4組	5組
丁度その組数である確率(Pk)	3.0%	11.5%	20.4%	22.2%	16.4%	8.8%
その組数以上である確率(Qk)	87.4%	84.4%	72.9%	52.5%	30.3%	13.9%
		6組	7組	8組	9組	10組以上
		3.6%	1.1%	0.3%	0.1%	0.0%
		5.0%	1.4%	0.3%	0.1%	0.0%

(注)「0組」は、「3人以上の誕生日が一致しているケース」も含めて考えた場合に、全員の誕生日が異なる確率としている。

これによると、50 人のグループでは、以下の状況になっている。

- ①全員の誕生日が異なる確率は「0 組」の数の 3.0%であることから、少なくとも誰かと誰かの誕生日が一致している確率は 97.0%となる。
- ②誕生日が一致するペアの数としては、「3 組」が最も多い。
- ③さすがに 7 組以上のペアが発生する確率は 1.4%と低くなるが、それでも 5 組のペアが発生する確率は 8.8%もあり、6 組のペアが発生する確率も 3.6%ある。

- ④一方で、全く誕生日が一致しないか、1組2人のペアの誕生日しか一致しない確率は、わずか14.5% (3.0%+11.5%)でしかない。このことはまた、誕生日が他の人と一致している人が3人以上 (1組でも3人以上又は2組以上) いる確率は、85.5%ということになる。
- ⑤2組以上のペアが発生する確率は72.9%、3組以上のペアが発生する確率は52.5%となる。
- ⑥上記の表の0組以上の発生確率が87.4%となっているが、これと100%との差異の12.6%は、今回の計算で考慮されていない、「少なくとも3人以上の誕生日が一致している組が1つは存在している確率」となる。
- ⑦即ち、例えば、上記の表の「3組」には、「1組が3人の誕生日が一致、2組 (あるいは3組) が2人の誕生日が一致」しているケース等は含まれていない。こうしたケースを含めれば、上記の表の確率はさらに高くなることになる。
- ⑧因みに、上記の表に基づくと、誕生日が一致するペアの数の期待値は、2.6組ということになる。50人いれば、平均して2.6組のペアの誕生日が一致していることになる。⑦で述べた3人以上の誕生日が一致しているケースも含めれば、さらに高い期待値になる。

前回の研究員の眼は、①の確率の高さについて触れていたが、今回の②以下の結果についても、一般の感覚からすると、再びかなり高い確率だと感じるのではないか、と思われる。

50人のグループで考えても、例えば誕生日が一致しているペアが5組あることも決して珍しくない、ということになる。

なお、上に述べたように、「少なくとも3人以上の誕生日が一致している組が1つは存在している確率」は12.6%であるが、これが「少なくとも4人以上の誕生日が一致している組が1つは存在している確率」となると1%未満でぐっと小さくなる。一定数の組のペアが一致しているケースに比べて、一定人数の誕生日が一致するケースは、より発生が限られている。

問題(その2)

次に、特定の個人に関しての発生確率を考える。

問題 8 自分と同じ誕生日の人がいない確率 S_0

問題 9 自分と誰かの誕生日が一致している確率(自分と同じ誕生日の人が他に1人以上いる確率)
 $T_1 (=1-S_0)$

問題 10 自分と同じ誕生日の人が他に丁度1人いる確率 S_1

問題 11 自分と同じ誕生日の人が他に2人以上いる確率 T_2

問題 12 自分と同じ誕生日の人が他に丁度2人いる確率 S_2

問題 13 自分と同じ誕生日の人が他に3人以上いる確率 T_3

問題 14 自分と同じ誕生日の人が他に丁度3人いる確率 S_3

以下、同様に、自分と同じ誕生日の人が丁度 k 人いる確率 S_k や自分と同じ誕生日の人が最低 k 人いる確率 T_k が考えられる。この場合、

$$T_k = S_1 + S_2 + \dots + S_k, \text{ あるいは、 } S_k = T_k - T_{(k-1)}$$

となる。

問題(その2)の解答

例えば、自分と同じ誕生日の人が他に丁度 2 人いるということは、2 人が自分と同じ誕生日（その確率はそれぞれ $1/365$ ）で、それ以外の 47 人は自分とは異なる誕生日（その確率はそれぞれ $364/365$ ）であり、自分以外の 49 人からの 2 人の組合せは ${}_{49}C_2$ 通りあるので、

$$S2 = {}_{49}C_2 \left(\frac{1}{365} \right)^2 \left(\frac{364}{365} \right)^{47}$$

で、結果として $S2=0.8\%$ となる。

以上をまとめると、以下の表の通りとなる。

自分と誕生日が一致する確率(50名の場合)

自分と誕生日が一致する人数(k)	0人	1人	2人	3人
丁度その人数である確率(Sk)	87.4%	11.8%	0.8%	0.0%
その人数以上である確率(Tk)	100.0%	12.6%	0.8%	0.0%

こちらの確率は、さすがに低いものとなる。

なお、人数が 100 名及び 200 名の場合には、以下の通りとなり、自分と同じ誕生日の人がいる確率はそれぞれ 23.8%、42.1%と高くなっていく。さらには、自分と同じ誕生日の人が 2 人以上いる確率もそれぞれ 3.1%、10.4%と高くなっていく。

自分と誕生日が一致する確率(100名の場合)

自分と誕生日が一致する人数(k)	0人	1人	2人	3人
丁度その人数である確率(Sk)	76.2%	20.7%	2.8%	0.2%
その人数以上である確率(Tk)	100.0%	23.8%	3.1%	0.3%

自分と誕生日が一致する確率(200名の場合)

自分と誕生日が一致する人数(k)	0人	1人	2人	3人
丁度その人数である確率(Sk)	57.9%	31.7%	8.6%	1.6%
その人数以上である確率(Tk)	100.0%	42.1%	10.4%	1.8%

まとめ

[以前の研究者](#)の眼と同様に、今回の結果についても驚かれた方が多いのではないかと思います。

ここでは誕生日をテーマにしているが、一般的に人間は、何かの事象の発生確率を想定する場合に、無意識的に自分を中心に起こるケースを想定して、その発生確率は低いものだと想定しているのではないかと。ところが、グループ全体として考える場合には、個人が想定しているよりもかなり高い確率でその事象が発生することになる。

このことは、物事を考えていく場合に何か示唆するものがあるのではないかと思います。

順列・組み合わせの問題については、中学・高校時代にかなり苦労された方も多いのではないかと思います。しかし、こうやって考えてみると、その解答を導き出すのは必ずしも易しくないとしても、その結果には感動させられることもあるのではないかと思います。これを機に、今一度若い頃に戻って、いろいろな順列・組み合わせが関係してくる確率の問題を考えてみるのも、頭の体操になってよいのではないかと。