Probabilité qu'au moins deux personnnes parmi n aient leur anniversaire le même jour

Robinson Cartez

Le calcul

Soit p la probabilité qu'au moins 2 personnes parmi n aient leur anniversaire le même jour de l'année.
Et soit $\overline{p}$ la probabilité que chacune des n personnes du groupe considéré aient leur anniversaire un jour différent.

On a alors que la somme des deux probabilités vaut l'événement certain

$$p + \overline{p} = 1$$

De plus, l'expérience montre que pour calculer une probabilité dont l'énoncé implique un "au moins", il vaut mieux travailler avec l'événement contraire, ici $\overline{p}$.

Il s'agit pour nous d'évaluer cette probabilité. Pour ce faire, on prends les n personnes une à une et on calcule la probabilité que leur date d'anniversaire tombe un jour différent de toutes les autres.

Pour la première c'est

$$\frac{365}{365}$$

pour la deuxième c'est

$$\frac{364}{365}$$

pour la troisième c'est

$$\frac{363}{365}$$

etc. Ce qui donne

$$\overline{p} = \frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{365 - n + 1}{365}$$

$$= \frac{365!}{(365 - n)!} \cdot \frac{1}{365^n}$$

Grâce à cette formule, on peut dire quel est la probabilité, dans un groupe de deux personnes ($n = 2$), que deux personnes au moins aient leur anniversaire le même jour :

$$n = 2$$

$$\overline{p} = \frac{365!}{363!} \cdot \frac{1}{365^2} = \frac{365 \cdot 364}{365 \cdot 365} = \frac{364}{365}$$

$$p = 1 - \overline{p} \approx 0.00274 \approx 0.3\%$$

Et si on a la même question pour un groupe de 5 personnes ($n = 5$) :

$$n = 5$$

$$\overline{p} = \frac{365!}{360!} \cdot \frac{1}{365^5} = \frac{365 \cdot 364 \cdot 363 \cdot 362 \cdot 361}{365 \cdot 365 \cdot 365 \cdot 365 \cdot 365}$$

$$p = 1 - \overline{p} \approx 0.02714 \approx 2.7\%$$

Le challenge

La question que les mathématiciens se sont posée est la suivante :

> "Combien de personnes faut-il pour qu'on ait 50% de chances de trouver au moins deux personnes ayant leur anniversaire le même jour ?"

La difficulté est que dans la formule close que l'on a trouvée pour $\overline{p}$, le n figure dans la parenthèse d'une factorielle : $(365 - n)!$. Il nous faut donc trouver un moyen de l'enlever.

Après avoir travaillé un moment sur la formule, en la tournant et la retournant, on arrive à l'écrire comme ceci :

$$\begin{aligned}\overline{p} &= 1 \cdot \frac{364}{365} \cdot \frac{363}{365} \cdot \ldots \cdot \frac{365-n+1}{365} \\ &= 1 \cdot (1 - \frac{1}{365}) \cdot (1 - \frac{2}{365}) \cdot \ldots \cdot (1 - \frac{n-1}{365})\end{aligned}$$

À défaut de trouver une formule exacte, les mathématiciens ont trouvé une formule approchée. L'efficacité de cette formule se comprend lorsque l'on réalise que n represente un nombre de personnes, donc un entier.

Une recherche dans un formulaire mathématique bien fourni, permet d'identifier

$$e^x = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + o(x^n)$$

comme candidate pour approcher chacun des facteurs de la formule pour $\overline{p}$ ci-dessus.

En effet, lorsque x est proche de zéro on a l'approximation suivante :

$$e^{-x} \approx (1 - x)$$

Puisque nous avons observé que dans un groupe de $365 + 1$ personnes on trouve à coup sûr au moins deux personnes ayant leur anniversaire le même jour, on peut considérer que les fractions

$$\frac{1}{365}, \frac{2}{365}, \frac{3}{365}, \frac{4}{365}, \text{ etc.}$$

comme “proches” de zéro. Nous allons donc utiliser

$$x = \frac{k}{365}, \quad \text{pour } k = 0, 1, \ldots, n-1$$

Par exemple, pour $k = 1$, $x = \frac{1}{365}$ on a

$$\begin{aligned} e^{-x} &\approx 0.997264024 \\ 1 - \frac{1}{365} &\approx 0.997260274\end{aligned}$$

Et pour $k = 2$, $x = \frac{2}{365}$ et on a

$$\begin{aligned} e^{-x} &\approx 0.994535533 \\ 1 - \frac{2}{365} &\approx 0.994520548\end{aligned}$$

Il suffit maintenant d'aligner ce que nous avons et de renverser la formule pour $\overline{p}$ en isolant n :

$$\begin{aligned}\overline{p} &= 1 \cdot (1 - \frac{1}{365}) \cdot (1 - \frac{2}{365}) \cdot \ldots \cdot (1 - \frac{n-1}{365}) \\ &= \prod_{k=0}^{n-1} (1 - \frac{k}{365}) \\ &\approx \prod_{k=0}^{n-1} e^{-\frac{k}{365}} \\ &= e^{-\frac{\sum_{k=0}^{n-1} k}{365}} \\ &= e^{-\frac{n(n-1)}{2} \cdot \frac{1}{365}}\end{aligned}$$

Première remarque. Le produit $e^p \cdot e^q = e^{p+q}$. D'où la somme à l'exposant de e.

Ensuite, la somme des m premiers entiers naturels vaut

$$\frac{m(m+1)}{2}$$

Dans la formule d'origine on a la somme des $n - 1$ premiers entiers (le zéro ne modifie pas le résultat).

Dernières remarques : on va appliquer la fonction logarithme naturel afin d'isoler complètement n. Et puisque cette expression montre n et n^2 nous allons utiliser les formules de Viète...

De

$$p + \overline{p} = 1$$

on tire que, la probabilité chercheé vaut

$$p = 1 - \overline{p}$$

et afin de travailler avec des vraies égalités, posons

$$r = 1 - e^{-\frac{n(n-1)}{730}}$$

sachant qu'ensuite nous aurons que

$$p \approx r$$

d'après l'approximation des facteurs de $\overline{p}$ trouvée plus haut. Isolons n

$$
\begin{aligned}
e^{-\dfrac{n(n-1)}{730}} &= 1-r \\
\Leftrightarrow -n(n-1) &= \ln(1-r)\cdot 730 \\
\Leftrightarrow n^2 - n + \ln(1-r)\cdot 730 &= 0
\end{aligned}
$$

$$\Delta = 1 - 4\cdot\ln(1-r)\cdot 730$$

Comme la probabilité $p \approx r$ est comprise entre 0 et 1, le discriminant calculé ci-dessus sera plus grand ou égal à zéro

$$\Delta \geq 0$$

il s'en suit que la quadratique a des solutions réelles, à savoir

$$n = \frac{1 \pm \sqrt{1-4\cdot\ln(1-r)\cdot 730}}{2}$$

Puisque c'est des approximations afin de trouver le nombre entier n, on suppose que $\overline{p} < 1$ et donc $p > 0$ et de plus $\ln(1-p) < 0$.

Ainsi, comme $p \approx r$, on a

$$n \approx \frac{1 + \sqrt{1-4\cdot\ln(1-p)\cdot 730}}{2}$$

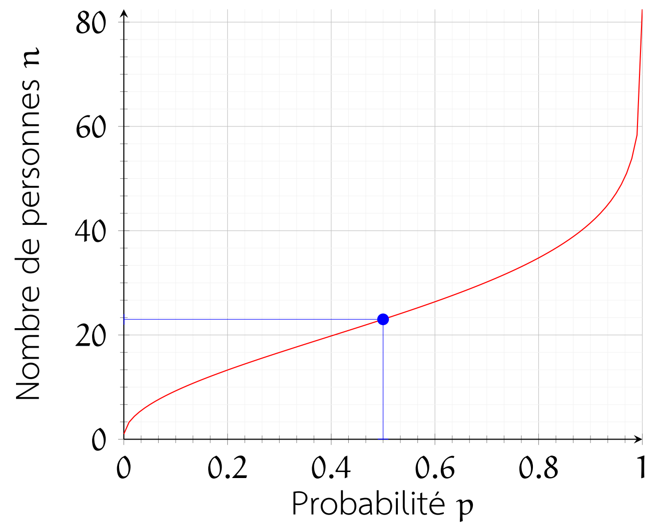
Qui est la formule (approchée) recherchée donnant le nombre de personnes d'une groupe en fonction de la probabilité p qu'au moins deux personnes dans le groupe aient leur anniversaire le même jour de l'année.

Donc, pour répondre à la question en encadré, posons $p = 0.5$ et calculons n :

$$n \approx \frac{1 + \sqrt{1-4\cdot\ln(0.5)\cdot 730}}{2} \approx 22.9999432$$

Il nous faut alors **un groupe de 23 personnes**.

Voici ce que donne le graphique du nombre de personnes en fonction de la probabilité souhaitée :



Et contrairement au raisonnement utilisant le principe des tiroirs de Dirichelet[1], on a besoin de moins de 366 personnes dans un groupe pour être sûr d'en trouver deux qui ont leur anniversaire le même jour de l'année.

Tableau des probabilités			
Nb pers.	Proba	Nb pers.	Proba
20	0.41	41	0.89
21	0.44	42	0.91
22	0.47	43	0.92
23	0.50	44	0.93
24	0.53	45	0.93
25	0.56	46	0.94
26	0.59	47	0.95
27	0.62	48	0.95
28	0.64	49	0.96
29	0.67	50	0.97
30	0.70	51	0.97
31	0.72	52	0.97
32	0.74	53	0.98
33	0.76	54	0.98
34	0.78	55	0.98
35	0.80	56	0.99
36	0.82	57	0.99
37	0.84	58	0.99
38	0.85	59	0.99
39	0.87	60	0.99
40	0.88	61	0.99

1. Principe qui affirme que si m chaussettes occupent n tiroirs, avec $m > n$, alors au moins un tiroir doit contenir strictement plus d'une chaussette.