Attente de trains

Dans une gare, un train est censé partir à 0 heure, puis un deuxième train à *d* heures (par exemple pour *d* = 1 à 1 heure, ou pour *d* = 2 à 2 heures), ces deux trains ayant la même destination. Plus précisément, le premier train peut avoir du retard, il part en fait entre 0 heure et 1 heure, avec équiprobabilité des temps de départ. Autrement dit, le temps de départ du premier train suit une loi uniforme sur [0 1]. Il en est de même pour le deuxième train qui part entre *d* heures et *d*+1 heures, toujours suivant une probabilité uniforme.

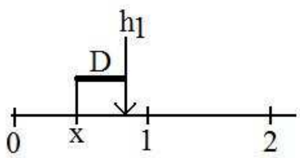
1) Un voyageur arrive à l'instant *x* (nombre réel donné) compris entre 0 et 1 heure. Il prend le premier train qui part. On veut connaître son temps d'attente moyen.

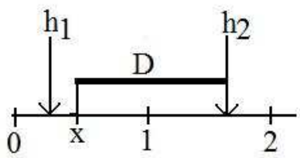
a) Faire le programme qui donne ce temps d'attente moyen.

Deux cas se présentent :

- Le premier train part entre le temps *x* et le temps 1, et c'est ce train que le voyageur prend. Peu importe le départ du deuxième train.
- Le premier train part avant le temps x, et le voyageur l'a raté. Il prend donc le deuxième train.

Ces deux situations sont ainsi visualisées, avec *D* correspondant au temps d'attente, h_1 étant le moment de départ du premier train, et h_2 celui du deuxième (on a pris le cas où *d* = 1) :





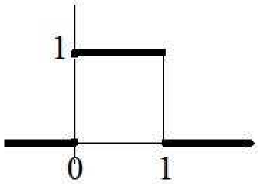
Dans le programme, cela se traduit par le tirage du nombre *h1* entre 0 et 1. S'il est supérieur à *x*, le temps d'attente est *D* = *h1* – *x*, et sinon, on procède au tirage du nombre *h2* entre *d* et *d*+1, et l'on a *D* = *h2* – *x*. Pout avoir la valeur moyenne de *D*, il suffit de faire un grand nombre *NBE* de tirages, en cumulant les valeurs de *D* obtenues. En divisant ce cumul par *NBE*, on obtient cette valeur moyenne.

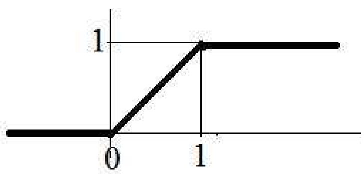
```
#include <stdio.h>
#include <stdlib.h>
#include <conio.h>
#include <time.h>
#define NBE 100000  /* nombre d'expériences */

int main()
{   int essai ;    float x,h1,h2,cumul,attentemoyenne,attenteth,T, delai;
    d=1.;
    srand(time(NULL));
   for(x=0.;x<=1. ;x+=0.01) /* on calcule la moyenne pour 100 valeurs de x */
     {    cumul=0.;
          for(essai=0; essai<NBE; essai++)
           {  h1= (float)rand()/32768.;
              if (h1>x  )  T=h1-x;
```

```
        else    { h2= d+(float)rand()/32768.;    T=h2-x;   }
        cumul+=T;
      }
    attentemoyenne=cumul/(float)NBE;
    attenteth=0.5*(1+(2*d-1)*x-x*x); /* formule théorique de l'attente, voir 2° */
   printf("%3.2f  (%3.2f )   ",attentemoyenne, attenteth);
  }
  getch();return 0;
}
```

b) Déterminer la formule théorique de ce temps d'attente en moyenne. Pour cela calculer d'abord la probabilité *p*(*D* < *t*), selon les valeurs prises par *t*. Ayant ainsi la fonction de répartition de la variable aléatoire *D*, en déduire sa densité, puis sa valeur moyenne.

On est dans un contexte de probabilités continues. La loi uniforme a pour densité 1 sur l'intervalle [01], et 0 ailleurs. La fonction de répartition s'obtient par intégration : $F(x) = \int_{-\infty}^{x} f(t)dt$, soit $F(x) = 0$ pour x négatif, $F(x) = x$ sur [0 1], et $F(x) = 1$ pour x supérieur à 1.





densité de la loi uniforme   fonction de répartition

Ici la loi uniforme s'applique au temps où part un train. En appelant T la variable aléatoire correspondant au temps de départ du premier train, on sait que la fonction de répartition F permet de calculer les probabilités :

$p(T < a) = F(a)$ et $p(a < T < b) = F(b) - F(a)$.

L'évènement $(D < t)$ signifie que le voyageur prend un train qui part entre les instants x et $x + t$. On a déjà :

- pour $t < 0$, $p(D < t) = 0$
- pour $t > d + 1 - x$, $p(D < t) = 1$ puisque dans ce cas le voyageur est sûr de partir avant le temps $d + 1$, départ au plus tard du deuxième train.

Les autres cas retrouvent les deux situations rencontrées précédemment :

- Cas où t est tel que $x + t \leq 1$, ou $t \leq 1 - x$. Le voyageur prend le premier train.

$p(D < t) = p$(départ du premier train entre x et $x+t$)
$= F(x+t) - F(x) = x + t - x = t$ (on est sur [0 1])

- Cas où t est tel que $t > 1 - x$ et aussi $x + t \leq d$, ou $t \leq d - x$, ce cas ne se distinguant du précédent que lorsque d est supérieur à 1.

$p(D < t) = p$(départ du premier train entre x et 1)
$= F(1) - F(x) = 1 - x$

- Cas où t est tel que $t > d - x$ et aussi $x + t \leq d + 1$, soit $t \leq d + 1 - x$,

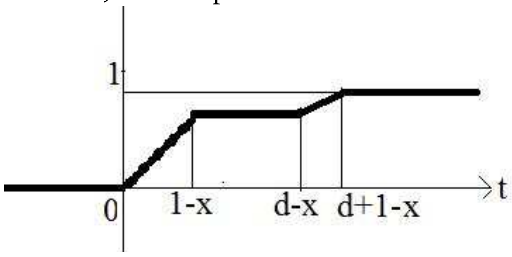
$p(D < t) = p$(départ du premier train entre x et 1 ou départ du deuxième train entre d et $x + t$ quand on a raté le premier)

= p(départ du premier train entre x et 1) + p(départ du deuxième train entre d et $x + t$ quand le premier est parti entre 0 et x)

= $F(1) - F(x) + F(x)\,F(x + t - d) = 1 - x + x\,(x + t - d) = 1 - (d + 1)x + x^2 + xt$.

On a utilisé le fait que la probabilité pour le premier train de partir avant x est égale à $F(x) - F(0) = F(x)$ et que celle du deuxième train de partir entre d et $x + t$ est $F(x + \mathrm{t} - d) - F(0)$ par simple décalage. A noter que la programmation précédente nous a aidé à trouver ce résultat.

La connaissance de $p(D < t)$ donne la fonction de répartition $G(t)$ de la variable aléatoire D, ainsi représentée :



Par dérivation de $G(t)$ par rapport à t, on en déduit la densité g de D :

$g(t) = 1$ sur $[0, 1 - x]$

$g(\mathrm{t}) = x$ sur $[d - x, d + 1 - x]$, et 0 ailleurs. L'espérance $E(D)$ en résulte :

$$E(D) = \int_0^{1-x} t\,g(t)dt + \int_{d-x}^{d+1-x} t\,g(t)dt$$

$$= \int_0^{1-x} t\,dt + \int_{d-x}^{d+1-x} x\,t\,dt = \left[t^2/2\right]_0^{1-x} + x\left[t^2/2\right]_{d-x}^{d+1-x}$$

$$= \frac{1}{2}(1 + (2d - 1)\,x - x^2) \text{ après calcul.}$$

Par exemple pour $d = 1$, $E(D)$ part de 0,5 (une demi-heure) en 0 pour revenir à 0,5 en 1, avec un maximum en 0,5 égal à 0,62 (environ 37 minutes). L'application de cette formule théorique dans le programme précédent donne une bonne concordance avec la valeur obtenue expérimentalement.

2) Maintenant le voyageur arrive entre 0 heure et 1 heure, autrement dit x est un nombre pris au hasard entre 0 et 1. Déterminer le temps d'attente moyen A du voyageur.

Commençons par la programmation. La valeur moyenne $E(D)$ obtenue ci-dessus dépend en fait de x, et peut être notée $E(D_x)$. Pour avoir A, il suffit de cumuler les valeurs de $E(D_x)$ pour de nombreuses valeurs de x prises à intervalles réguliers entre 0 et 1, par exemple pour 100 valeurs :

$A = \sum_{k=0,k<100} E(D_{k/100})/100$. Il suffit d'ajouter ce cumul dans le programme précédent :

```
srand(time(NULL)); cumulattente=0.;
for(x=0.;x<1. ;x+=0.01)
{ cumul=0.;
  for(essai=0; essai<NBE;essai++)
    {  h1= (float)rand()/32768.;
       if (h1>x  )  T=h1-x;
       else { h2= d+(float)rand()/32768. ;  T=h2-x; }
```

```
        cumul+=T;
      }
    attentemoyenne=cumul/(float)NBE;
    cumulattente+=attentemoyenne;
  }
  moyenneattentes=cumulattente/100.; /* il s'agit de A */
  printf("   ****  %3.2f   (%3.2f) ",moyenneattentes, 1./12.+0.5*d);
```

La formule donnant A correspond à une somme d'aires de petits rectangles de base 1/100 et de hauteur $E(D_{k/100})$. En termes théoriques, cela s'écrit : $A=\int_0^1 E(D_x)dx$ puisque c'est l'aire entre la courbe de $E(D_x)$ et l'axe des x sur [0 1]. D'où :

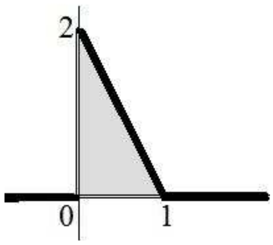
$$A=\frac{1}{2}\int_0^1 (1+(2d-1)x-x^2)dx=\frac{1}{2}\left[x+(2d-1)\frac{x^2}{2}-\frac{x^3}{3}\right]_0^1$$

$$=\frac{1}{12}+\frac{d}{2}$$

Cette valeur reste proche de d / 2.

B- Nous reprenons le même problème mais en changeant les retards pris par chaque train. A lieu de suivre une loi uniforme sur [0 1], le départ du premier train suit une loi linéaire décroissante, de densité $f(x) = 2(1 - x)$ sur [0 1] et nulle ailleurs. Et de même pour le deuxième train, entre d et d+1.

1) Vérifier qu'il s'agit bien d'une densité de probabilité, et donner la fonction F de répartition correspondante.

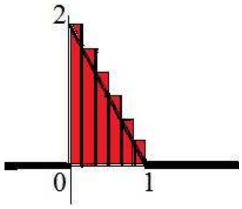


La fonction f est continue sur $\mathbb{R}$, sauf en 0 où elle présente une discontinuité, et d'autre part : $\int_{-\infty}^{+\infty} f(t)dt=\int_0^1 2(1-t)dt=$ aire du triangle $=1$. Tout cela lui octroie le droit d'être une densité, ici associée à la variable aléatoire correspondant au temps de départ du premier train (ou à son retard).

La fonction de répartition est $F(x)$ = $\int_{-\infty}^{x} f(t)dt=\int_0^x 2(1-t)dt=2x-x^2$ sur [0 1] et d'autre part $F(x) = 0$ pour x négatif et $F(x) = 1$ pour x supérieur à 1.

2) Faire un programme permettant d'obtenir cette distribution décroissante de probabilités. Autrement dit, comment faire lorsque l'on dispose seulement d'une distribution uniforme, celle donnée par l'ordinateur ?

Commençons par passer en mode discret, en remplaçant la fonction densité par un diagramme en forme de rectangles. L'intervalle [0 1] est découpé en un certain nombre de petits intervalles de même longueur (nous avons pris *nbintervalles* = 10 ou *nbintervalles* = 100 dans le programme).



Prenons l'exemple où le nombre d'intervalles est 10. On va fabriquer 10 casiers pouvant contenir un nombre décroissant de numéros. Le casier 0 peut contenir les 10 numéros de 0 à 9. Le casier 1 peut contenir les 9 numéros de 10 à 18, le casier 2 les 8 numéros de 19 à 26, et ainsi de suite jusqu'à la boîte 9 qui peut contenir un numéro, en l'occurrence 54 (soit 10 . 11 / 2).

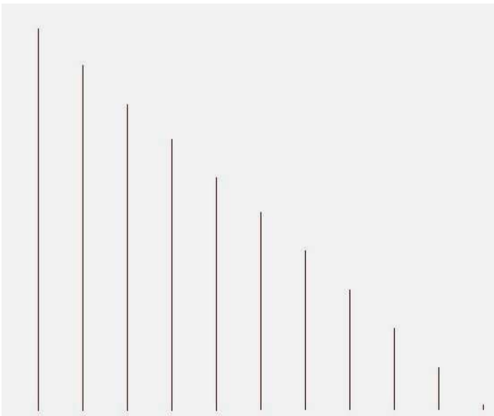
En faisant tomber dans ces casiers une pluie de numéros au hasard (avec une densité uniforme), les casiers vont être atteints un nombre décroissant de fois, suivant une disposition régulière, celle que nous voulons réaliser. Les numéros tirés au hasard vont de 0 à 54 pour *nbintervalles* = 10, plus généralement chaque numéro est compris entre 0 et *nbintervalles** *nbintervalles* / 2 = *nombrei*, ce dernier nombre n'étant pas pris. Quand un numéro tombe dans un casier, il s'agit maintenant de savoir quel est le nombre associé à ce casier (entre 0 et 9 si le nombre de casiers (*nbintervalles*) est 10). C'est ce que va faire la fonction *tirage*(). Prenons un exemple : *numero* = 22 (le nombre d'intervalles étant 10). En conditions initiales, on met un compteur à -1. Puis on commence par enlever 10, il reste 22 – 10 = 12, le compteur passe à 0 (*compteur* augmente de 1 à chaque étape). Ensuite on enlève 9 (ce nombre *enlever* diminue de un à chaque étape) :   = 12 – 9 = 3, le compteur passe à 1. On enlève 8 au reste : 3 – 8 < 0, le compteur passe à 2 et l'on s'arrête quand le reste est négatif. On a trouvé qu'il s'agit du casier numéro 2, et c'est ce que ramène la fonction *tirage*(). Si le nombre d'intervalles est 10, cette fonction ramène un nombre entre 0 et 9, ces nombres ayant une distribution de probabilité régulièrement décroissante. Si le nombre d'intervalles est 100, les nombres ramenés sont entre 0 et 99.

```
int tirage(void)
{int numero,reste,enlever,compteur;
  numero=((float)rand()/32768.)*(float)nombrei;
  reste=numero; enlever=nbintervalles; compteur=-1;
  do {   reste-=enlever; enlever-=1; compteur++;  }
  while (reste>=0);
  return compteur;
}
```

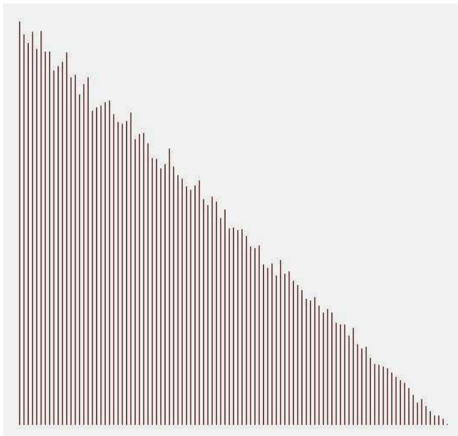
En répétant ces tirages un grand nombre de fois, on obtient un histogramme des fréquences qui correspond bien à cette décroissance régulière. C'est ce que fait le programme qui suit.

```
#define xorig 100
#define yorig 550
#define NBE 70000
SDL_Surface * ecran;    Uint32 rouge,blanc;
int nbintevalles,nombrei,resultat , freq[1000],essai,i;
```

```
int main ( int argc, char** argv )
{
   SDL_Init( SDL_INIT_VIDEO );
   ecran=SDL_SetVideoMode(800,600,32, SDL_HWSURFACE|SDL_DOUBLEBUF);
   blanc=SDL_MapRGB(ecran->format,255,255,255);
   rouge=SDL_MapRGB(ecran->format,120,0,0);
   SDL_FillRect(ecran,0,blanc);
   srand(time(NULL));
   nbintervalles=10;   nombrei= nbintervalles*(nbintevalles+1)/2;
   for(i=0;i<=nbintervalles;i++) freq[i]=0;
   for(essai=0;essai<NBE; essai++)
     {  resultat=tirage();  freq[resultat]++; }
   for(i=0;i<=nbintervalles;i++)    ligne(xorig+500./nbintervalles*i , yorig ,
                              xorig+500/nbintervalles*i ,  yorig-(float)freq[i]/300.*delta, rouge);
   SDL_Flip( ecran);  pause(); return 0;
}
```



nombre d'intervalles = 10



nombre d'intervalles = 100

3) Utiliser les programmes précédents pour avoir le temps moyen d'attente du voyageur.

```
#define NBE 100000
int nbintervalles,nombrei;
int main()
{   int essai ;  float x,h1,h2,cumul,attentemoyenne,attenteth,T,d;
    d=1; nbintervalles=100;       srand(time(NULL));
     nombrei= nbintervalles*(nbintervalles+1)/2;
    for(x=0.;x<=1. ;x+=0.01)
      { cumul=0.;
         for(essai=0; essai<NBE;essai++)
           { h1=(float)tirage();  h1=h1/(float)nbintervalles;
              if (h1>x  )  T=h1-x;
              else   { h2= d+(float)tirage()/(float)nbintervalles; T=h2-x;   }
              cumul+=T;
           }
        attentemoyenne=cumul/(float)NBE;
        attenteth=1./3.+5.*x/3.-7*x*x/3.+2*x*x*x/3.; /* formule ci-dessous */
        printf("%3.2f: %3.2f (%3.2f )  ",x,attentemoyenne,attenteth);
```

```
    }
  getch();return 0;
}
```

4) Faire le calcul théorique du temps d'attente moyen *D*, dans le cas où *d* = 1 (le deuxième train part une heure après le premier).

La densité associée au retard du premier train est, rappelons-le, $f(t) = 2(1 - t)$ pour t entre 0 et 1, et nulle ailleurs. La fonction de répartition est $F(t) = 2t - t^2$ sur [0 1]. Passons à la fonction G de répartition du temps d'attente du voyageur, pour $d = 1$. C'est, en utilisant les formules déjà trouvées au A- :

- Pour $0 \leq t < 1 - x$, $G(t) = F(x+t) - F(x) = - t^2 + 2(1 - x)\, t$, d'où une densité $g(t) = - 2t + 2(1 - x)$.
- Pour $1 - x \leq t \leq 2 - x$, $G(t) = 1 - F(x) + F(x)\, F(t + x - 1)$ $= - (2x - x^2)t^2 + (2x - x^2)(4 - 2x)t + cte$, après calcul, d'où la densité $g(t) = (2x - x^2)(- 2t + 4 - 2x)$.

Ce sont les seuls intervalles où la densité n'est pas nulle. On en déduit par intégration le temps d'attente moyen :

$$D = \int_0^{1-x} (-2t + 2(1-x))\, t\, dt + \int_{1-x}^{2-x} \left(2x - x^2\right)\left(-2t + 4 - 2x\right) t\, dt$$

$$= \left[-2t^3/3 + (1-x)t^2\right]_0^{1-x} + \left[(2x - x^2)(2t^3/3 + (2-x)t^2\right]_{1-x}^{2-x}$$

$$= (1/3)\,(1 - x)^3 + (2x - x^2)(4/3 - x) \quad \text{(calculs abrégés)}$$

$$= (1/3) + (5/3)\, x - (7/3)\, x^2 + (2/3)\, x^3.$$

En notant D comme une fonction de x, on peut vérifier que $D(0) = 1/3$, qui est la valeur moyenne du retard du premier train (on vérifie en effet que $\int_0^1 t\, f(t) dt = \int_0^1 2t(1-t) dt = 1/3$) et l'on aussi $D(1) = 1/3$, puisque le voyageur est aussi juste à l'heure pour le départ officiel du deuxième train. Le temps d'attente atteint un maximum pour x de l'ordre de 0,5, et ce maximum est de l'ordre de 0,67.