

par le dé est supérieure à celle attendue dans le cas où vous continuez à lancer le dé après avoir rejeté le premier lancer. Cela vous placerait dans un nouveau jeu où vous n'avez droit qu'à quatre lancers, dont l'espérance mathématique est également inconnue au départ. La stratégie optimale dans un problème à quatre lancers, à son tour, est de s'arrêter au premier lancer si la valeur obtenue est supérieure au gain espéré en continuant avec un problème à trois lancers, et ainsi de suite. On aboutit au problème à un lancer, pour lequel il n'y a qu'une stratégie, à savoir s'arrêter, et la récompense attendue est l'espérance mathématique d'un unique lancer de dé, dont nous avons vu qu'elle vaut 3,5 (voir la figure 6).

Cette information nous donne à son tour la stratégie optimale d'un problème à deux lancers : s'arrêter au premier lancer si la valeur est supérieure à ce que vous espérez gagner en continuant, c'est-à-dire supérieure à 3,5. Nous connaissons donc maintenant la stratégie optimale pour un

problème à deux lancers (s'arrêter au premier lancer si c'est un 4, un 5 ou un 6, continuer sinon) et cela permet de calculer la récompense espérée.

Dans un problème à deux lancers, on gagne 4, 5 ou 6 au premier lancer avec une probabilité de $1/6$ chacun, et l'on s'arrête. Sinon (quand le premier lancer est un 1, un 2 ou un 3, situation qui se produit une fois sur deux), on continue, auquel cas on s'attend à gagner 3,5 en moyenne. Ainsi, la récompense attendue pour le problème à deux lancers est :

$$4(1/6) + 5(1/6) + 6(1/6) + (1/2)(3,5) = 4,25.$$

Stratégie optimale

Cela fournit maintenant la stratégie optimale pour un problème à trois lancers, à savoir s'arrêter si le premier lancer est un 5 ou un 6 (c'est-à-dire supérieur à 4,25), et sinon continuer et ne s'arrêter que si le second lancer est un 4, 5 ou 6, et dans le cas contraire poursuivre avec un troisième

et dernier lancer. La connaissance de cette récompense attendue pour trois lancers donne à son tour la stratégie optimale pour un problème à quatre lancers : ne s'arrêter au premier lancer que s'il donne un 5 ou un 6, s'arrêter au second lancer si c'est un 5 ou un 6, au troisième si c'est un 5 ou un 6, au quatrième si c'est un 4, un 5 ou un 6, et sinon aller jusqu'au lancer final. Cette stratégie garantit que vous gagnerez 5,12 Krugerrands en moyenne, et aucune stratégie n'est meilleure.

La méthode de la récurrence à rebours est polyvalente, et fonctionne aussi bien lorsque les valeurs du processus ne sont pas indépendantes (comme elles sont supposées l'être lorsqu'on lance plusieurs fois le dé), ou s'il s'agit de minimiser une espérance mathématique telle que le coût.

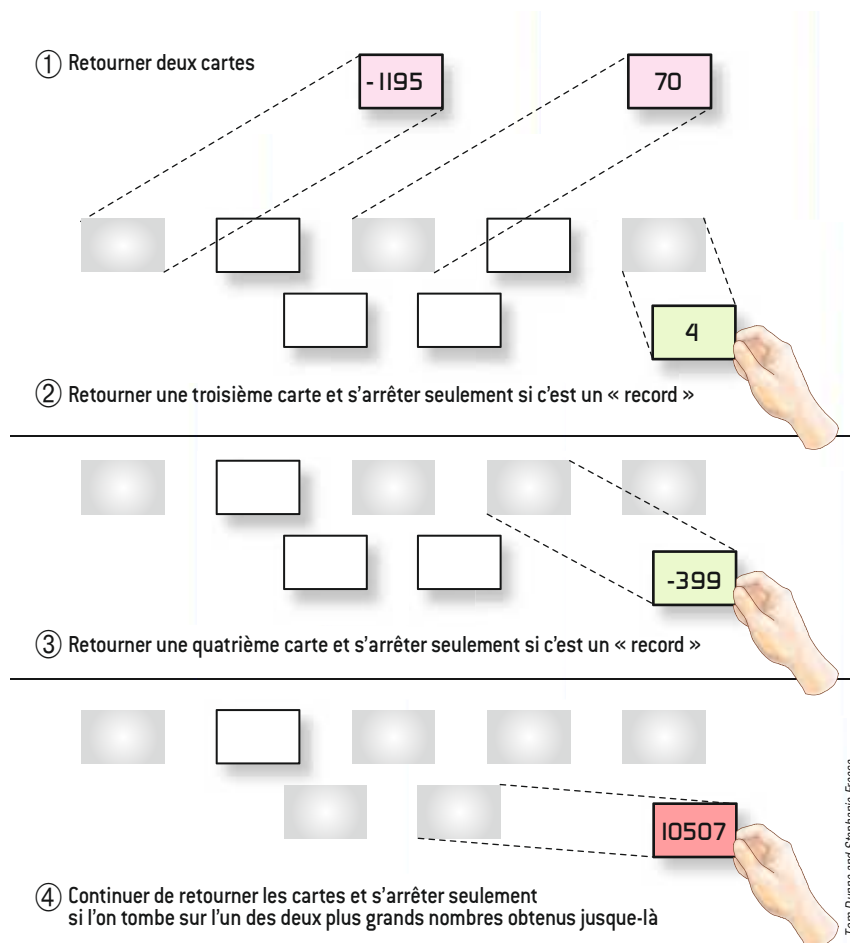
Supposons qu'une entreprise doive acheter son approvisionnement hebdomadaire en énergie le lundi, mardi ou mercredi précédent et que la probabilité des prix à venir puisse être estimée sur la base de statistiques passées. Par exemple, si le lundi l'acheteur a l'occasion de payer l'énergie pour la semaine suivante à un prix de 100, il peut savoir d'expérience qu'il y a une probabilité $1/2$ pour que le prix du mardi soit 110, et 90 sinon. De plus, il sait que si le prix est égal à 110 le mardi, alors le prix sera de 115 le mercredi avec une probabilité égale à $1/3$, et de 100 autrement ; et que s'il vaut 90 le mardi, il a autant de chances de valoir 100 que 85 le mercredi.

Par récurrence à rebours, on voit que la règle optimale pour le mardi est de ne pas acheter si le prix est de 110, puisque 110 est supérieur à l'espérance mathématique du prix s'il attend mercredi pour acheter, espérance égale à :

$$(1/3)(115) + (2/3)(100) = 105.$$

De la même façon, si le prix mardi est de 90, il est optimal d'acheter. En remontant jusqu'au lundi, puisque 100 est supérieur à l'espérance de prix s'il continue (à savoir $(1/2)(105) + (1/2)(90) = 97,5$), il est optimal de ne pas acheter le lundi. Ces éléments réunis constituent la stratégie optimale d'arrêt.

Dans le cas où l'information est complète, avec pour objectif de s'arrêter sur l'une des k plus grandes valeurs, on ignoreait quelles étaient les meilleures probabilités possibles de gagner pour des séquences finies générales de variables aléatoires indépendantes. En utilisant à la fois la récurrence normale et celle à rebours, et une classe de distributions



3. S'IL S'AGIT DE S'ARRÊTER dès l'obtention d'une des k meilleures valeurs, on peut là aussi augmenter ses chances. Dans un scénario où l'on désire obtenir un des deux meilleurs choix sur sept, on peut gagner deux fois sur trois avec la stratégie indiquée.