

ESSEC - 2009 - VOIE E

(adapté pour Scilab)

Correction proposée par G. Dupont.

Ce sujet comporte trois problèmes de décision inspirés de situations concrètes.
Ces problèmes sont indépendants.

Problème 1 : prédire le dernier succès

Présentation : soit un entier $n \geq 1$. On répète n fois, de façon indépendante, une même expérience qui conduit à un succès avec la probabilité $p \in]0, 1[$ ou à un échec avec la probabilité $1 - p$.

Le jeu proposé est de deviner quand aura lieu le dernier succès. A chaque succès, on peut décider d'annoncer ou non qu'il s'agit du dernier de toute la série d'expériences. On ne peut faire qu'une annonce par partie.

Le jeu est gagné si, à l'issue des n expériences, on a fait une annonce et qu'elle s'est révélée exacte. Le jeu est perdu si l'on n'a pas fait d'annonce ou si l'on s'est trompé en annonçant le dernier succès.

Stratégie : on choisit un entier $s \in \llbracket 1, n \rrbracket$, et on laisse passer les $n - s$ premières expériences. Ensuite, dès qu'un succès se présente, on annonce que ce sera le dernier.

On note P_s la probabilité de gagner en utilisant cette stratégie.

1. Montrer que la stratégie est gagnante si et seulement si il y a exactement un succès lors des s dernières expériences.
2. En déduire une expression de P_s en fonction de p et de s .
3. Montrer l'équivalence : $\frac{P_{s+1}}{P_s} \geq 1 \Leftrightarrow s \leq \frac{1}{p} - 1$.
4. En déduire que la probabilité P_s est maximale pour une ou deux valeurs de s .
5. Un exemple : on lance 10 fois un dé bien équilibré, et on doit prédire quand survient le dernier six. Quel choix convient-il de faire ?

Solution :

1. S'il n'y a aucun succès après l'annonce, la stratégie est perdante. S'il y a deux succès ou plus, cela signifie que l'on va annoncer qu'il va y avoir au moins un succès après celui que l'on aura annoncé comme le dernier, et la stratégie est perdante. Enfin, s'il y a un seul succès, on l'annoncera comme étant le dernier et il n'y en aura effectivement aucun autre derrière, la stratégie est donc gagnante dans ce cas.
2. Notons, pour tout $i \in \llbracket 1, n \rrbracket$, l'événement S_i : « faire un succès à la i -ème épreuve ». Si on note G_s l'événement correspondant gagner à la stratégie choisie, on a d'après la question précédente,

$$G_s = \bigsqcup_{i=n-s+1}^n \left(S_i \cap \left(\bigcap_{\substack{j=n-s+1 \\ j \neq i}}^n \bar{S}_j \right) \right).$$

Ainsi,

$$\begin{aligned} P_s &= \mathbf{P}(G_s) \\ &= \mathbf{P}\left(\bigsqcup_{i=n-s+1}^n \left(S_i \cap \left(\bigcap_{\substack{j=n-s+1 \\ j \neq i}}^n \bar{S}_j\right)\right)\right) \\ &= \sum_{i=n-s+1}^n \left(\mathbf{P}(S_i) \times \prod_{\substack{j=n-s+1 \\ j \neq i}}^n \mathbf{P}(\bar{S}_j)\right) \\ &= \sum_{i=n-s+1}^n p(1-p)^{s-1} \\ &= sp(1-p)^{s-1}. \end{aligned}$$

Ainsi,

$$P_s = sp(1-p)^{s-1}.$$

3. On a

$$\frac{P_{s+1}}{P_s} = \frac{(s+1)p(1-p)^s}{sp(1-p)^{s-1}} = \frac{s+1}{s}(1-p).$$

Ainsi,

$$\begin{aligned} \frac{P_{s+1}}{P_s} \geq 1 &\Leftrightarrow \frac{s+1}{s}(1-p) \geq 1 \\ &\Leftrightarrow (s+1)(1-p) \geq s \\ &\Leftrightarrow sp \leq 1-p \\ &\Leftrightarrow s \leq \frac{1}{p} - 1. \end{aligned}$$

En conclusion,

$$\frac{P_{s+1}}{P_s} \geq 1 \Leftrightarrow s \leq \frac{1}{p} - 1.$$

4. Pour tout $s \in \llbracket 1, n \rrbracket$, $P_s > 0$ donc

$$\frac{P_{s+1}}{P_s} \geq 1 \Leftrightarrow P_{s+1} \geq P_s.$$

Ainsi, tant que $s \leq \frac{1}{p} - 1$, on a $P_{s+1} \geq P_s$, puis après on a $P_{s+1} \leq P_s$, avec des inégalités strictes quand la première inégalité l'est. On a donc trois cas :

- Si $n \leq \frac{1}{p} - 1$, cela signifie que P_n est maximale.
- Si $\frac{1}{p}$ est entier, on a

$$P_{\frac{1}{p}} = P_{\frac{1}{p}-1}$$

et toutes les autres valeurs de P_s sont strictement inférieures.

- Enfin, si $\frac{1}{p}$ n'est pas un entier, alors la valeur maximale est $P_{\lfloor \frac{1}{p} \rfloor}$.

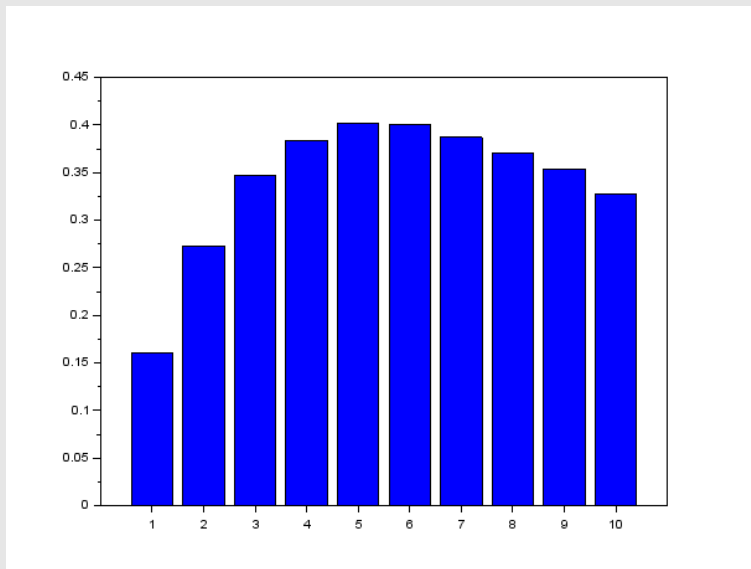
5. On est dans le cas où $n = 10$ et $p = \frac{1}{6}$ de sorte que $\frac{1}{p} - 1 = 5 \leq 10$. Ainsi, P_5 et P_6 sont maximales. Autrement dit, la stratégie optimale consiste à laisser passer 4 ou 5 lancers et à annoncer le premier succès suivant comme étant le dernier.

Remarque : Les scripts Scilab ci-dessous permettent de simuler ces expériences et de calculer les fréquences de succès des différentes stratégies.

```
function G = strategies(n,p)
L = listerresultats(n,p)
G = zeros(2,n)
R = zeros(1,n)
for s=1:n
    G(1,s) = s
    R(s) = sum(L(n-s+1:n))
    if (R(s)==1)
        G(2,s) = 1
    else
        G(2,s) = 0
    end
end
endfunction

function P = Pstrategies(n,p,N)
P = zeros(2,n)
for i =1:n
    P(1,i) = i
end
for i = 1:N
    G = strategies(n,p)
    for i=1:n
        if (G(2,i) == 1)
            P(2,i) = P(2,i)+1/N
        end
    end
end
end
bar(P(1,1:n),P(2,1:n))
endfunction
```

La commande `Pstrategies(10,1/6,10000)` a donné le graphe suivant, en accord avec la réponse à la question 5.



□

Problème 2 : chercher une place de parking

Présentation : on est en voiture au départ d'une rue infiniment longue et à sens unique. On doit se rendre à un point d'arrivée située à une certaine distance du point de départ et on cherche à se garer le plus près possible de l'arrivée. À partir d'où doit-on commencer à accepter une place libre ?

Mise en place : au départ on est au numéro 0 de la rue. Pour chaque entier naturel n , il y a une place de parking au numéro n , qui peut être libre avec la probabilité $p \in]0, 1[$. On suppose que p ne dépend pas de n et que les occupations de places se font indépendamment les unes des autres. L'arrivée est au numéro d .

Stratégie : on se donne $s \in \llbracket 0, d \rrbracket$, et on conduit sans s'arrêter jusqu'au numéro s de la rue. On accepte alors la première place libre à partir du numéro s (inclus).

On note X le numéro de la place trouvée par cette méthode. La distance à l'arrivée est $|X - d|$ et l'espérance $D_s = \mathbf{E}[|X - d|]$ est la distance moyenne à l'arrivée.

1. Loi de X .

- Déterminer l'univers-image $X(\Omega)$.
- Pour tout $k \in \mathbf{N}$, on note A_k l'événement « la place au numéro k est occupée ». Pour $n \in X(\Omega)$, exprimer l'événement $(X = n)$ en fonction des événements A_k .
- Déterminer la loi de X .
- Vérifier que $X - s + 1$ suit une loi géométrique.
- En déduire l'espérance de X .

2. Calcul de $D_s = \mathbf{E}[|X - d|]$.

- Montrer que la variable aléatoire $|X - d|$ admet une espérance.

- Établir :
$$D_s = \sum_{n=s}^{+\infty} (n - d) \mathbf{P}(X = n) - 2 \sum_{n=s}^d (n - d) \mathbf{P}(X = n).$$

- Soit $x \in \mathbf{R} \setminus \{1\}$, donner la valeur de la somme $\sum_{k=0}^N x^k$ en fonction de N et x , et en déduire une expression de

la somme $\sum_{k=0}^N kx^k$.

- En déduire :
$$\sum_{n=s}^d (n - d) \mathbf{P}(X = n) = \frac{1}{p} + s - d - 1 - \frac{(1 - p)^{d-s+1}}{p}.$$

- Montrer finalement :
$$D_s = d - s + 1 - \frac{1}{p} + \frac{2}{p}(1 - p)^{d-s+1}.$$

3. Optimisation.

- Simplifier $D_{s+1} - D_s$ et en déduire que D_s est minimale pour s le plus petit entier strictement supérieur à $\sigma_p = d + \frac{\ln 2}{\ln(1 - p)}$.

- Montrer que si $p \geq \frac{1}{2}$, D_s est minimale pour $s = d$.

4. Exemple : il y a en moyenne 1 place sur 10 de libre, à quelle distance de l'arrivée doit-on commencer à chercher une place ?

On utilisera l'encadrement suivant : $2^{-\frac{1}{6}} < 0,9 < 2^{-\frac{1}{7}}$.

5. Simulation informatique.

L'algorithme ci-dessous permet de simuler la recherche de place.

```
p = input('Probabilité de place libre ? ')
x = rand()
-----
while(-----)
    x = rand()
    k = k+1
end
disp(k,'Place trouvée: ')
disp(abs(k-d),'Distance: ')
```

(a) Laquelle de ces instructions manque à la troisième ligne ?

i. $k = s-1$

ii. $k = s$

iii. $k = s+1$

(b) Compléter la quatrième ligne.

Solution :

1. Loi de X .

(a) X peut prendre toutes les valeurs entières à partir de s . Ainsi,

$$X(\Omega) = \llbracket s, +\infty \llbracket.$$

(b) Pour $n \geq s$, $(X = n)$ signifie que toutes les places entre s et $n-1$ étaient prises et que la place n était libre. Autrement dit,

$$(X = n) = \left(\bigcap_{k=s}^{n-1} A_k \right) \cap \overline{A_n}.$$

(c) On a $X(\Omega) = \llbracket s, +\infty \llbracket$ et, pour tout $n \geq s$,

$$\begin{aligned} \mathbf{P}(X = n) &= \mathbf{P} \left(\left(\bigcap_{k=s}^{n-1} A_k \right) \cap \overline{A_n} \right) \\ &= \left(\prod_{k=s}^{n-1} \mathbf{P}(A_k) \right) \times \mathbf{P}(\overline{A_n}) \quad (\text{par indépendance}) \\ &= (1-p)^{n-s} p. \end{aligned}$$

Ainsi,

$$\forall n \geq s, \quad \mathbf{P}(X = n) = (1-p)^{n-s} p.$$

(d) On a $(X - s + 1)(\Omega) = \mathbf{N}^*$ et, pour tout $k \in \mathbf{N}^*$,

$$\begin{aligned} \mathbf{P}(X - s + 1 = k) &= \mathbf{P}(X = k + s - 1) \\ &= (1-p)^{k+s-1-s} p \\ &= (1-p)^{k-1} p. \end{aligned}$$

Ainsi,

$$X - s + 1 \rightsquigarrow \mathcal{G}(p).$$

(e) On a

$$\begin{aligned} \mathbf{E}[X] &= \mathbf{E}[X - s + 1 + s - 1] \\ &= \mathbf{E}[X - s + 1] + s - 1 \\ &= \frac{1}{p} + s - 1 \quad (\text{car } X - s + 1 \rightsquigarrow \mathcal{G}(p)). \end{aligned}$$

Ainsi,

$$\mathbf{E}[X] = \frac{1}{p} - s + 1.$$

2. Calcul de $D_s = \mathbf{E}[|X - d|]$.

(a) On a $|X - d| \leq |X| + d$ et $|X|$ admet une espérance, donc $|X - d|$ aussi.

(b) On a

$$\begin{aligned} \sum_{n=s}^{+\infty} (n-d)\mathbf{P}(X=n) - 2 \sum_{n=s}^d (n-d)\mathbf{P}(X=n) &= \sum_{n=d+1}^{+\infty} (n-d)\mathbf{P}(X=n) - \sum_{n=s}^d (n-d)\mathbf{P}(X=n) \\ &= \sum_{n=d+1}^{+\infty} (n-d)\mathbf{P}(X=n) + \sum_{n=s}^d (d-n)\mathbf{P}(X=n) \\ &= \sum_{n=d+1}^{+\infty} |n-d|\mathbf{P}(X=n) + \sum_{n=s}^d |n-d|\mathbf{P}(X=n) \\ &= \sum_{n=s}^{+\infty} |n-d|\mathbf{P}(X=n) \\ &= \mathbf{E}[|X-d|] \quad (\text{théorème de transfert}). \end{aligned}$$

Ainsi,

$$\mathbf{E}[|X-d|] = \sum_{n=s}^{+\infty} (n-d)\mathbf{P}(X=n) - 2 \sum_{n=s}^d (n-d)\mathbf{P}(X=n).$$

(c) Soit $x \in \mathbf{R} \setminus \{1\}$. On a

$$\sum_{k=0}^N x^k = \frac{1-x^{N+1}}{1-x}.$$

Ainsi,

$$\begin{aligned} \sum_{k=0}^N kx^k &= x \sum_{k=0}^N kx^{k-1} \\ &= x \left(\sum_{k=0}^N x^k \right)' \\ &= x \left(\frac{1-x^{N+1}}{1-x} \right)' \\ &= \frac{x}{(1-x)^2} (-(N+1)x^N(1-x) + (1-x^{N+1})) \\ &= \frac{x - (N+1)x^{N+1} + Nx^{N+2}}{(1-x)^2}. \end{aligned}$$

On obtient donc

$$\sum_{k=0}^N kx^k = \frac{x - (N+1)x^{N+1} + Nx^{N+2}}{(1-x)^2}.$$

(d) Ici, il faut prendre son courage à deux mains pour trouver la forme demandée même s'il n'y a pas de réelle difficulté. Procédons méthodiquement. Commençons par poser $q = 1 - p$.

$$\begin{aligned}
\sum_{n=s}^d (n-d)\mathbf{P}(X=n) &= \sum_{n=s}^d (n-s+s-d)\mathbf{P}(X=n) \\
&= \sum_{n=s}^d ((n-s) + (s-d))q^{n-s}p \\
&= \sum_{n=s}^d (n-s)q^{n-s}p + \sum_{n=s}^d (s-d)q^{n-s}p \\
&= p \sum_{n=s}^d (n-s)q^{n-s} + p(s-d) \sum_{n=s}^d q^{n-s} \\
&= p \sum_{k=0}^{d-s} kq^k + p(s-d) \sum_{k=0}^{d-s} q^k \quad (\text{en posant } k = n-s) \\
&= p \sum_{k=0}^N kq^k + pN \sum_{k=0}^N q^k \quad (\text{en posant } N = d-s \text{ pour alléger}) \\
&= p \frac{q - (N+1)q^{N+1} + Nq^{N+2}}{(1-q)^2} - pN \frac{1-q^{N+1}}{1-q} \quad (\text{d'après 2.c}) \\
&= \frac{q - (N+1)q^{N+1} + Nq^{N+2}}{p} - N + Nq^{N+1} \quad (\text{car } p = 1-q) \\
&= -N + \frac{1}{p} (q - Nq^{N+1} - q^{N+1} + Nq^{N+2} + Npq^{N+1}) \\
&= -N + \frac{1}{p} (Nq^{N+1}(p-1) + q - q^{N+1} + Nq^{N+2}) \\
&= -N + \frac{1}{p} (Nq^{N+2} + q - q^{N+1} + Nq^{N+2}) \quad (\text{car } p = 1-q) \\
&= -N + \frac{1}{p} (1-p - (1-p)^{N+1}) \quad (\text{car } p = 1-q) \\
&= -N + \frac{1}{p} - 1 - \frac{(1-p)^{N+1}}{p} \\
&= \frac{1}{p} + s - d - 1 - \frac{(1-p)^{d-s+1}}{p} \quad (\text{car } N = d-s).
\end{aligned}$$

Ainsi,

$$\boxed{\sum_{n=s}^d (n-d)\mathbf{P}(X=n) = \frac{1}{p} + s - d - 1 - \frac{(1-p)^{d-s+1}}{p}.}$$

(e) Reprenons la formule établie en 2.b et injectons le résultat de 2.d. On obtient :

$$\begin{aligned}
D_s &= \sum_{n=s}^{+\infty} (n-d)\mathbf{P}(X=n) - 2 \sum_{n=s}^d (n-d)\mathbf{P}(X=n) \\
&= \sum_{n=s}^{+\infty} (n-d)\mathbf{P}(X=n) - 2 \left(\frac{1}{p} + s - d - 1 - \frac{(1-p)^{d-s+1}}{p} \right).
\end{aligned}$$

Par ailleurs,

$$\begin{aligned}
\sum_{n=s}^{+\infty} (n-d)\mathbf{P}(X=n) &= \sum_{k=1}^{+\infty} (k+s-d-1)\mathbf{P}(X=k+s-1) \quad (\text{en posant } k=n-s+1) \\
&= \sum_{k=1}^{+\infty} (k+s-d-1)\mathbf{P}(X-s+1=k) \\
&= \sum_{k=1}^{+\infty} k\mathbf{P}(X-s+1=k) + (s-d-1) \sum_{k=1}^{+\infty} \mathbf{P}(X-s+1=k) \\
&= \frac{1}{p} + (s-d-1) \quad (\text{car } X-s+1 \rightsquigarrow \mathcal{G}(p)).
\end{aligned}$$

Ainsi,

$$\begin{aligned}
D_s &= \frac{1}{p} + (s-d-1) - 2 \left(\frac{1}{p} + s-d-1 - \frac{(1-p)^{d-s+1}}{p} \right) \\
&= \frac{1}{p} + s-d-1 - \frac{2}{p} - 2s + 2d + 2 + \frac{2}{p}(1-p)^{d-s+1} \\
&= -\frac{1}{p} - s + d + 1 + \frac{2}{p}(1-p)^{d-s+1}.
\end{aligned}$$

On a donc bien établi

$$D_s = -\frac{1}{p} - s + d + 1 + \frac{2}{p}(1-p)^{d-s+1}.$$

3. Optimisation.

(a) On a

$$\begin{aligned}
D_{s+1} - D_s &= -1 + \frac{2}{p}q^{d-s}(1-q) \\
&= 2q^{d-s} - 1.
\end{aligned}$$

Ainsi,

$$\begin{aligned}
D_{s+1} \leq D_s &\Leftrightarrow D_{s+1} - D_s \leq 0 \\
&\Leftrightarrow 2q^{d-s} - 1 \leq 0 \\
&\Leftrightarrow q^{d-s} \leq \frac{1}{2} \\
&\Leftrightarrow \exp((d-s)\ln(q)) \leq \frac{1}{2} \\
&\Leftrightarrow (d-s)\ln(q) \leq \ln\left(\frac{1}{2}\right) \\
&\Leftrightarrow (d-s) \geq -\frac{\ln 2}{\ln q} \\
&\Leftrightarrow s \leq d + \frac{\ln 2}{\ln q}.
\end{aligned}$$

Ainsi, D_{s+1} décroît tant que $s \leq d + \frac{\ln 2}{\ln q}$. Il s'ensuit que :

$$D_s \text{ est minimale quand } s \text{ est le plus petit entier strictement supérieur à } \sigma_p = d + \frac{\ln 2}{\ln q}.$$

- (b) Si $p \geq \frac{1}{2}$, on a $\sigma_p \geq d - 1$ de sorte que l'on aura un maximum pour $s \geq d$. Par ailleurs, $D_{s+1} - D_s = 2(1-p)^{d-s} - 1 \leq 2\frac{1}{2^{d-s}} - 1$ est négatif dès que $s > d$. Ainsi :

D_s est minimale quand $s = d$.

4. Dire qu'il y a en moyenne 1 place sur 10 de libre signifie que $p = \frac{1}{10}$ et donc

$$\sigma_p = d + \frac{\ln 2}{\ln(0,9)}.$$

Or on a

$$\begin{aligned} 2^{-\frac{1}{6}} < 0,9 < 2^{-\frac{1}{7}} &\Leftrightarrow \exp\left(-\frac{1}{6} \ln 2\right) < 0,9 < \exp\left(-\frac{1}{7} \ln 2\right) \\ &\Leftrightarrow -\frac{1}{6} \ln 2 < \ln(0,9) < -\frac{1}{7} \ln 2 \\ &\Leftrightarrow -\frac{1}{6} > \frac{\ln(0,9)}{\ln 2} > -\frac{1}{7} \\ &\Leftrightarrow -6 > \frac{\ln(2)}{\ln(0,9)} > -7 \\ &\Leftrightarrow -6 > \frac{\ln(2)}{\ln(0,9)} > -7 \end{aligned}$$

donc

$$d - 7 < \sigma_p < d - 6.$$

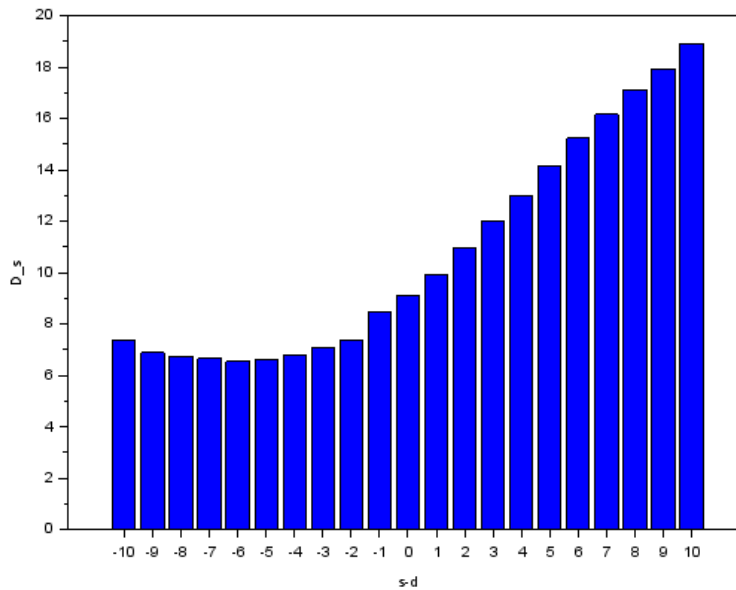
Il faut donc commencer à chercher à six places de l'arrivée.

5. Simulation informatique.

- (a) Il s'agit de mettre $k = s$ car si le test `while` est échoué, cela signifie que c'est la place occupée avant l'exécution du bloc correspondant qui sera retenue.
- (b) Le programme complet est donc :

```
p = input('Probabilité de place libre ? ')
x = rand()
k = s
while(x>p)
    x = rand()
    k = k+1
end
disp(k,'Place trouvée: ')
disp(abs(k-d),'Distance: ')
```

Remarque : Une simulation pour comparer chacune des stratégies sur des échantillons de taille 5000 avec $p = \frac{1}{10}$ a amené au graphique suivant, où les abscisses correspondent à $d - s$, c'est-à-dire là où on commence à chercher la place par rapport à l'arrivée, et les ordonnées correspondent à D_s , la distance moyenne à cette arrivée. On y observe que la distance est effectivement minimale quand on prend la sixième place avant l'arrivée.



Pour l'obtenir, on a exécuté `optimal(10,1/10,5000)` avec le script suivant (la valeur 10 étant arbitraire, elle correspond à là où l'on a situé l'arrivée) :

```
function r = distance(s,d,p)
    x = rand()
    k = s
    while (x>p)
        x = rand()
        k = k+1
    end
    r = abs(k-d)
endfunction

function r = strategie(s,d,p,N)
    X = zeros(1,N)
    for i=1:N
        X(i) = distance(s,d,p)
    end
    r = mean(X)
endfunction

function optimal(d,p,N)
    nmax = 2*d+1
    X = zeros(1,nmax)
    for i =1:nmax
        X(i) = strategie(i-1,d,p,N)
    end
    clf()
    x = 1:nmax
    x = x-d-1
    bar(x,X)
    xlabel('s-d')
    ylabel('D_s')
endfunction
```

□

Problème 3 : vendre par petites annonces

Présentation : On met en vente un objet dans les petites annonces d'un journal. On reçoit chaque jour une nouvelle offre (et une seule), que l'on peut accepter ou refuser. Cette décision est définitive : en cas de refus, on ne pourra plus accepter cette offre dans les jours qui suivent ; en cas d'acceptation, on gagne le montant de l'offre et la parution s'arrête. Le nombre d'offres est a priori illimité, mais le journal facture un coût $c > 0$ pour chaque jour de parution. Quand doit-on accepter l'offre proposée ?

Mise en place : On fait les hypothèses suivantes.

- Pour $k \in \mathbf{N}^*$, on note X_k l'offre du k -ième jour. Les variables X_k sont indépendantes et suivent toutes la même loi qu'une variable aléatoire X .
- X est à valeurs dans \mathbf{R}_+ , et admet une densité notée f . On notera F la fonction de répartition.
- X admet une espérance notée m .

On appelle N le numéro de l'offre acceptée, c'est une variable aléatoire à valeurs dans \mathbf{N}^* , et G le gain final que l'on tire de la vente.

On a ainsi $G = X_N - Nc$.

Stratégie : On se donne une valeur $s \in \mathbf{R}_+$, et on choisit d'accepter la première offre supérieure ou égale à s . On cherche une valeur de s qui maximise le gain moyen $\mathbf{E}[G]$.

1. Expliquer pourquoi on peut supposer que s est tel que $F(s) \in [0, 1[$.
Cette condition sera vérifiée dans toute la suite du sujet.
2. Calcul de l'espérance de G .
 - (a) Justifier que N suit une loi géométrique dont on exprimera le paramètre en fonction de $F(s)$.
Donner l'espérance de N .
 - (b) Justifier : $\mathbf{P}(X_N < s) = 0$.
 - (c) Soit $x \geq s$.

i. Justifier : pour tout $n \in \mathbf{N}^*$, $(X_N > x) \cap (N = n) = (X_n > x) \cap \left(\bigcap_{k=1}^{n-1} (X_k < s) \right)$.

ii. En déduire : $\mathbf{P}(X_N \leq x) = \frac{F(x) - F(s)}{1 - F(s)}$.

(d) Déterminer une densité de X_N .

(e) Montrer que X_N admet une espérance.

(f) Montrer que G admet une espérance, donnée par $\mathbf{E}[G] = \frac{1}{1 - F(s)} \left(\int_s^{+\infty} x f(x) dx - c \right)$.

3. Optimisation.

On pose $g(s) = \mathbf{E}[G]$.

(a) Montrer que $\lim_{s \rightarrow +\infty} g(s) = -\infty$ et interpréter ce résultat.

(b) Que vaut $g(0)$? Interpréter la valeur trouvée.

(c) Montrer que si $c \geq m$, alors $g(s) \leq 0$ pour toute valeur de s .

On suppose dans toute la suite que $c < m$.

(d) Montrer que g est dérivable et mettre sa dérivée sous la forme $g'(s) = \frac{f(s)h(s)}{(1 - F(s))^2}$, où h est une fonction à préciser.

(e) Montrer que h est décroissante sur \mathbf{R}_+ .

(f) Montrer que $h(s)$ est négatif pour s suffisamment grand.

(g) En déduire que h s'annule au moins une fois sur \mathbf{R}_+ .

(h) Soit σ un réel positif tel que $h(\sigma) = 0$.

i. Montrer que g est maximale en σ .

ii. Montrer que $g(\sigma) = \sigma$.

iii. En déduire l'unicité de σ .

4. Variations en fonction de c .

L'espérance de G dépend en fait de s et de c . On la note dorénavant $g(s, c)$. La question précédente prouve qu'à c fixé, $g(s, c)$ est maximale pour une valeur unique que l'on note maintenant σ_c , et qui vérifie $g(\sigma_c, c) = \sigma_c$.

(a) Soit c et c' deux réels positifs tels que $c \leq c'$.

Vérifier que $g(s, c) \geq g(s, c')$ pour tout s .

(b) En déduire $\sigma_c \geq \sigma_{c'}$.

(c) La fonction $c \mapsto \sigma_c$ est ainsi décroissante. Ce résultat était-il prévisible?

5. Un exemple : la loi uniforme.

On suppose que X suit la loi uniforme sur un intervalle $[a, b]$ avec a et b réels positifs.

(a) Calculer $g(s, c)$.

(b) Montrer qu'à c fixé, $\mathbf{E}[G]$ est maximale lorsque $s = b - \sqrt{2(b-a)c}$.

6. Simulation informatique.

L'algorithme ci-dessous propose d'expérimenter la stratégie dans le cas où X suit la loi uniforme sur $[a, b]$. Compléter les instructions manquantes.

```
n = 0
a = input('Valeur de a:')
b = input('Valeur de b:')
x = rand()
y = -----
while (y < s)
    x = rand()
    y = -----
    -----
end
disp(y-n*c, 'Gain : ')
```

Solution :

1. On a $F(s) = \mathbf{P}(X \leq s) \in [0, 1]$. Si $F(s) = 1$, on a alors $\mathbf{P}(X \geq s) = \mathbf{P}(X > s) = 1 - F(s) = 0$, ce qui implique que l'on n'acceptera aucune offre.

On peut donc supposer $F(s) \in [0, 1[$.

2. (a) N correspond au rang du premier succès (=offre supérieure au seuil s) dans une répétition indéfinie d'expériences de Bernoulli indépendantes de paramètre $p = \mathbf{P}(X \geq s) = 1 - F(s)$. Il s'ensuit que N suit une loi géométrique de paramètre $1 - F(s)$.

Ainsi,

$$N \rightsquigarrow \mathcal{G}(1 - F(s)) \quad \text{et} \quad \mathbf{E}[N] = \frac{1}{1 - F(s)}.$$

(b) X_N est le montant de l'offre acceptée. Puisque l'on n'accepte que des offres supérieures ou égales à s , on a $\mathbf{P}(X_N \geq s) = 1$ et donc :

$$\mathbf{P}(X_N < s) = 0.$$

(c) Soit $x \geq s$.

i. Pour tout $n \in \mathbf{N}^*$, $(X_N > x) \cap (N = n)$ signifie que l'offre acceptée est l'offre numéro n et que cette offre est supérieure à x . En d'autres termes, toutes les offres jusqu'à la $(n - 1)$ -ième ont été refusées (elles étaient inférieures au seuil s) et la n -ième était supérieure à x . Mathématiquement, cela s'écrit

$$(X_N > x) \cap (N = n) = (X_n > x) \cap \left(\bigcap_{k=1}^{n-1} (X_k < s) \right).$$

ii. On a

$$\mathbf{P}(X_N \leq x) = 1 - \mathbf{P}(X_N > x)$$

et, d'après la formule des probabilités totales appliquée au système complet d'événements $\{(N = n)\}_{n \in \mathbf{N}^*}$, on a

$$\begin{aligned} \mathbf{P}(X_N > x) &= \sum_{n=1}^{+\infty} \mathbf{P}((X_N > x) \cap (N = n)) \\ &\stackrel{(2.c.i)}{=} \sum_{n=1}^{+\infty} \mathbf{P}\left((X_n > x) \cap \left(\bigcap_{k=1}^{n-1} (X_k < s)\right)\right) \\ &= \sum_{n=1}^{+\infty} \mathbf{P}(X_n > x) \prod_{k=1}^{n-1} \mathbf{P}(X_k < s) \quad (\text{par indépendance}) \\ &= \sum_{n=1}^{+\infty} (1 - F(x)) \prod_{k=1}^{n-1} F(s) \\ &= (1 - F(x)) \sum_{n=1}^{+\infty} F(s)^{n-1} \\ &= \frac{1 - F(x)}{1 - F(s)}. \end{aligned}$$

Ainsi,

$$\mathbf{P}(X_N \leq x) = 1 - \frac{1 - F(x)}{1 - F(s)} = \frac{F(x) - F(s)}{1 - F(s)}.$$

(d) Si on note F_{X_N} la fonction de répartition de X_N , on a

$$F_{X_N}(x) = \begin{cases} 0 & \text{si } x < s \\ \frac{F(x) - F(s)}{1 - F(s)} & \text{sinon.} \end{cases}$$

Puisque F est la fonction de répartition de X qui est à densité, elle est continue et \mathcal{C}^1 sauf éventuellement en un nombre fini de points. Il s'ensuit que c'est aussi le cas pour F_{X_N} et donc X_N admet pour densité la fonction f_{X_N} obtenue en dérivant F_{X_N} en tout point où elle est dérivable. Ainsi, on peut prendre pour densité

$$f_{X_N} : x \mapsto \begin{cases} 0 & \text{si } x < s \\ \frac{f(x)}{1 - F(s)} & \text{sinon.} \end{cases}$$

(e) On a

$$\begin{aligned} \int_{-\infty}^{+\infty} |x| f_{X_N}(x) dx &= \int_s^{+\infty} |x| \frac{f(x)}{1 - F(s)} dx \\ &= \frac{1}{1 - F(s)} \int_s^{+\infty} x f(x) dx \end{aligned}$$

mais cette dernière intégrale est convergente car X admet une espérance. Ainsi,

$$X_N \text{ admet une espérance et } \mathbf{E}[X_N] = \frac{1}{1-F(s)} \int_s^{+\infty} xf(x) dx.$$

- (f) On a $G = X_N - Nc$. Ainsi, G admet une espérance comme combinaison linéaire de variables aléatoires admettant une espérance. En outre

$$\begin{aligned} \mathbf{E}[G] &= \mathbf{E}[X_N - Nc] \\ &= \mathbf{E}[X_N] - c\mathbf{E}[N] \\ &\stackrel{(2.a)}{=} \frac{1}{1-F(s)} \int_s^{+\infty} xf(x) dx - \frac{c}{1-F(s)} \\ &\stackrel{(2.e)}{=} \frac{1}{1-F(s)} \left(\int_s^{+\infty} xf(x) dx - c \right). \end{aligned}$$

On a donc bien

$$\mathbf{E}[G] = \frac{1}{1-F(s)} \left(\int_s^{+\infty} xf(x) dx - c \right).$$

3. Optimisation.

On pose $g(s) = \mathbf{E}[G]$.

- (a) On a $\lim_{s:+\infty} F(s) = 1^-$ donc $\lim_{s:+\infty} \frac{1}{1-F(s)} = +\infty$. Par ailleurs

$$\int_s^{+\infty} xf(x) dx = \int_{-\infty}^{+\infty} xf(x) dx - \int_{-\infty}^s xf(x) dx = m - \int_{-\infty}^s xf(x) dx \xrightarrow{s:+\infty} m - m = 0$$

donc

$$\int_s^{+\infty} xf(x) dx - c \xrightarrow{s:+\infty} -c < 0$$

et donc

$$\lim_{s:+\infty} \frac{1}{1-F(s)} \left(\int_s^{+\infty} xf(x) dx - c \right) = -\infty.$$

- (b) On a

$$g(0) = \frac{1}{1-F(0)} \left(\int_0^{+\infty} xf(x) dx - c \right) = m - c.$$

En fait $g(0)$ correspond au gain moyen si on accepte n'importe quelle offre positive. En pratique cela signifie que l'on acceptera la première offre, on gagnera en moyenne l'offre moyenne (qui est m) minorée du prix de publication d'un jour (qui est c). On a donc bien :

$$g(0) = m - c.$$

- (c) Si $c \geq m$, on a

$$\int_s^{+\infty} xf(x) dx - c \leq \int_0^{+\infty} xf(x) dx - c = m - c \leq 0$$

et $\frac{1}{1-F(s)} > 0$ et donc

$$g(s) = \frac{1}{1 - F(s)} \left(\int_s^{+\infty} x f(x) dx - c \right) \leq 0.$$

- (d) g est dérivable comme quotient de fonctions dérivables dont le dénominateur ne s'annule pas et, pour tout $s \geq 0$, on a

$$\begin{aligned} g'(s) &= \frac{-sf(s)(1 - F(s)) + \left(\int_s^{+\infty} x f(x) dx - c \right) f(s)}{(1 - F(s))^2} \\ &= \frac{f(s)}{(1 - F(s))^2} \left(\int_s^{+\infty} x f(x) dx - c - s + sF(s) \right) \\ &= \frac{f(s)}{(1 - F(s))^2} h(s), \end{aligned}$$

avec

$$h(s) = \int_s^{+\infty} x f(x) dx - c - s + sF(s).$$

- (e) h est dérivable et on a, pour tout $s \geq 0$,

$$h'(s) = -sf(s) - 1 + F(s) + sf(s) = F(s) - 1 < 0$$

Ainsi,

h est strictement décroissante sur \mathbf{R}_+ .

- (f) On a vu en 3.a que $g(s)$ tend vers $-\infty$ quand s tend vers $+\infty$. Il s'ensuit que g ne peut pas être toujours croissante. Il existe donc $s_0 \in \mathbf{R}$ tel que $g'(s_0) < 0$. Mais

$$g'(s_0) = \frac{f(s_0)}{(1 - F(s_0))^2} h(s_0)$$

or $\frac{f(s_0)}{(1 - F(s_0))^2} > 0$ donc $h(s_0) < 0$. Par décroissance de h (3.e), il vient

$$\forall s \geq s_0, \quad h(s) < 0.$$

- (g) On a $h(0) = \int_0^{+\infty} x f(x) dx - c = m - c > 0$ et $h(s_0) < 0$. La fonction h étant continue, il suit du théorème des valeurs intermédiaires qu :

il existe $\sigma \in [0, s_0]$ tel que $h(\sigma) = 0$.

- (h) Soit σ un réel positif tel que $h(\sigma) = 0$.

- i. h étant décroissante, h est positive sur $[0, \sigma]$ et négative sur $[\sigma, 0]$. Il s'ensuit que g' est positive sur $[0, \sigma]$ et négative sur $[\sigma, 0]$. Ainsi :

g admet un maximum global en σ .

ii. On a

$$\begin{aligned}
 h(\sigma) = 0 &\Leftrightarrow \int_{\sigma}^{+\infty} xf(x) dx - c - \sigma + \sigma F(\sigma) = 0 \\
 &\Leftrightarrow \int_{\sigma}^{+\infty} xf(x) dx - c = \sigma(1 - F(\sigma)) \\
 &\Leftrightarrow \frac{1}{1 - F(\sigma)} \int_{\sigma}^{+\infty} xf(x) dx - c = \sigma \\
 &\Leftrightarrow g(\sigma) = \sigma.
 \end{aligned}$$

Ainsi,

$$g(\sigma) = \sigma.$$

iii. Si σ et τ sont des réels tels que $g(\sigma) = g(\tau) = 0$, il suit alors de 3.h.i que g admet un maximum en σ et en τ et, d'après les variations observées, que les valeurs de g en σ et en τ coïncident. On a donc $g(\sigma) = g(\tau)$. Or il suit de 3.h.ii que $g(\sigma) = \sigma$ et $g(\tau) = \tau$. En combinant ces deux égalités, on obtient $\sigma = \tau$. Autrement dit,

$$\exists! \sigma \in \mathbf{R}_+, \quad g(\sigma) = 0.$$

4. Variations en fonction de c .

(a) Supposons $c \leq c'$. Pour tout $s \in \mathbf{R}_+$, on a

$$\begin{aligned}
 g(s, c) - g(s, c') &= \frac{1}{1 - F(s)} \left(\int_s^{+\infty} xf(x) dx - c - \int_s^{+\infty} xf(x) dx + c' \right) \\
 &= \frac{c' - c}{1 - F(s)} \geq 0.
 \end{aligned}$$

Ainsi,

$$c \leq c' \quad \Rightarrow \quad g(s, c) \geq g(s, c').$$

(b) On a

$$\sigma_{c'} \stackrel{(3.h)}{=} g(\sigma_{c'}, c') \stackrel{(4.a)}{\leq} g(\sigma_{c'}, c) \stackrel{(3.h)}{\leq} g(\sigma_c, c) \stackrel{(3.h)}{=} \sigma_c.$$

Ainsi,

$$c \leq c' \quad \Rightarrow \quad \sigma_c \geq \sigma_{c'}.$$

(c) Ce résultat était prévisible puisqu'il signifie que le gain moyen décroît quand le coût de publication de l'annonce augmente, ce qui est parfaitement logique.

5. Un exemple : la loi uniforme.

On suppose que X suit la loi uniforme sur un intervalle $[a, b]$ avec a et b réels positifs.

(a) Si $s > b$, on n'accepte aucune offre et le gain est négatif. On peut donc supposer $s \leq b$.

Si $s < a$, on a $F(s) = 0$ et donc

$$g(s, c) = \int_a^b xf(x) dx - c = \frac{a+b}{2} - c.$$

Si $s \in [a, b]$, on a $F(s) = \frac{s-a}{b-a}$ et donc $1 - F(s) = \frac{b-s}{b-a}$ et ainsi

$$\begin{aligned} g(s, c) &= \frac{b-a}{b-s} \left(\int_s^b \frac{x}{b-a} dx - c \right) \\ &= \frac{b-a}{b-s} \left(\frac{b^2 - s^2}{2(b-a)} - c \right) \\ &= \frac{b+s}{2} - c \frac{b-a}{b-s}. \end{aligned}$$

On a ainsi,

$$\forall s \leq b, \quad g(s, c) = \begin{cases} \frac{a+b}{2} - c & \text{si } s < a \\ \frac{b+s}{2} - c \frac{b-a}{b-s} & \text{sinon.} \end{cases}$$

On remarque au passage que $g(a, c) = \frac{a+b}{2} - c$ de sorte que $s \mapsto g(s, c)$ est continue en a .

- (b) g étant constante sur $[0, a[$ et continue en a , il suffit de l'étudier sur $[a, b]$ pour déterminer son maximum. D'après la question 3.h, on sait que g sera maximale en s tel que $g(s) = s$.

Soit $s \in \mathbf{R}$. On a

$$\begin{aligned} g(s) = s &\Leftrightarrow \frac{b-s}{2} - c \frac{b-a}{b-s} = s \\ &\Leftrightarrow \frac{b^2 - s^2 - 2c(b-a)}{2(b-s)} = s \\ &\Leftrightarrow b^2 - s^2 - 2c(b-a) = 2(b-s)s \\ &\Leftrightarrow s^2 - 2bs + (b^2 - 2c(b-a)) = 0. \end{aligned}$$

On reconnaît un polynôme de degré 2 en s . Son discriminant est

$$\Delta = 4b^2 - 4(b^2 - 2c(b-a)) = 8c(b-a) > 0.$$

Il y a donc deux solutions dans \mathbf{R} :

$$s_1 = b - \sqrt{2c(b-a)} \quad \text{et} \quad s_2 = b + \sqrt{2c(b-a)}.$$

Puisque $s_2 > b$, ce n'est pas une solution dans $[a, b]$. Par contre, on a bien $s_1 < b$ et puisque $c < m = \frac{a+b}{2}$, on a aussi $s_1 > a$ de sorte que

$$g \text{ atteint son maximum en } s = b - \sqrt{2(b-a)c}.$$

6. Simulation informatique.

```
n = 0
a = input('Valeur de a:')
b = input('Valeur de b:')
x = rand()
y = a+(b-a)*rand()
while (y < s)
    x = rand()
    y = a+(b-a)*rand()
n = n+1
```

```
end  
disp(y-n*c,'Gain : ')
```

