

Supposons que des candidat.e.s en nombre fini et connu se présentent à une entrevue d'emploi dans un ordre aléatoire et qu'après chaque entretien, la direction puisse offrir le poste à la personne interviewée ou rejeter sa candidature mais que la décision, une fois prise, soit irrévocable. Quelle stratégie adopter pour maximiser les chances de recruter la meilleure candidature ?

La règle des 37 %

Christian Genest
Université McGill

Ce problème d'optimisation aux origines obscures est attesté dans la littérature scientifique depuis le milieu des années 1950. Formulé tantôt comme la quête d'un.e secrétaire, d'une propriété, d'un.e partenaire de vie, voire d'une dot, il doit sa popularité à la forme étonnante de sa solution, qui fait intervenir le nombre d'Euler, $e = 2,71828...$

Comme nous le verrons, la stratégie optimale consiste à laisser passer un certain nombre de candidatures et à sélectionner ensuite, si possible, la première personne qui se révèle meilleure que toutes celles qui l'ont précédée. Fait remarquable, il se trouve qu'à mesure que croît le nombre de candidatures, la proportion d'entre elles à rejeter d'emblée tend vers $1/e \approx 37\%$. Qui plus est, la probabilité que la meilleure personne soit retenue s'approche de $1/e$ et la probabilité que nul ne soit embauché tend aussi vers $1/e$. Nous évoquerons ensuite certaines variantes du problème.

Formulation précise

Supposons que le nombre $n \geq 2$ de candidatures au poste soit connu à l'avance. Dénотons ces candidatures C_1, \dots, C_n dans leur ordre d'arrivée et supposons qu'il soit possible de les classer sans ambiguïté. Ainsi, après avoir interviewé l'individu C_n , la personne chargée du recrutement possède un classement relatif de C_1, \dots, C_n et peut choisir ou non d'offrir le poste à C_n en fonction de cette information. On dénote C_S la candidature retenue à la fin de ce processus, s'il en est.

Si l'on savait à l'avance que les candidatures se présenteraient dans l'ordre, de la meilleure à la pire, on embaucherait évidemment la première personne interviewée. De même, si les candidatures étaient classées dans l'ordre

contraire, on offrirait alors le poste à la toute dernière personne rencontrée. On poserait ainsi $S = 1$ dans le premier cas et $S = n$ dans le second.

Pour que le problème ait un intérêt, supposons plutôt que les entrevues se déroulent dans un ordre aléatoire, de sorte que la probabilité que la candidature C_i soit la meilleure, dénoté $C_i = C_M$ soit $1/n$. Exprimé en ces termes, le problème consiste à identifier une règle de décision qui maximise la probabilité que C_S , la candidature sélectionnée, s'il en est, soit effectivement C_M .

Exemple dans le cas $n = 3$

Supposons que trois candidatures soient numérotées 1, 2, 3, de la pire à la meilleure. Il y a alors six ordres d'arrivée possibles. L'un d'eux serait $(C_1, C_2, C_3) = (1, 2, 3)$, les cinq autres étant $(1, 3, 2)$, $(2, 1, 3)$, $(2, 3, 1)$, $(3, 1, 2)$ et $(3, 2, 1)$. On les suppose équiprobables.

Si par exemple on choisit d'offrir le poste au premier venu, on embauchera la meilleure personne si et seulement si les candidat.e.s sont interviewé.e.s dans l'ordre $(3, 1, 2)$ ou $(3, 2, 1)$. La probabilité de succès associée à cette stratégie est donc $2/6 = 1/3$.

On peut toutefois faire mieux en rejetant d'emblée la première candidature et en offrant le poste à la deuxième personne si elle est jugée meilleure que la première. Sinon, on fait l'offre à la troisième personne, pourvu qu'elle se classe mieux que les deux précédentes.

Selon cette stratégie, la meilleure candidature sera retenue dans les trois cas suivants : $(1, 3, 2)$, $(2, 1, 3)$ et $(2, 3, 1)$. La probabilité de succès est donc $3/6 = 1/2 > 1/3$. Aussi cette stratégie est-elle préférable à la première, quoiqu'elle ait le désavantage de ne pas toujours mener

à une embauche. En effet, personne ne sera retenu si les entrevues ont lieu dans l'ordre (3, 1, 2) ou (3, 2, 1), tandis que dans le cas (1, 2, 3), on sélectionnera la seconde candidature, qui n'est pas la meilleure.

Forme de la règle optimale

Fixons $i \in \{1, \dots, n\}$ et supposons que les candidat.e.s C_1, \dots, C_i aient été interviewé.e.s. Pour maximiser ses chances de recruter la meilleure candidature, la direction ne devrait offrir le poste à la personne C_i que si les deux conditions suivantes sont réunies :

- La candidature doit être éligible, c'est-à-dire s'avérer la meilleure parmi C_1, \dots, C_i .
- La probabilité que $C_i = C_M$ doit être plus grande que pour n'importe quelle stratégie qui pourrait être employée une fois que la candidature C_i aura été rejetée, soit

$$\Pr(E_i) > \Pr(F_i), \quad (1)$$

où E_i est l'événement que $C_i = C_M$ sachant que C_i est éligible, tandis que F_i est l'événement « recruter C_M avec la meilleure stratégie fondée sur C_{i+1}, \dots, C_n ». Or, le terme de gauche de l'inégalité (1) est strictement croissant en i , puisque $\Pr(E_i) = i/n$. En effet, la probabilité que C_i soit éligible n'est autre que la probabilité que la meilleure des i premières candidatures soit la dernière, soit $1/i$. Par ailleurs, la probabilité que la candidature i soit la meilleure de toutes est $1/n$. Remarquez que si $C_i = C_M$, alors C_i est la meilleure et est donc automatiquement éligible. Par la définition de probabilité conditionnelle, on a donc

$$\Pr(E_i) = (1/n) / (1/i) = i/n.$$

En revanche, le terme de droite de l'inégalité (1) est strictement décroissant en i , puisque si $i < j$, toute stratégie qui s'appuie sur C_{j+1}, \dots, C_n fait partie des stratégies fondées sur C_{i+1}, \dots, C_n .

Par conséquent, la meilleure stratégie est forcément de la forme suivante :

- Rejeter les $r-1$ premières candidatures, pour un certain $r \in \{1, \dots, n\}$.
- Choisir la première candidature $i \geq r$ qui s'avère éligible, c'est-à-dire qui se classe mieux que toutes les précédentes.

La figure 1 illustre cette règle pour deux valeurs de r , soient 3 et 6, lorsque le nombre de candidatures est $n=10$. Dans la figure, la taille de chaque personne reflète sa valeur et le x indique la meilleure d'entre elles. La zone bleue correspond à la phase d'exploration et la zone jaune à celle pendant laquelle la direction s'autorise à faire une offre ou pas. La zone verte regroupe les candidatures non rencontrées. Enfin, l'individu encerclé est celui qui est choisi au terme du processus. Dans le volet inférieur, on sélectionne la meilleure personne du lot ; dans le volet supérieur, ce n'est pas le cas.

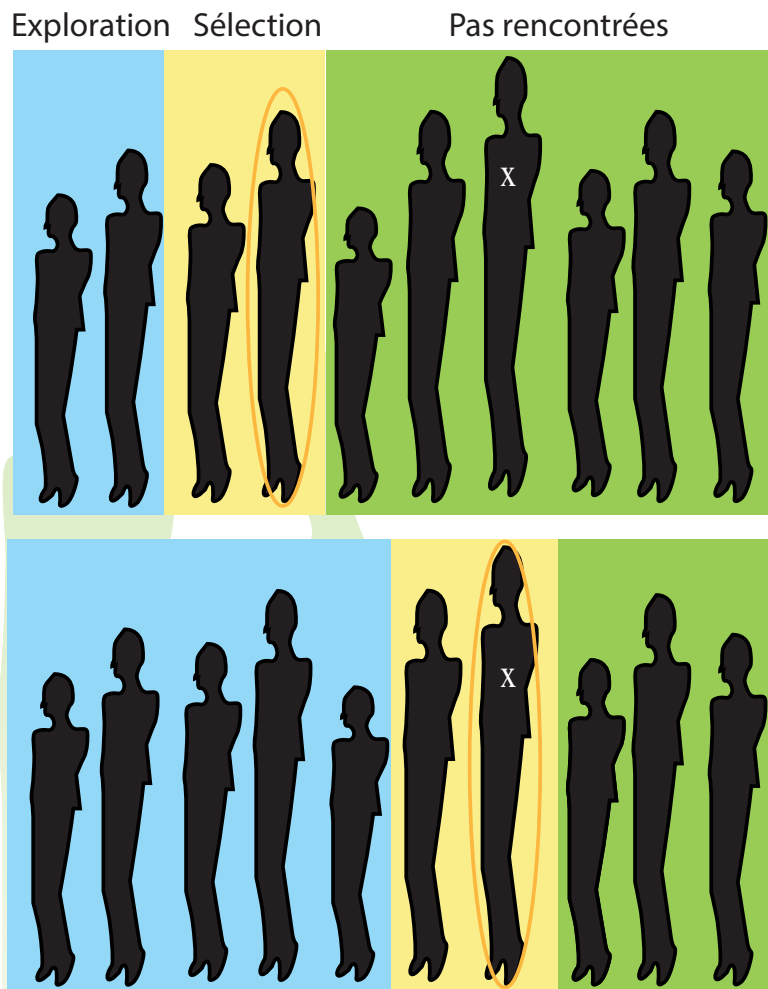


Figure 1.

Illustration de la stratégie d'embauche selon que l'on rejette d'emblée les $r = 2$ premières ou les $r = 5$ premières candidatures d'un lot de $n = 10$, la personne la plus grande étant la meilleure.

Détermination de la valeur de r

Étant donné $r \in \{1, \dots, n\}$, dénotons par $P_n(r)$ la probabilité que la stratégie ci-dessus conduise à choisir la meilleure candidature dans l'ensemble $\{C_1, \dots, C_n\}$. Quand $r = 1$, la stratégie se réduit à sélectionner la première personne interviewée, puisque l'ensemble $\{C_1, \dots, C_{r-1}\}$ est vide. On a alors $S = 1$ et la probabilité que $C_S = C_M$ est $1/n$. On a donc $P_n(1) = 1/n$.

Si $r \geq 2$, la probabilité $P_n(r)$ peut être calculée en distinguant $n - r + 1$ cas distincts, selon que pour $i \in \{r, \dots, n\}$, la candidature C_i est à la fois éligible et la meilleure de toutes. On peut donc écrire

$$P_n(r) = \sum_{i=r}^n \Pr(C_i = C_S \text{ et } C_i = C_M).$$

Par la formule de probabilité conditionnelle, on a

$$\begin{aligned} \Pr(C_i = C_S \text{ et } C_i = C_M) \\ = \Pr(C_i = C_S \mid C_i = C_M) \times \Pr(C_i = C_M) \end{aligned}$$

et $\Pr(C_i = C_M) = 1/n$ du fait que l'ordre dans lequel les candidatures sont examinées est aléatoire.

Par ailleurs, si $C_i = C_M$, c'est-à-dire si C_i est la meilleure candidature, elle ne pourra être sélectionnée que si la meilleure des $i - 1$ candidatures qui la précèdent se trouve parmi les $r - 1$ candidatures qui ont été rejetées d'emblée. Si tel n'était pas le cas, on se serait arrêté avant !

Sachant que l'ordre dans lequel les candidatures sont étudiées est aléatoire, cet événement se produit avec probabilité $(r-1)/(i-1)$. Ainsi, pour tout $r \in \{2, \dots, n\}$, la probabilité recherchée est

$$\begin{aligned} P_n(r) &= \sum_{i=r}^n \frac{r-1}{i-1} \times \frac{1}{n} \quad (2) \\ &= \frac{r-1}{n} \left(\frac{1}{r-1} + \frac{1}{r} + \dots + \frac{1}{n-1} \right). \end{aligned}$$

Remarquons que pour tout $n \geq 2$, on a

$$\begin{aligned} P_n(2) &= \frac{1}{n} \left(\frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{n-1} \right) \\ &> \frac{1}{n} = P_n(1), \end{aligned}$$

tel que déjà illustré plus haut dans le cas $n = 3$. Comme on l'a aussi observé dans cet exemple, le rejet automatique d'un certain nombre de candidatures entraîne un risque de ne retenir personne à l'issue du processus d'embauche. La probabilité que cela se produise est la même que celle que la meilleure candidature figure parmi les $r - 1$ candidatures rejetées, soit $Q_n(r) = (r - 1)/n$. Notons en outre que $P_n(n) = 1/n$, comme il se doit.

Calcul de r et $Q_n(r)$ pour n donné

Étant donné une valeur de $n \geq 2$, on peut aisément calculer la valeur de $P_n(r)$ pour tout choix de $r \in \{2, \dots, n\}$. Si $n = 5$, par exemple, on pose successivement $r = 2, 3, 4, 5$ dans la formule (2) et on trouve

$$P_5(2) = \frac{1}{5} \left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} \right) \approx 0,417,$$

$$P_5(3) = \frac{2}{5} \left(\frac{1}{2} + \frac{1}{3} + \frac{1}{4} \right) \approx 0,433,$$

$$P_5(4) = \frac{3}{5} \left(\frac{1}{3} + \frac{1}{4} \right) \approx 0,35,$$

$$P_5(5) = \frac{4}{5} \left(\frac{1}{4} \right) \approx 0,20.$$

Par conséquent, la probabilité $P_5(r)$ est maximisée lorsque $r = 3$, ce que l'on dénote $r_5 = 3$. De plus, la probabilité de ne sélectionner personne est alors égale à $Q_5(r_5) = 2/5 = 0,4$.

De façon générale, on peut démontrer que pour une valeur de n donnée, la fonction $r \mapsto P_n(r)$ est « unimodale ». On entend par là que cette fonction est d'abord croissante en r , qu'elle atteint éventuellement un sommet en une valeur r_n , après quoi elle se met à décroître. Une illustration de cette remarque se trouve à la figure 2, qui correspond au cas $n = 10$.

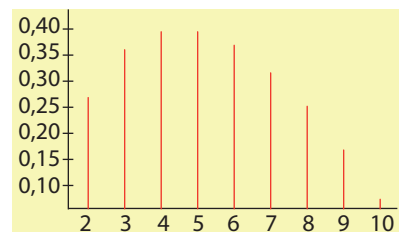


Figure 2.

Graphique des valeurs de $P_n(r)$ lorsque $n = 10$ et que r varie entre 2 et 10. La fonction $r \mapsto P_{10}(r)$ est unimodale et son maximum se trouve en $r = 4$, soit $P_{10}(4) \approx 0,39869$, valeur légèrement supérieure à $P_{10}(5) \approx 0,39825$.

Le tableau 1 donne, pour chaque valeur de $n \in \{2, \dots, 5, 10, 20, 30, 40, 50, 100\}$, la valeur de $r \in \{2, \dots, n\}$, dénotée r_n qui maximise $P_n(r)$. La valeur de $P_n(r_n)$ y est également précisée, de même que la probabilité $Q_n(r_n)$ de ne sélectionner personne en adoptant cette stratégie.

En consultant le tableau 1, on observe que la probabilité $P_n(r_n)$ est monotone décroissante en n , ce qui reflète le fait que même avec une stratégie optimale, il devient de plus en plus ardu d'identifier la meilleure personne à mesure que le nombre de candidatures augmente.

n	r_n	$P_n(r_n)$	$Q_n(r_n)$
2	2	0,5	0,5
3	2	0,5	0,33
4	2	0,458	0,25
5	3	0,433	0,4
10	4	0,399	0,3
20	8	0,384	0,35
30	12	0,379	0,367
40	16	0,376	0,375
50	19	0,374	0,36
100	38	0,371	0,37

Tableau 1. Valeur r_n de r permettant de maximiser $P_n(r)$ et valeurs de $P_n(r_n)$ et $Q_n(r_n)$ correspondantes pour $n \in \{2, 3, 4, 5, 10, 20, 30, 40, 50, 100\}$.

Limite de r_n et $P_n(r_n)$ quand $n \rightarrow \infty$

Il est instructif de déterminer comment le rapport r_n/n varie au fur et à mesure que n augmente. Les données du tableau 1 suggèrent que ce rapport tend à décroître, bien que pas de façon monotone. Soit t la limite de r_n/n quand $n \rightarrow \infty$. Puisqu'alors on a $1/n \rightarrow 0$, il s'ensuit que

$$Q_n(r_n) = (r_n - 1)/n = r_n/n - 1/n \rightarrow t.$$

Par ailleurs, on peut montrer que si $n \rightarrow \infty$, on a

$$P_n(r_n) \rightarrow P(t) = -t \ln(t). \tag{3}$$

Se référer à l'encadré pour une démonstration de ce fait.

Or, t est par définition la proportion qui maximise $P(t)$. En prenant la dérivée de $P(t)$ par rapport à t , on trouve que $P'(t) = -\ln(t) - 1$ et $P'(t) = 0$ si et seulement si $t = 1/e$. Puisqu'en

outre $P''(t) = -1/t < 0$ pour tout $t \in]0, 1[$, on conclut qu'asymptotiquement, c'est-à-dire lorsque $n \rightarrow \infty$, la proportion optimale de candidatures à rejeter d'office est de $1/e \approx 37\%$, tel qu'annoncé. Ce faisant, la probabilité $Q_n(r_n)$ de ne recruter personne est aussi égale à $1/e$. Et puisque $P(1/e) = 1/e$, l'emploi de cette règle donne à son utilisateur une probabilité $1/e \approx 37\%$ d'identifier la meilleure candidature.

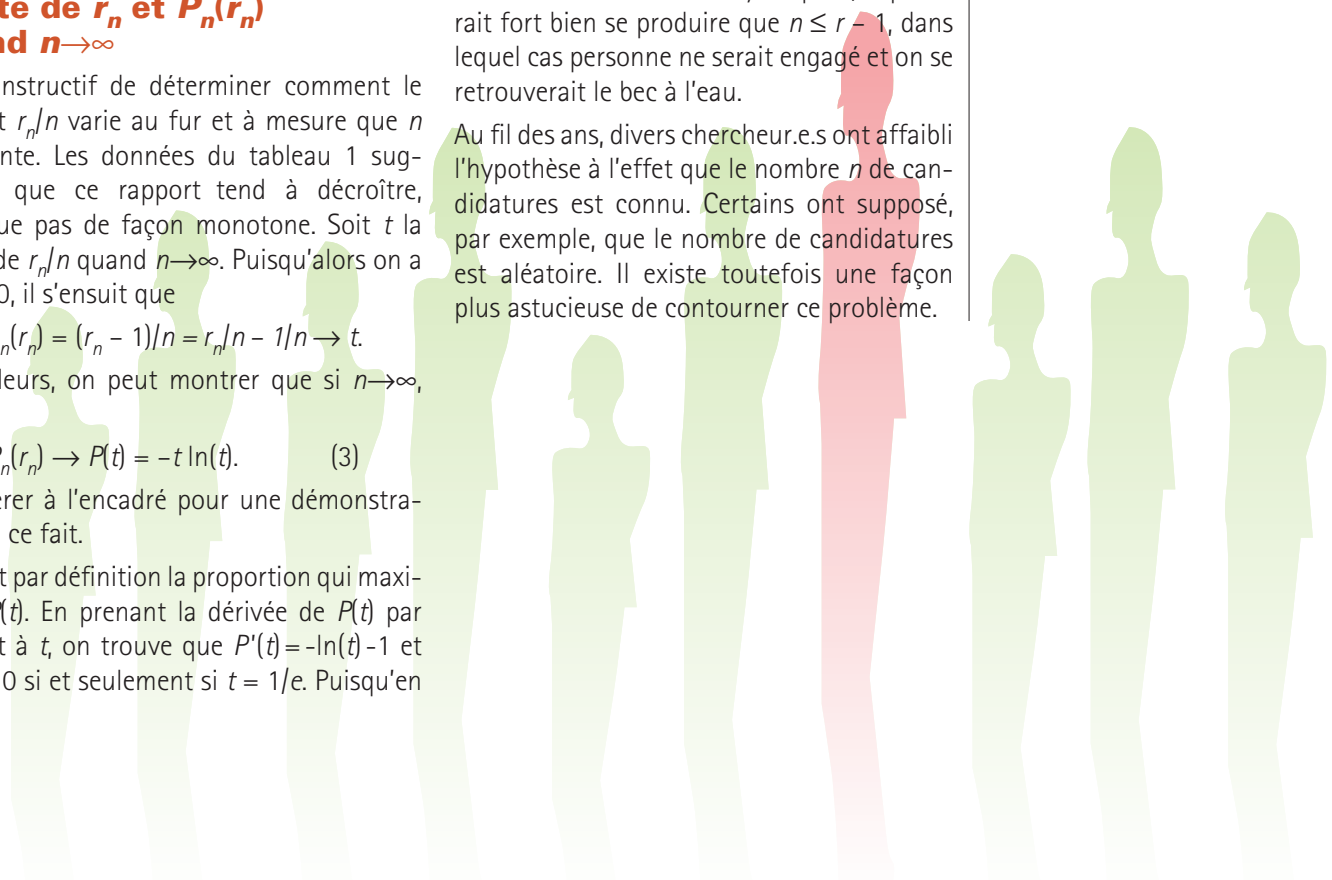
La fraction $1/e$ est donc tout à la fois la proportion optimale de candidatures à rejeter d'office, la probabilité d'identifier la meilleure candidature et la probabilité de n'embaucher personne, ce qui est remarquable !

Limitations

Bien que séduisante, la solution du problème précédent n'a de sens que si les candidat.e.s au poste ne sont pas informé.e.s de la stratégie qui sera employée pour opérer la sélection. Autrement, personne ne se bousculerait au portillon pour passer l'entrevue tôt !

Par ailleurs, il est crucial que le nombre total de candidatures, n , soit connu à l'avance. Sinon, il est impossible de fixer d'entrée de jeu le nombre $r - 1$ de dossiers qui seront écartés d'emblée. Si on s'y risquait, il pourrait fort bien se produire que $n \leq r - 1$, dans lequel cas personne ne serait engagé et on se retrouverait le bec à l'eau.

Au fil des ans, divers chercheur.e.s ont affaibli l'hypothèse à l'effet que le nombre n de candidatures est connu. Certains ont supposé, par exemple, que le nombre de candidatures est aléatoire. Il existe toutefois une façon plus astucieuse de contourner ce problème.



Une approche différente

Au début des années 1980, le mathématicien allemand Thomas Bruss a proposé une formulation ingénieuse du problème du choix de la meilleure candidature dans laquelle on ne fixe pas le nombre de candidatures à examiner, mais plutôt la période de temps pendant laquelle on cherche à combler le poste. Cette approche a aussi le mérite d'être plus proche de la pratique.

Soit $[0, T]$ l'intervalle de temps prédéterminé à l'intérieur duquel on cherche à combler le poste. Formulons en outre les hypothèses suivantes quant au processus d'arrivée des candidatures :

- Un nombre aléatoire N de candidatures est observé dans l'intervalle $[0, T]$.
- Sachant $N = n$, les temps d'arrivée des n candidatures sont mutuellement indépendants et tous de même loi de probabilité représentée par une densité continue f concentrée sur l'intervalle $[0, T]$.
- Les $n!$ classements possibles des candidatures, au plan de leur aptitude à l'emploi, sont équiprobables.

Démonstration de l'énoncé (3)

Pour démontrer l'énoncé (3), on exploite le fait que pour tout entier m suffisamment grand, la somme $1 + 1/2 + \dots + 1/m$ est à peu près égale à $\ln(m) + e$. Ainsi, pour de grandes valeurs de r et m , on a

$$\sum_{i=r}^n \frac{1}{i} = \sum_{i=1}^n \frac{1}{i} - \sum_{i=1}^{r-1} \frac{1}{i} \approx \ln\left(\frac{n}{r-1}\right).$$

On peut alors déduire de la formule (2) que quand $n \rightarrow \infty$,

$$\begin{aligned} P_n(r_n) &= \frac{r_n - 1}{n} \times \sum_{i=r_n-1}^{n-1} \frac{1}{i} \\ &\approx \frac{r_n - 1}{n} \times \ln\left(\frac{n-1}{r_n-2}\right). \end{aligned}$$

Comme -1 et -2 deviennent négligeables lorsque n est grand, on peut écrire

$$P_n(r_n) \approx \frac{r_n}{n} \times \ln\left(\frac{n}{r_n}\right).$$

Puisque $r_n/n \rightarrow t$ quand $n \rightarrow \infty$, il s'ensuit que

$$P_n(r_n) \rightarrow P(t) = -t \ln(t).$$

La figure 3 montre une densité de probabilité continue f dont toute la masse est concentrée sur l'intervalle $[0, 1]$. L'aire sous la courbe est égale à l'unité. La partie en bleu correspond au quantile d'ordre $1/e \approx 37\%$ de cette loi. Dans cet exemple précis, le quantile d'ordre 37% est d'environ 0,63.

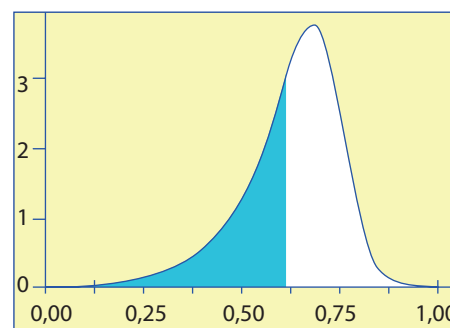


Figure 3. Une densité de probabilité sur l'intervalle $[0, 1]$ dont le nombre 0,628 est le 37^e centile. L'aire tracée en bleu, à gauche de 0,628, représente 37% de l'aire sous la courbe.

En général, soit τ le quantile d'ordre $1/e$ de la loi continue f , c'est-à-dire le plus petit nombre $\tau \in [0, T]$ à la gauche duquel se situe une proportion de $1/e$ de la masse de probabilité :

$$\int_0^{\tau} f(x) dx = 1/e.$$

Considérons la stratégie consistant à rejeter d'emblée toutes les candidatures reçues dans l'intervalle $[0, \tau]$ et à offrir ensuite le poste à la première personne éligible, c'est-à-dire qui se classe mieux que toutes celles qui la précèdent (en supposant qu'une telle personne se présente). Dans ces conditions, Bruss est parvenu à démontrer les faits suivants :

- La probabilité de recruter ainsi la meilleure candidature qui se présente dans l'intervalle $[0, T]$ est supérieure ou égale à $1/e$, peu importe f et la loi du nombre N de candidatures.
- La probabilité que cette stratégie ne permette d'embaucher personne est aussi de $1/e$.

Ce résultat, publié en 1984 dans les prestigieuses *Annals of Probability*, a généré beaucoup d'émoi dans la communauté mathématique, du fait que $1/e$ constitue ici une borne inférieure !

Et quoi encore ?

Il existe de nombreuses autres variantes du problème. Citons notamment :

- La possibilité d'exploiter non pas seulement un classement relatif mais plutôt la valeur intrinsèque des candidatures, telle que le prix demandé dans la recherche d'une propriété.
- La possibilité de revenir en arrière et d'offrir le poste à une personne déjà interviewée.
- La possibilité que certaines personnes refusent une offre.

On pourrait aussi imaginer qu'il y ait plus d'un poste à combler, ou inversement que l'on puisse sélectionner plus d'une personne pour un seul poste et que l'on considère l'opération comme réussie dès lors que la meilleure candidature figure dans le lot.

Une autre option consisterait à chercher dès le départ à recruter, non pas la meilleure personne possible, mais plutôt la seconde meilleure. Cet objectif pourrait être motivé, entre autres, par la crainte que la meilleure candidature reçoive plusieurs offres et préfère aller ailleurs.

Dans ce dernier cas, on peut montrer que la meilleure stratégie consiste à rejeter la première moitié des candidatures et à offrir le poste à la première personne qui se qualifie ensuite comme meilleure deuxième. La probabilité de succès associée à cette stratégie est d'environ $\frac{1}{4}$. Curieusement, il appert qu'il est plus facile de choisir la meilleure candidature que la seconde en lice !

Conclusion

S'il faut tirer une leçon de toutes ces considérations, c'est sans doute qu'en amour ou en affaires, il est généralement sage de voir venir et de prendre le temps d'explorer un peu le marché avant de s'engager... Gare, donc, aux décisions hâtives !

Choix d'une épouse

Dans un article de synthèse sur la règle des 37% paru en 1989 dans la revue *Statistical Science*, le mathématicien américain Tom Ferguson relate comment le célèbre astronome allemand Johannes Kepler (1571-1630) s'y est pris pour choisir sa seconde épouse, après que la première ait été emportée par le choléra en 1611.

Dans un processus élaboré qu'il a décrit dans une lettre adressée au Baron Strahlendorf en octobre 1613, Kepler raconte avoir considéré pas moins de 11 candidates. Après avoir pris en compte toutes sortes de variables (qualités et défauts de ces dames, existence ou taille de la dot, négociations avec les parents, conseils des amis, etc.), Kepler a finalement opté pour la cinquième.

On peut remarquer facétieusement que $r_{11} = 5$. Cependant, Kepler ne s'en est pas tenu exactement aux règles du problème tel qu'énoncé ici. En particulier, bien qu'il ait été fortement attiré par la candidate n° 5 après l'avoir rencontrée, il a ensuite fait une proposition de mariage à la candidate n° 4 sur l'avis de ses amis. Cette offre ne s'est toutefois pas concrétisée.

Pour de plus amples détails, consulter l'article de Ferguson ou la biographie de Kepler réalisée par Arthur Koestler parue en 1960 chez Anchor Books (Garden City, NY).

Dédicace

Je souhaite dédicacer cet article à Thomas Bruss et Tom Ferguson, que j'ai côtoyés au quotidien en 1997-98 à la faveur d'une année d'étude et de recherche à l'Université libre de Bruxelles. Ils m'ont beaucoup appris et je les considère comme des amis. Cet article s'inspire énormément de leurs travaux.