

Licence Creative Commons



MAJ: 10 février 2013

Probabilités pour l'informatique (I)



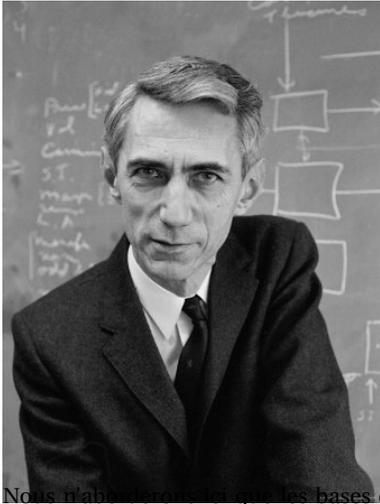
Guillaume CONNAN

TABLE DES MATIÈRES

1 Scilab for dummies	4
1.1 Avant de commencer	5
1.2 Scilab comme super-calculatrice	5
1.2.1 Les opérations de base	5
1.2.2 Un brin d'informatique	6
1.2.3 Définir une fonction	7
1.3 L'objet de base de Scilab : la matrice	7
1.3.1 Utilisation d'une matrice ligne	7
1.4 Les graphiques	9
1.5 La programmation	10
1.5.1 Les tests	10
1.5.2 Boucles for	12
1.5.3 Boucles while	12
1.6 Probabilités et statistiques	13
2 Dénombrement	14
2.1 Rappels de théorie des ensembles	15
2.1.1 Ensembles finis	15
2.1.2 Parties d'un ensemble	15
2.2 Une dose d'algèbre générale	15
2.2.1 Relation d'ordre	15
2.2.2 Loi de composition interne	16
2.2.3 Lien avec les opérateurs logiques	16
2.2.4 Partition	16
2.2.5 Produit cartésien	16
2.3 Quelques résultats sur les cardinaux	16
2.3.1 Cardinal de $\mathcal{P}(E)$	16
2.3.2 Cardinal d'une partition	17
2.3.3 Cardinaux et inclusion	17
2.3.4 Cardinal d'un produit cartésien	17
2.3.5 Formule du crible	17
2.4 Dénombrement	18
2.4.1 Nombre de permutations	18
2.4.2 Nombre d'applications injectives - arrangements sans répétitions	18
2.4.3 Combinaisons sans répétition	19
2.5 Triangle de Pascal - Binôme de Newton	19
2.6 EXERCICES	21
3 Espace probabilisé discret	28
3.1 Probabilités?	29
3.2 Avant la formalisation	29
3.3 Espace probabilisable - Espace probabilisé	30
3.4 Probabilités conditionnelles	32
3.4.1 Un exemple pour comprendre	32
3.4.2 Définition	32
3.4.3 Arbre	32
3.4.4 Formule de BAYES	33
3.4.5 Indépendance	33
3.5 Variables aléatoires réelles finies	34
3.5.1 Définition	34
3.5.2 Loi de probabilité	35
3.5.3 Variables aléatoires indépendantes	36
3.5.4 Fonctions de répartition	37
3.5.5 Espérance mathématique	37
3.5.6 Variance	38
3.5.7 Linéarité de l'espérance	38

3.5.8	Théorème de König-Huygens	38
3.5.9	V.a.r. centrée réduite	38
3.5.10	Fonction indicatrice d'un évènement	39
3.6	Quelques lois discrètes classiques	39
3.6.1	Loi uniforme (discrète)	39
3.6.2	Loi de Bernoulli : to be or not to be	39
3.6.3	Loi binomiale	40
3.6.4	Loi hypergéométrique	40
3.6.5	Loi géométrique	41
3.6.6	Loi de Pascal	41
3.6.7	Loi binomiale négative	41
3.6.8	Loi de Poisson	42
3.7	Somme de v.a.r. indépendantes - Stabilité	42
3.7.1	Convolution	42
3.7.2	Somme de v.a.r. indépendantes	42
3.7.3	Applications	43
3.8	EXERCICES	44
4	Initiation aux processus aléatoires	55
4.1	Vision dynamique des probabilités - Automates	56
4.1.1	Découverte	56
4.1.2	Premier exemple - Règles de parcours	56
4.1.3	Deuxième exemple - Valeur moyenne	58
4.2	Comportement asymptotique des chaînes de MARKOV	59
4.3	Exercices	63
4.3.1	Chaînes absorbantes - automates	63
4.3.2	Avec le calcul matriciel	64
5	Loi normale	65
5.1	Densité de probabilité	66
5.2	Tout sur la loi normale	66
6	Simulation informatique et probabilités	68

Introduction



Claude Elwood SHANNON (1916 - 2001) est un mathématicien-inventeur-jongleur américain qui, suite à son article « *A mathematical theory of communications* » paru en 1948, est considéré comme le fondateur de la *théorie de l'information* qui est bien sûr une des bases de...l'*informatique*.

L'idée est d'étudier et de quantifier l'« information » émise et reçue : quelle est la compression maximale de données digitales? Quel débit choisir pour transmettre un message dans un canal « bruité »? Quel est le niveau de sûreté d'un chiffrement?...

La théorie de l'information de SHANNON est fondée sur des modèles probabilistes : leur étude est donc un préalable à l'étude de problèmes de réseaux, d'intelligence artificielle, de systèmes complexes.

Par exemple, dans le schéma de communication présenté par SHANNON, la source et le destinataire d'une information étant séparés, des perturbations peuvent créer une différence entre le message émis et le message reçu. Ces perturbations (bruit de fond thermique ou acoustique, erreurs d'écriture ou de lecture, etc.) sont de nature *aléatoire* : il n'est pas possible de prévoir leur effet. De plus, le message source est par nature *imprévisible* du point de vue du destinataire (sinon, à quoi bon le transmettre).

Nous réviserons donc les bases de la théorie des probabilités afin de faciliter l'étude de la théorie de l'information qui sera abordée par certain(e)s à leur sortie de l'IUT. L'information étant transmise sous forme digitale, nous étudierons principalement les probabilités discrètes sans pour autant négliger le monde analogique : nous aborderons donc aussi les probabilités continues.

La théorie des probabilités est une branche des mathématiques qui permet d'étudier les phénomènes où le hasard intervient et permet de formaliser le raisonnement en présence d'informations partielles. C'est une science *déductive* au même titre que les autres branches des mathématiques. Vous avez « tourné » autour des probabilités depuis de nombreuses années mais sans vraiment définir les notions que vous aviez abordées : c'est un peu à l'image de cette partie des mathématiques qui a fait son apparition aux alentours du XVI^e siècle pour permettre aux princes d'estimer leur espérance de gain mais qui n'a reçu sa formalisation rigoureuse qu'au début du XX^e siècle sous l'impulsion du russe KOLMOGOROV, alors qu'à titre de comparaison, la géométrie avait subi le même sort par EUCLIDE...2200 ans plus tôt!

La statistique est elle une science *inductive* : elle part d'observations élémentaires et on cherche à induire des lois générales.

Nous veillerons donc à séparer leur étude même si ces domaines interagissent largement.

D'ailleurs, le domaine d'application des probabilités en informatique dépasse la théorie de l'information : on évoquera l'algorithme de MILLER-RABIN qui permet de générer des grands nombres entiers « fortement probablement premiers », la gestion des files d'attente (de personnes, de fichiers, dans un réseau,...), etc.

Scilab for dummies



Scilab est un logiciel libre, multiplateforme de calcul numérique orienté informatique industrielle et ingénierie. Certains d'entre vous l'utiliseront peut-être dans leurs projets de seconde année. Il dispose également d'outils performants pour traiter les statistiques et les probabilités. Le logiciel R était également bien placé mais son typage plus que flou ne pouvait convenir à des informaticien(ne)s... Il y avait aussi la piste Caml : on peut tout faire en Caml et d'ailleurs un collègue de la Fac de maths travaille actuellement à fabriquer un scilab made in Caml mais pour l'instant, il est plus utile d'utiliser une boîte à outils bien faite et largement répandue dans le monde industriel, ou tout au moins sous la forme de son clone au prix exorbitant, Matlab...

1 Avant de commencer

SCILAB est téléchargeable rapidement à l'adresse suivante :

<http://www.scilab.org/>

Le logiciel est accompagné d'une documentation très complète. Vous pouvez compléter votre apprentissage en parcourant « Le guide du calcul avec des logiciels libres » paru chez Dunod en 2008 :-)

De nombreux TP sont présents sur le site de Scilab référencé ci-dessus.

Il faut savoir que scilab, comme matlab, est basé sur un noyau de fonctions en Fortran ou en C déjà compilées donc rapides d'utilisation. En revanche, tout ce que vous créez vous-même sera interprété ce qui ralentira l'exécution par rapport à d'autres langages. De plus, la gestion des nombres n'est pas toujours optimale car ils sont codés en complexes double précision ce qui peut être lourd.

2 Scilab comme super-calculatrice

2.1 Les opérations de base

Scilab peut fonctionner de manière classique :

```
--> 3+2
```

Si on ajoute un « ; » à la fin de la ligne, celle-ci est lue par SCILAB mais le résultat n'est pas affiché : c'est le mode silencieux.

Mais Scilab fait du calcul numérique ; les résultats sont donc donnés sous forme de nombres à virgule flottante :

```
--> 1+sqrt(2)
```

sachant que sqrt désigne la racine carrée (SQare RooT en anglais...).

Navigation et historique

On ne peut pas recommencer un calcul en plaçant le curseur de la souris sur une ligne précédente déjà évaluée. En revanche, avec les touches  et , on peut naviguer dans l'historique et réafficher des commandes précédentes.

Remarque

Par défaut, les résultats sont affichés avec 10 caractères. Pour avoir plus de précision, on utilise format :

```
--> format(20); 1+sqrt(2)
```

```
2.41421356237309492
```

On peut vouloir utiliser l'écriture scientifique d'un nombre : il faut préciser l'option 'e' :

```
--> format('e',20); 1+sqrt(2)
```

```
2.4142135623731D+00
```

Le D+00 signifie mathématiquement $\times 10^0$.

On peut revenir à la configuration par défaut avec l'option 'v' :

```
--> format('v',10)
```

Certaines constantes utiles sont déjà implémentées comme π :

```
--> %pi
```

```
3.1415927
```

et exp(1) :

```
--> %e
```

2.7182818

Pour les fonctions usuelles, la syntaxe est assez standard. Veuillez toutefois noter que le logarithme népérien se note `log` :

```
--> log(%e^2)
```

2

Notez une particularité des logiciels de calcul numérique :

```
--> sin(%pi)
```

1.225D - 16

C'est le « zéro » fixé par Scilab qui est affiché...

On peut faire malgré tout des calculs avec des nombres plus petits :

```
--> 1D-25 - 1D-26
```

9.000D - 26

La valeur du « zéro » de Scilab est donnée par `%eps` :

```
--> %eps
```

2.220D - 16

Cela peut jouer des tours pour l'addition :

```
--> N=1D30; P=1D10;
```

```
--> N-N+D
```

```
--> N+D-N
```

Il faudra donc faire attention au moment de faire des tests d'égalité. Nous étudierons la norme IEEE 754 plus en détail en deuxième année.

On peut travailler avec des nombre complexes, le nombre de carré -1 se notant `%i` :

```
--> (3+%i)^2
```

```
--> abs(1+%i)
```

2 2 Un brin d'informatique

2 2 1 Types

Scilab travaille avec des objets de différents types et ne les traite pas de la même manière. Il faut distinguer :

- les réels et les complexes ;
- les booléens : `%T` pour vrai (True) et `%F` pour faux (False) ;
- les polynômes ;
- les chaînes de caractères entre apostrophes ;
- les listes ;
- les fonctions ;
- et bien d'autres encore...

On obtient le type d'un objet avec `typeof` :

```
--> typeof(%e)
```

constant

```
--> typeof(' %e ')
```

string

2 2 2 Opérateurs

Scilab possède les opérateurs arithmétiques habituels (+), (-), (*), (/) mais aussi une division inversée (\) et les opérateurs pointés dont nous parlerons plus tard.

Il sera également important d'utiliser les opérateurs booléens qui testent l'égalité (==), la non-égalité (<>), les inégalités strictes et larges (<), (>), (<=), (>=). Le résultat est un booléen (%T ou %F) :

```
--> 1+%eps/2==1
```

T

Ce dernier résultat s'explique du fait que la différence entre les deux termes est inférieure à %eps

Enfin il existe des opérateurs logiques qui seront également utiles dans les tests : la négation (~), l'opérateur ET (&) et l'opérateur OU (|). Par exemple :

```
--> ~(3<3) == (3>=3)
```

T

La négation de $3 < 3$ est bien $3 \geq 3$.

2 2 3 Affectation

Pour donner un nom à une case mémoire où nous voulons stocker un objet, on utilise (=) :

```
--> a=3;
```

```
--> a^2
```

```
--> ans^2
```

La commande ans reprend en effet le résultat de la commande précédente.

La commande who permet de rappeler quelles sont les variables affectées.

Pour supprimer une variable affectée, on utilise clear :

```
--> clear a
```

La commande clear utilisée toute seule effacera toutes les variables affectées.

2 3 Définir une fonction

On peut définir toutes sortes de fonctions, d'une ou plusieurs variables, numériques ou non à l'aide de la syntaxe suivante :

```
--> function y=truc(x)
      y=x*log(x)
endfunction
```

On utilise ensuite la fonction créée de manière naturelle :

```
--> truc(2)
```

1.3862944

3**L'objet de base de Scilab : la matrice****3 1 Utilisation d'une matrice ligne**

Même si on ne s'en aperçoit pas tout de suite, tout est matrice pour Scilab, c'est-à-dire un sorte de tableau. Cela n'est pas naturel à un(e) élève sortant du lycée mais cela peut présenter des avantages. Occupons-nous d'abord des matrices-ligne, c'est-à-dire d'une sorte de tableau comportant une seule ligne.

Par exemple, nous voulons obtenir les valeurs prises par une fonction en certaines valeurs :

```
--> va = [1 3 74 100*%pi]
```

1 3 74 314.15927

Si on veut que ces valeurs soient incrémentées régulièrement :

```
--> x = [0:2:11]
```

0 2 4 6 8 10

On obtient ainsi une suite arithmétique de nombres de premier terme 0, de raison 2 et inférieurs à 11. Par défaut, l'incrément est de 1 :

```
--> y = [0:9]
```

0 1 2 3 4 5 6 7 8 9

L'incrément peut être négatif et non entier :

```
--> X = [0:-2.5:-13]
```

0 -2.5 -5 -7.5 -10 -12.5

On a facilement les valeurs de quelques suites obtenues à partir des précédentes :

```
--> x^2
```

0 4 16 36 64 100

On peut extraire une valeur de ces matrices :

```
--> x(3)
```

4

ou une partie de ces valeurs :

```
--> x(1:3)
```

0 2 4

Le dernier élément d'une matrice est obtenu à l'aide de \$:

```
--> x($)
```

On peut extraire des composantes vérifiant une condition avec la commande find qui renvoie les rangs répondant à la demande :

```
--> find(x>5)
```

4 5 6

On peut ajouter une valeur à une matrice :

```
--> Xnew=[X,32]
```

0 -2.5 -5 -7.5 -10 -12.5 32

ou concaténer deux matrices :

```
--> GrandX=[x,[1 2]]
```

0 2 4 6 8 10 1 2

On peut effectuer des opérations entre ces vecteurs. Par exemple, essayons de les multiplier terme à terme :

```
--> x*X
```

Attention ! Les opérateurs arithmétiques (*), (/) et (\) effectuent des opérations sur des matrices. Souvenez-vous que ces opérations n'ont presque rien à voir avec celles sur les entiers. En particulier, $a * b \neq b * a$ en général.

Pour effectuer ces dernières opérations « terme à terme », il faut utiliser les opérateurs pointés sans oublier de laisser une espace entre le premier terme et le point :

```
--> x .*X
```

```
0 -5 -20 -45 -80 -125
```

Des matrices particulières peuvent être utiles : **zeros (m, n)** et **ones (m, n)** renvoient des matrices remplies respectivement de 0 et de 1.

Pour extraire des lignes ou des colonnes ou des éléments particuliers :

```
A=[1,2,3;
   4,5,6;
   7,8,9]
A(1,3)
A(1,:)
A(:,1)
A($,:)
```

4

Les graphiques

On commence par créer un vecteur de 10000 valeurs régulièrement espacées entre 0 et 100 :

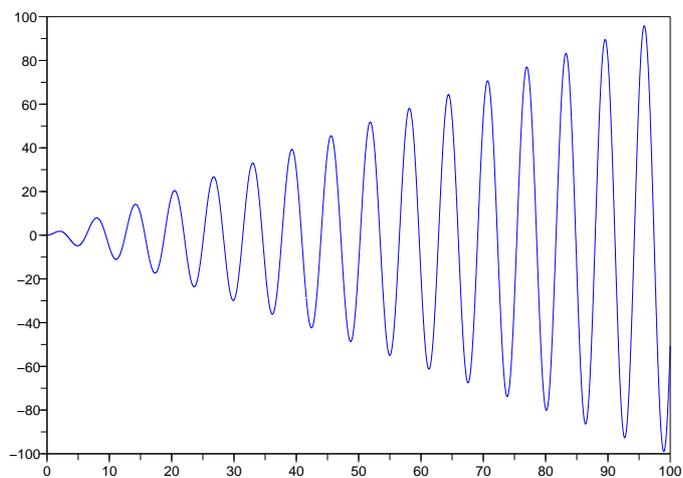
```
--> x=linspace(0,100,10000);
```

On aurait pu aussi bien poser :

```
--> x=0:.01:100
```

Nous voulons obtenir la représentation graphique de la fonction $x \mapsto x \cdot \sin(x)$ sur $[0; 100]$:

```
--> plot(x,x .*sin(x))
```



Pour rajouter d'autres courbes sur la figure, nous avons plusieurs possibilités :

- demander une autre courbe :

```
--> plot(x,x)
```

```
--> plot(x,-x)
```

– soit tout demander en même temps. Pour éviter les superpositions, il faut donc effacer la figure précédente grâce à la commande `clf` :

```
--> clf; plot(x,x .*sin(x),x,x,x,-x)
```

On peut modifier le style des courbes à partir des menus de la fenêtre graphique ou en ligne de commandes :

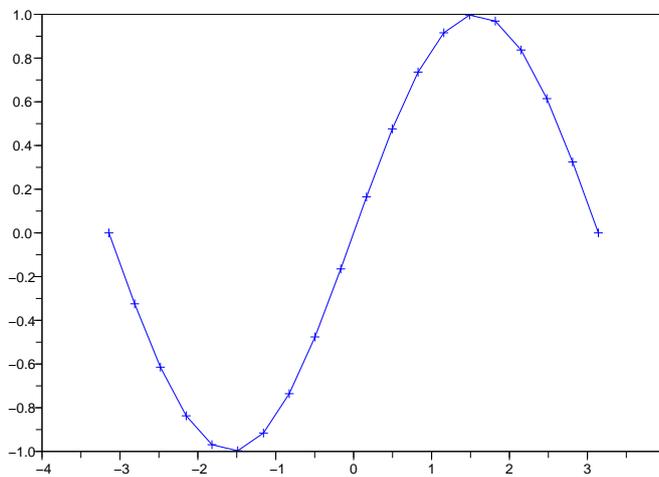
```
--> X=linspace(-%pi,%pi,20);
```

ou bien

```
--> X=-%pi:%pi/10:%pi
```

```
--> clf; plot(X,sin(X),"+-b")
```

L'option `+b` signifie qu'on représente chaque point par une croix (+), qu'on les relie par un trait continu (-), en bleu (b).



Il existe d'autres raccourcis :

Couleur	jaune	magenta	cyan	rouge	vert	bleu	blanc	noir
Symbole	y	m	c	r	g	b	w	k

Style de ligne	plein	tirets	pointillés	mixte
Symbole	-	-	:	-.

Style de point	plus	rond	astérique	point	croix
Symbole	+	o	*	.	x

N'hésitez pas à explorer l'aide sur `gce` et `gda` pour plus de « configurabilité ».

5 La programmation

5 1 Les tests

5 1 1 Si...alors...sinon

Construisons par exemple une fonction donnant la valeur absolue d'un nombre ^a

a. Cette fonction existe déjà sur Scilab (`abs`) mais il nous faut un exemple simple...

```
--> function y = ABS(x)
    if x >= 0 then
        y=x
    else
        y=-x
    end
endfunction
```

En indentant son programme pour mettre en valeur les différentes structures créées, on le rend plus lisible.

5 1 2 Éditeur de programme

Il existe un éditeur de texte dédié à Scilab (onglet Editor) qui permet de travailler à part sur son programme avant de l'exécuter.



```
File Edit Search Execute Debug Scheme Options Windows
1 function y=ABS(x)
2     if x>=0 then
3         y=x
4     else
5         y=-x
6     end
7 endfunction
```

On clique alors sur Execute → Load into scialb ou on appuie sur **Ctrl** + **L** pour exécuter le programme et l'envoyer vers Scilab. Si une erreur est décelée, le « débogueur » est là pour vous aider à la déceler.

De retour dans Scilab, on appuie sur **Entrée** et on peut utiliser la fonction ABS comme tout à l'heure. Cela ne sera vraiment utile que pour de longs programmes.

On peut également, depuis l'éditeur, enregistrer le programme puis utiliser la commande `exec le_chemin_vers_le_f`

```
--> exec('./ABS.sci')
```

5 1 3 Emploi de return et de elseif

Dans l'exemple précédent il n'y avait qu'une seule alternative : x ou $-x$.

Intéressons-nous maintenant à des programmes dont les structures conditionnelles sont imbriquées.

Par exemple, construisons une fonction Sol qui donne les solutions d'une équation du second degré de la forme $ax^2 + bx + c = 0$. On pourrait construire notre programme comme précédemment :

```
function y=Sol(a,b,c)
d=b^2-4*a*c
if d>0 then
    y=[(-b-sqrt(d))/(2*a) (-b+sqrt(d))/(2*a)]
else
    if d<0 then
        y=[(-b-%i*sqrt(-d))/(2*a) (-b+%i*sqrt(-d))/(2*a)]
    else
        y=-b/(2*a)
    end
end
endfunction
```

Cependant, il faut être très prudent et ne pas oublier de `end` et bien savoir se retrouver dans l'imbrication de conditions.

Il existe un moyen plus efficace qui permet de rester dans le même test : la commande `elseif` :

```
function y=Sol(a,b,c)
d=b^2-4*a*c
if d>0 then
    y=[(-b-sqrt(d))/(2*a) (-b+sqrt(d))/(2*a)]
elseif d<0 then
    y=[(-b-%i*sqrt(-d))/(2*a) (-b+%i*sqrt(-d))/(2*a)]
else
    y=-b/(2*a)
end
endfunction
```

Un moyen qui peut parfois s'avérer encore plus efficace est d'utiliser `return` qui permet de sortir d'un programme et de gagner ainsi du temps de compilation :

```
function y=Sol(a,b,c)
d=b^2-4*a*c
  if d>0 then
    y=[(-b-sqrt(d))/(2*a) (-b+sqrt(d))/(2*a)]
    return(y)
  end
  if d<0 then
    y=[(-b-%i*sqrt(-d))/(2*a) (-b+%i*sqrt(-d))/(2*a)]
    return(y)
  end
  y=-b/(2*a)
endfunction
```

5 2 Boucles for

Par exemple, calculons la somme S des 25 premiers entiers. Il faut commencer par initialiser S :

```
--> S=0;
```

```
--> for i=1:25 do S=S+i; end
```

Rappelez-vous que `1:25` crée la liste des entiers de 1 à 25. L'emploi du mode silencieux après le `S=S+i` permet d'éviter d'afficher tous les résultats intermédiaires.

On peut aussi boucler sur les éléments d'une matrice.

5 3 Boucles while

Reprenons le calcul de la somme des 25 premiers entiers :

```
--> S=0; i=1;
```

```
--> while i<=25 do S=S+i; i=i+1; end
```

6 Probabilités et statistiques

Voici quelques outils qui nous seront utiles en TD.

rand(l, c) renvoie une matrice de l lignes et c colonnes dont les coefficients sont des nombres de [0, 1[aléatoirement choisis selon la distribution uniforme.

Par exemple, pour simuler 100 sorties d'un dé à six faces, on peut faire :

```
-->m=floor(6*rand(10,10)+1)
```

m =

1.	4.	6.	4.	2.	6.	5.	3.	6.	4.
1.	5.	2.	6.	5.	6.	5.	6.	5.	5.
6.	2.	1.	1.	5.	1.	4.	6.	5.	2.
5.	1.	1.	1.	4.	2.	5.	2.	1.	1.
2.	1.	5.	5.	4.	4.	3.	6.	6.	4.
5.	3.	1.	2.	6.	6.	6.	1.	5.	6.
3.	5.	1.	3.	3.	2.	3.	2.	6.	2.
3.	3.	2.	5.	1.	3.	1.	1.	6.	2.
2.	5.	5.	1.	4.	5.	4.	4.	3.	4.
3.	6.	4.	1.	3.	5.	6.	5.	6.	3.

La fonction **tabul(mat)** est très pratique : elle renvoie une matrice de deux colonnes avec les valeurs prises par les coefficients de **mat** sur la première colonne et leurs effectifs respectifs sur la deuxième.

```
t = tabul(m)
```

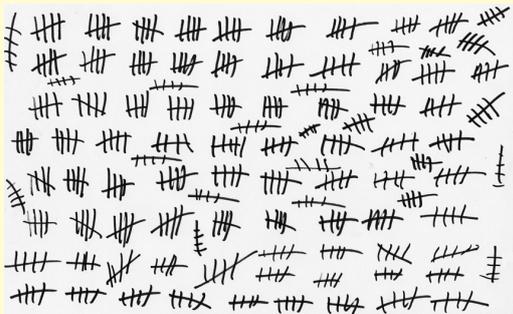
Pour obtenir une représentation graphique sous forme d'un diagramme en bâtons, on utilise **bar(liste des ab**

```
-->bar(t(:,1),t(:,2))
```

On peut préférer un camembert (une tarte en anglais...)

```
-->pie(r(:,2),r(:,1),string(r(:,1)))
```

Dénombrement



Lors de l'étude des probabilités discrètes, il est primordial de compter le nombre d'événements réalisables. On met donc en « correspondance » des éléments de \mathbb{N} avec des événements : on dénombre.

Cela peut parfois être simple : combien y a-t-il d'issues à une expérience de pile ou face ?...

Mais on peut souvent avoir besoin d'outils performants que nous allons présenter après quelques petits rappels.

1 Rappels de théorie des ensembles

1 1 Ensembles finis

Il est aisé de comprendre intuitivement ce qu'est un ensemble fini. En voici une définition :

Définition 2 - 1

Ensemble fini

On dit qu'un ensemble E est fini si, et seulement si, il existe un entier naturel n et une bijection φ de $\llbracket 1, n \rrbracket$ sur E .

Cela signifie en fait que chaque élément d'un ensemble fini E porte un numéro de dossier et qu'on connaît le nombre total de dossiers : n . Ce nombre jouant un rôle important, on lui donne un nom :

Définition 2 - 2

Cardinal d'un ensemble fini

L'entier n défini précédemment est appelé cardinal de E et noté $\text{Card}(E)$ ou $|E|$.

Si on considère l'ensemble formé des pages de cet ouvrage, il constitue un ensemble fini : quel est son cardinal ?

Par convention, l'ensemble vide est de cardinal nul : $|\emptyset| = 0$.

1 2 Parties d'un ensemble

Définition 2 - 3

Parties d'un ensemble

Soit E un ensemble. L'ensemble formé de tous les sous-ensembles de E est appelé ensemble des parties de E et noté $\mathcal{P}(E)$.

Il faudra donc faire bien attention au moment d'utiliser les symboles \in et \subseteq . Ainsi :

$$A \in \mathcal{P}(E) \iff A \subseteq E$$

Si $E_2 = \{a, b\}$ alors $\mathcal{P}(E_2) = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$.

Nous aurons également besoin de parler des éléments de E qui n'appartiennent pas à une de ses parties.

Définition 2 - 4

Complémentaire d'une partie

Soit A une partie de E . On appelle **complémentaire** de A dans E l'ensemble des éléments de E qui n'appartiennent pas à A . On note cet ensemble \overline{A}_E , $\complement_E A$ ou plus simplement \overline{A} quand il n'y a pas d'ambiguïté.

Par exemple, le complémentaire de l'ensemble des filles dans une classe est l'ensemble des garçons de cette classe. La notation permet de faire le lien avec la négation d'une proposition : la négation de « $x \in A$ » est « $x \in \overline{A}$ »...

2 Une dose d'algèbre générale

2 1 Relation d'ordre

Soit A , B et C trois éléments de $\mathcal{P}(E)$.

Considérons la relation d'inclusion (\subseteq) qui relie deux éléments quelconques de $\mathcal{P}(E)$.

On obtient facilement que :

- $A \subseteq A$: on dit que la relation \subseteq est réflexive ;
- $A \subseteq B \wedge B \subseteq A \implies A = B$: on dit que la relation \subseteq est antisymétrique ;
- $A \subseteq B \wedge B \subseteq C \implies A \subseteq C$: on dit que la relation \subseteq est transitive.

Théorème 2 - 1

la relation d'ordre \subseteq

Ces trois propriétés étant vérifiées, on dit que \subseteq est une **relation d'ordre** sur $\mathcal{P}(E)$.

Vous vérifierez que la relation \leq sur \mathbb{R} a exactement les mêmes propriétés : vous commencez à entrevoir l'essence de l'algèbre qui est de dégager et d'étudier des structures les plus générales possibles afin de simplifier l'étude de phénomènes qui paraissent différents.

2 2 Loi de composition interne

Définition 2 - 5

réunion d'ensembles

La réunion de deux ensembles A et B notée $A \cup B$ est l'ensemble des éléments appartenant à A **OU** B.

Le « ou » correspond bien sûr à celui de la définition ?? page ??, c'est-à-dire qu'il est inclusif.

- Comme $A \cup B$ est encore un élément de $\mathcal{P}(E)$, on dit que \cup est une **loi de composition interne** ;
- comme $A \cup (B \cap C) = (A \cup B) \cap C$, on dit que la loi \cup est **associative** ;
- comme $A \cup B = B \cup A$, on dit que la loi \cup est **commutative** ;
- comme $A \cup \emptyset = A$, on dit que \emptyset est un **élément neutre** pour la loi \cup ;
- comme $A \cup A = A$, on dit que la loi \cup est **idempotente**.

On peut remarquer que ces propriétés sont vérifiées par l'addition sur \mathbb{N} , l'élément neutre étant 0.

Définition 2 - 6

intersection d'ensembles

L'intersection de deux ensembles A et B notée $A \cap B$ est l'ensemble des éléments appartenant à A **ET** B.

Vous vérifierez que la loi \cap a les mêmes propriétés que \cup avec cette fois comme élément neutre E.

2 3 Lien avec les opérateurs logiques

Comme souvent en mathématiques, on retrouve des structures semblables dans des domaines différents. Nous avons déjà découvert les liens entre \cup et OU puis entre \cap et ET.

On peut également faire le rapprochement entre l'implication des propositions et l'inclusion des ensembles, « vrai » correspondant « \in ».

Ainsi, si deux ensembles A et B vérifient $A \subseteq B$, alors il est impossible de trouver un élément x tel que $x \in A$ et $x \notin B$. Tous les autres cas sont possibles.

De même, si $A \implies B$, il est impossible que A soit vraie et B soit fausse. Toutes les autres propositions sont vraies.

On fait le même rapprochement entre l'équivalence des propositions et l'égalité des ensembles. Merveilleuses mathématiques...

2 4 Partition

On a souvent besoin de découper un ensemble en tranches (souvenez-vous de votre cours de terminale sur les probabilités totales...). Voici une définition utile :

Définition 2 - 7

Partition d'un ensemble

La famille $(A_i)_{i \in [1, n]}$ de parties non vides d'un ensemble E réalise une **partition** de cet ensemble si, et seulement si :

$$\forall i \neq j \ A_i \cap A_j = \emptyset \quad \text{et} \quad \bigcup_{i=1}^n A_i = E$$

Par exemple, une partie A de E et son complémentaire dans E réalise une partition de E car $A \cap \bar{A} = \emptyset$ et $A \cup \bar{A} = E$.

2 5 Produit cartésien

Définition 2 - 8

Produit cartésien

Soit $(A_i)_{i \in [1, n]}$ des parties d'un ensemble E. Le produit $\prod_{i=1}^n A_i = A_1 \times A_2 \times \dots \times A_n$ est l'ensemble des n-listes (x_1, x_2, \dots, x_n) telles que $x_1 \in A_1, \dots, x_n \in A_n$.

Soit $A = \{1, 3, 5, 7, 9\}$ et $B = \{0, 2, 4, 6, 8\}$. Alors $(1, 0)$ est un élément de $A \times B$ et $(0, 1)$ est un élément de $B \times A$. L'ordre des ensembles dans le produit et donc des éléments dans les n-listes est important !

3 Quelques résultats sur les cardinaux

3 1 Cardinal de $\mathcal{P}(E)$

Soit $E_2 = \{a, b\}$ et $E_3 = \{a, b, c\}$. En fait, $E_3 = E_2 \cup \{c\}$. Les parties de E_3 sont donc les parties de E_2 et ces mêmes parties auxquelles on adjoint c.

En effet, $\mathcal{P}(E_3) = \{\emptyset, \{a\}, \{b\}, \{a, b\}, \emptyset \cup \{c\}, \{a\} \cup \{c\}, \{b\} \cup \{c\}, \{a, b\} \cup \{c\}\}$.

Ainsi $|\mathcal{P}(E_3)| = 2 \times |\mathcal{P}(E_2)|$.

Notons de manière générale E_n un ensemble à n éléments avec $n \in \mathbb{N}$ et $E_{n+1} = E_n \cup \{\alpha\}$ avec $\alpha \notin E_n$. On montre facilement que $|\mathcal{P}(E_{n+1})| = 2 \times |\mathcal{P}(E_n)|$. Or $\mathcal{P}(E_0) = \{\emptyset\}$ donc $|\mathcal{P}(E_0)| = 1$.

une rapide récurrence permet donc de prouver que :

Théorème 2 - 2

Cardinal de $\mathcal{P}(E)$

Si $|E| = n$ alors $|\mathcal{P}(E)| = 2^n$

3 2 Cardinal d'une partition

Il est assez immédiat d'obtenir le résultat suivant :

Cardinal d'une partition

Si la famille $(A_i)_{i \in [1, n]}$ de parties non vides d'un ensemble E réalise une **partition** de cet ensemble alors :

Théorème 2 - 3

$$|E| = \sum_{i=1}^n |A_i|$$

Pour compter les élèves du lycée, il suffit de compter les élèves de chaque classe et d'additionner les résultats...

3 3 Cardinaux et inclusion

Théorème 2 - 4

$\forall (A, B) \in \mathcal{P}(E) \times \mathcal{P}(E), A \subseteq B \implies |A| \leq |B|$

Pour le prouver, il suffit d'utiliser une petite ruse (bientôt) habituelle : A et $\overline{A} \cap B$ forment une partition de B donc, d'après le théorème 2 - 3, $|A| + |\overline{A} \cap B| = |B|$, donc $|A| \leq |B|$

Cette application n'est pas une équivalence ! Prenez $A = \{1\}$ et $B = \{3; 4\}$

3 4 Cardinal d'un produit cartésien

Voici un autre résultat qui sera bien utile en probabilités :

Cardinal d'un produit cartésien

Soit $(A_i)_{i \in [1, n]}$ des parties d'un ensemble E .

Théorème 2 - 5

$$\left| \prod_{i=1}^n A_i \right| = \prod_{i=1}^n |A_i|$$

En effet, pour chaque n -liste (x_1, \dots, x_n) , il y a $|A_1|$ choix pour la première composante, $|A_2|$ choix pour la deuxième, etc.

3 5 Formule du crible

Nous n'en verrons qu'un cas particulier que vous devez déjà connaître :

Théorème 2 - 6

Formule du crible pour deux ensembles

$\forall (A, B) \in \mathcal{P}(E) \times \mathcal{P}(E), |A \cup B| = |A| + |B| - |A \cap B|$

Pour le démontrer, nous allons former une partition de $A \cup B$ avec $A \cap \overline{B}$, $\overline{A} \cap B$ et $A \cap B$.

Mais on peut également « partitionner » A avec $A \cap \overline{B}$ et $A \cap B$ puis « partitionner » B avec $B \cap \overline{A}$ et $B \cap A$.

On obtient donc que $|A| = |A \cap \overline{B}| + |A \cap B|$ puis que $|B| = |B \cap \overline{A}| + |B \cap A|$.

On en déduit que $|A| + |B| = |A \cap \overline{B}| + |B \cap \overline{A}| + |A \cap B| + |B \cap A|$ et enfin que :

$$|A| + |B| - |A \cap B| = |A \cap \overline{B}| + |B \cap \overline{A}| + |A \cap B| = |A \cup B|$$

4 Dénombrement

4 1 Nombre de permutations

Définition 2 - 9

Permutation

Soit E un ensemble. Une permutation de E est une application bijective de E sur E . On note $\mathfrak{S}(E)$ l'ensemble des permutations de E .

Généralement, on note les permutations sous forme d'une matrice où la première ligne correspond aux éléments de E et où la deuxième ligne correspond aux images des éléments de E .

Par exemple, les permutations de $\{a, b, c\}$ sont :

$$\begin{pmatrix} a & b & c \\ a & b & c \end{pmatrix} \begin{pmatrix} a & b & c \\ a & c & b \end{pmatrix} \begin{pmatrix} a & b & c \\ b & a & c \end{pmatrix} \begin{pmatrix} a & b & c \\ b & c & a \end{pmatrix} \begin{pmatrix} a & b & c \\ c & a & b \end{pmatrix} \begin{pmatrix} a & b & c \\ c & b & a \end{pmatrix}$$

On peut même se contenter d'écrire (a, b, c) , (a, c, b) , (b, a, c) , (b, c, a) , (c, a, b) , (c, b, a) si l'on est sûr que la première ligne est toujours (a, b, c) .

Théorème 2 - 7

Nombre de permutations

Il y a $n!$ permutations d'un ensemble de n éléments.

La démonstration s'effectue par récurrence. Il y a bien sûr $1!$ façon de permuter 1 élément.

On peut donc supposer qu'il existe au moins un entier k tel qu'il y ait $k!$ permutations d'un ensemble à k éléments.

Considérons un ensemble à $k + 1$ éléments. On en choisit 1 : il y a $k + 1$ choix possibles. Ensuite, il s'agit d'ordonner les k éléments restant ce qui laisse $k!$ possibilités.

Il y a donc finalement $(k + 1) \times k! = (k + 1)!$ permutations au total.

On retrouve dans l'exemple précédent qu'il y a $3! = 6$ permutations de $\{a, b, c\}$.

On conviendra que $0! = 1$.

On a déjà rencontré les permutations lors de notre découverte de la cryptographie. Elles jouent également un rôle important dans les algorithmes de tri par exemple et aussi dans certains algorithmes concernant les graphes que vous étudierez en deuxième année, dans les mélanges de cartes, dans l'étude du Rubik's cube, et bien d'autres domaines encore...

Remarque

Formule de STIRLING

Pour n « grand », on peut estimer que :

$$n! \underset{+\infty}{\sim} \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

4 2 Nombre d'applications injectives - arrangements sans répétitions

Une source veut envoyer à chacun de ses cinq clients un fichier différent parmi les vingt-six situés dans un certain répertoire.

Combien de possibilités a la source pour envoyer un fichier à chacun des 26 clients ?

On peut modéliser cette situation : il s'agit de compter les applications injectives (pourquoi « injectives » ?) de l'ensemble E des clients vers l'ensemble F des fichiers. Il ne peut y avoir répétition donc il y a 26 choix de fichiers pour le premier client, 25 pour le deuxième, etc. ce qui donne $26 \times 25 \times 24 \times 23 \times 22 \times 21$ possibilités.

On peut généraliser.

Définition 2 - 10

Arrangement sans répétition

Une application injective d'un ensemble E de cardinal p dans un ensemble F de cardinal n (avec $p \leq n$) est appelée un arrangement sans répétition de p parmi n .

Un exemple fameux est la recherche de tiercés dans l'ordre (sans ex-æquo...) : pourquoi ?

Nous pouvons formuler cette propriété plus synthétiquement. En effet

$$A_n^p = n(n-1)(n-2)\cdots(n-p+1) \times \frac{(n-p)(n-p-1)(n-p-2)\cdots \times 2 \times 1}{(n-p)(n-p-1)(n-p-2)\cdots \times 2 \times 1} = \frac{n!}{(n-p)!}$$

d'où

Nombre d'arrangements

Le nombre d'arrangements sans répétition de p parmi n est

Théorème 2 - 8

$$A_n^p = n(n-1)(n-2)\cdots(n-p+1) = \frac{n!}{(n-p)!}$$

Preuve ?

4 3 Combinaisons sans répétition

Il n'y a pas assez de gagnants au loto. Les règles en sont donc modifiées. Il s'agit maintenant de trouver 3 numéros parmi 5. Combien y a-t-il de grilles (combinaisons) possibles ?

Par exemple, on pourrait dire que j'ai cinq manières de choisir le premier numéro, quatre choix pour le deuxième et trois choix pour le troisième, donc il y a $5 \times 4 \times 3$ grilles différentes mais dans ce cas, je compte des 3-listes ordonnées alors que les 3-listes (1, 2, 3) et (3, 2, 1) correspondent à la même grille ou combinaison $\{1, 2, 3\}$.

Posons une petite définition pour clarifier les débats. Donnons en fait un nom à une grille du loto, c'est à dire à un sous-ensemble (une partie) contenant p éléments d'un plus grand ensemble contenant n éléments.

Combinaison sans répétition

Soit n et p deux entiers naturels et E un ensemble contenant n éléments. Un sous-ensemble de E contenant p éléments est appelé une **combinaison** de p éléments de E ou encore une **p -combinaison** d'éléments de E .

Définition 2 - 11

Or ce qui nous intéresse, c'est le nombre de ces combinaisons, donc introduisons une notation :

Nombre de combinaisons sans répétitions

Le nombre de p -combinaisons d'un ensemble E contenant n éléments est noté $\binom{E}{p}$ ou encore $C_{|E|}^p = C_n^p$

Définition 2 - 12

Revenons à notre mini-loto. Considérons une grille quelconque (i.e. une 3-combinaison de l'ensemble des 5 numéros) : par exemple $\{2, 4, 5\}$. Nous avons vu dans un paragraphe précédent qu'il y a 3! façons d'ordonner ces nombres. Finalement, il y a $C_5^3 \times 3!$ suites de 3 nombres ordonnées : c'est le nombre d'arrangements de p parmi n . Or nous en avons comptées $5 \times 4 \times 3$ tout à l'heure. Nous en déduisons finalement que

$$C_5^3 = \frac{A_5^3}{3!} = \frac{5 \times 4 \times 3}{3!}$$

Il est alors possible de généraliser la formule suivante :

Théorème 2 - 9

$$C_n^p = \frac{A_n^p}{p!} = \frac{n(n-1)(n-2)\cdots(n-(p-1))}{p!} = \frac{n!}{p!(n-p)!}$$

Preuve ?

5 Triangle de Pascal - Binôme de Newton

À l'aide des formules précédentes, on prouve facilement (faites-le !) le résultat suivant :

Théorème 2 - 10

$$C_n^p = C_n^{n-p}$$

On établit ensuite, toujours par le calcul, la relation suivante, dite **Relation de Pascal** même si les mathématiciens chinois l'avaient mise en évidence avant lui :

Relation de PASCAL

Théorème 2 - 11

$$C_n^p = C_{n-1}^p + C_{n-1}^{p-1}$$

Il est possible de démontrer cette formule à l'aide d'un raisonnement ensembliste : pour former des groupes de p éléments dans un ensemble en contenant n , on distingue un élément n_0 quelconque :



- soit le groupe le contient et alors il y a C_{n-1}^{p-1} choix des $p-1$ éléments distincts de n_0 parmi les $n-1$ restant ;
- soit le groupe ne le contient pas et il faut donc choisir p éléments parmi les $n-1$ éléments distincts de n_0 : il y a donc cette fois C_{n-1}^p .

Cette relation permet de démontrer la formule du binôme de Newton (que nous admettrons) :

Formule du binôme

Soit \mathbb{A} un anneau commutatif. Dans $\mathbb{A}[X]$, pour tout entier naturel non nul n :

$$(1 + X)^n = \sum_{k=0}^n C_n^k X^k$$

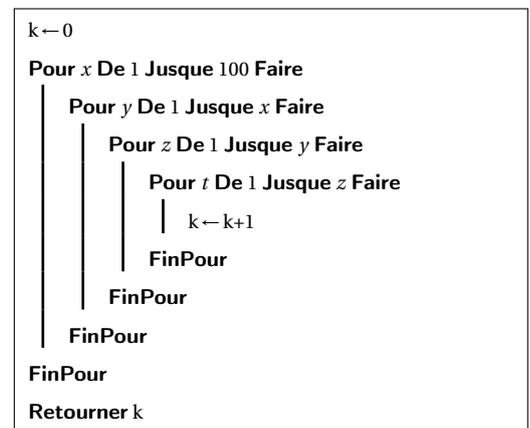
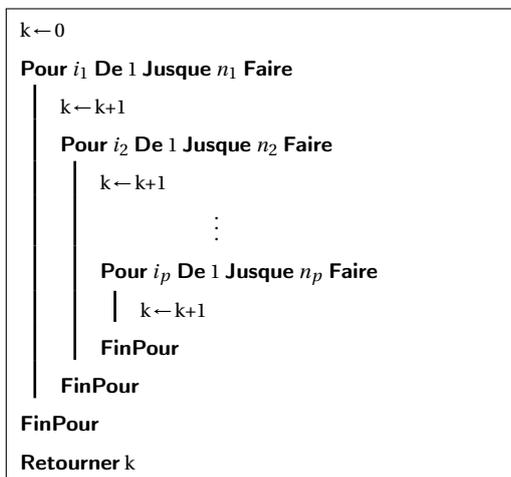
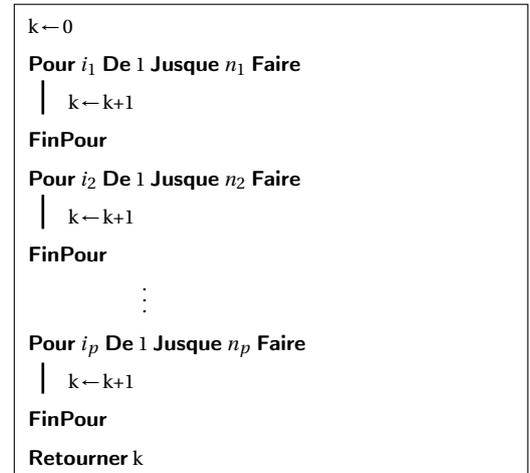
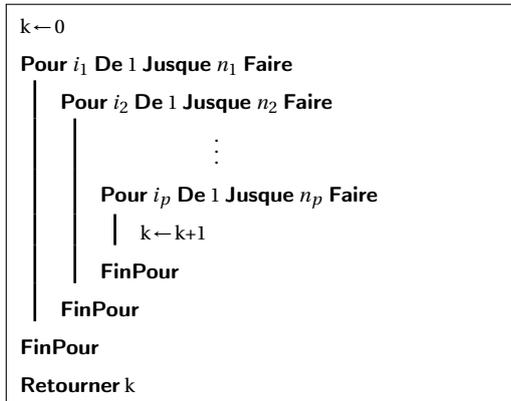
et plus généralement :

$$\forall (a, b) \in \mathbb{A}^2, \quad (a + b)^n = \sum_{k=0}^n C_n^k a^k b^{n-k}$$

Théorème 2 - 12

Exercice 2 - 1 Pour

Que vaut k à la fin ?

**Exercice 2 - 2 Chaînes**

Combien y a-t-il d'octets commençant par 1 ou se terminant par 00 ? Combien y en a-t-il ne contenant pas deux 1 consécutifs ? Combien y en a-t-il qui contiennent la chaîne 000 ou la chaîne 111 ?

Exercice 2 - 3 Principe des tiroirs

Si on a $n + 1$ ou plus chaussettes à ranger dans n tiroirs, il y a au moins un tiroir qui en contient 2 ou plus. Cette affirmation assez simple est aussi appelé *Principe de DIRICHLET* ou même *pigeon hole principle* par les anglo-saxons.

Démontrez le *principe des tiroirs généralisé* : si on a n chaussettes à ranger dans p tiroirs, alors il existe au moins un tiroir contenant au moins $\left\lceil \frac{n}{p} \right\rceil$ chaussettes. Vous pourrez raisonner par l'absurde

Exercice 2 - 4 Téléphone

On suppose que les numéros de téléphone sont de la forme 0Z-YX-XX-XX-XX avec X un nombre entier entre 0 et 9, Y un nombre entier entre 2 et 9 et Z un entier entre 1 et 9.

En France, il y a environ 40 millions d'abonnements à un « fixe » et 69 millions à un « mobile » (données du quatrième trimestre 2011).

Est-on tranquille ?

Supposons qu'en Syldavie, il y ait 109 millions d'abonnés à un téléphone fixe et que le Y soit fixé par opérateur et qu'il n'y ait que deux opérateurs. De combien de zones a-t-on besoin ?

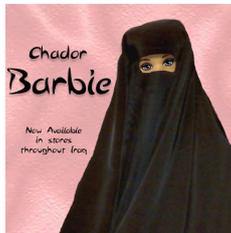
Exercice 2 - 5 Câbles

Il y a 1000 ordinateurs et 50 imprimantes. Combien faut-il au minimum de câbles pour que 50 ordinateurs pris au hasard puissent, à tout moment, accéder chacun à une imprimante différente ?

Exercice 2 - 6 Nombre d'applications

Combien y a-t-il d'applications d'un ensemble de cardinal n vers un ensemble de cardinal p ?

Exercice 2 - 7 Barbie



Un étudiant de l'IUT d'informatique de Klow en Syldavie a reçu pour Noël sept Barbie et sept séries de dix robes de couleurs différentes. Il habille chaque Barbie d'une de ces robes.

1. Combien y a-t-il d'habillages possibles selon la couleur ?
2. Combien y a-t-il d'habillages de sorte que chaque Barbie ait une robe de couleur différente des autres ?
3. Combien y a-t-il d'habillages de sorte qu'au moins deux Barbie aient une robe de la même couleur ?
4. Combien y a-t-il d'habillages de sorte qu'exactement deux Barbie aient une robe de la même couleur ?

Exercice 2 - 8 Jus

Lors de la soirée d'intégration de l'IUT de Klow, 9 étudiants ont bu un verre d'eau, 25 un verre de jus de tomate, 14 un verre de jus de betterave, 7 un verre de jus de tomate et un verre d'eau, 4 un verre d'eau et un verre de jus de betterave, 10 un verre de jus de betterave et un verre de jus de tomate. Les étudiants d'Info s'occupaient des boissons et ceux de GEA des entrées or ces derniers ont perdu leur liste et ne savent plus combien d'étudiants ont participé à la soirée. Les étudiants d'Info, connaissant les statistiques sur les consommations de boissons, pourront-ils corriger la gaffe des GEA sachant également que lors de la valse, tous les participants dansaient en couple ?

Exercice 2 - 9 k-uplets

Soit E un ensemble fini non vide de cardinal n .

- a) Combien y a-t-il de parties de E formées de k éléments ?
- b) Combien y a-t-il de k -uplets d'éléments de E ?
- c) Combien y a-t-il de k -uplets d'éléments deux à deux distincts de E ?
- d) Combien y a-t-il de k -uplets d'éléments deux à deux distincts de E , tel que le premier élément est le plus petit et le dernier élément est le plus grand ?
- e) Combien y a-t-il de k -uplets d'éléments de E ordonnés dans l'ordre strictement croissant ?

Exercice 2 - 10 Arrangements

Comment modéliser le tiercé brillamment organisé par le PMU ? (Qui est E ? Qui est F ? Quelle est l'application de E dans F ?) Combien y a-t-il de tiercés possibles (sans ex-æquo) dans une course de 19 partants ? Modélisez la recherche du nombre d'arrangements de p parmi n à l'aide d'un tirage de boules dans une urne.

Exercice 2 - 11 L'âge du capitaine

Le capitaine des pompiers de New-York est marié et a quatre enfants dont un mange un yaourt aux fruits tous les matins. Il ne fume pas, aime regarder des films de gladiateurs et réside à l'angle de la 1^{ère} avenue et de la 33^{ème} rue. La caserne se trouve à l'angle de la 9^{ème} avenue et de la 40^{ème} rue. Il s'y rend tous les jours à pied en sifflant « $\text{\textcircled{D}}$ du schöner Westerland » et sans perdre de temps (i.e. dans le sens des numéros croissants aussi bien pour les rues que pour les avenues). Sachant qu'il a commencé à travailler le jour de ses 18 ans, et sachant qu'il n'est jamais passé deux fois par le même chemin, qu'il boîte légèrement de la jambe droite après avoir participé au championnat d'Écosse de lancement d'enclume, qu'il est sourd de l'oreille gauche depuis qu'il a plongé dans un lac gelé pour sauver son petit frère John qui était en train de se noyer après que la glace sur laquelle il patinait craqua, quel est l'âge (maximum) du capitaine ?

Exercice 2 - 12 Stirling strikes back

Donnez un équivalent de C_{2n}^n quand n tend vers l'infini.

Exercice 2 - 13 Mélange de cartes



On dispose d'un jeu de 52 cartes et on le mélange en *queue d'aronde* : on divise le jeu en deux parties de 26 cartes et on intercale les cartes de la moitié gauche dans la moitié droite sans modifier l'ordre de chaque

partie bien sûr. Est-ce que quatre mélanges en queue d'aronde successifs permettent de fournir un ordre aléatoire des cartes ?

Il faudrait préciser d'abord cette notion : combien y a-t-il d'ordres possibles des cartes ?

Existe-t-il alors des ordres qui ne pourront être obtenus par notre succession de mélanges ? Pour répondre à cette question, il faudrait calculer le nombre d'ordres qu'on peut obtenir par la succession de quatre mélanges... Imaginez que vous disposiez de 26 cartes bleues et 26 rouges. Vous les mélangez avec la méthode de la queue d'aronde : pouvez-vous reconstituer le jeu initial ? Combien y a-t-il alors d'ordres possibles après le premier mélange ?



Exercice 2 - 14 Poker

Une main au poker est constituée de 5 cartes tirées d'un jeu de 52 cartes. Combien y a-t-il de mains contenant des carrés (XXXXY) ? des fulls (XXXYY) ? des brelans (XXXYZ) ? des doubles paires (XXYYZ) ? des paires (XXYZA) ?

Deux lettres identiques (par exemple XX) correspondent à deux cartes de même hauteur (par exemple deux dames).

Exercice 2 - 15 Binôme

Soit E un ensemble de cardinal n . Si on remplace a et b par 1 dans la formule du binôme, quel résultat classique retrouve-t-on ?

Exercice 2 - 16 Bibosses

1. En utilisant la relation de PASCAL, déterminez une fonction CAML récursive de signature :

```
val binom1 : int * int -> int = <fun>
```

telle que $\text{binom1}(n, p)$ calcule C_n^p .

On pourra alors afficher le triangle de PASCAL à l'aide de la fonction suivante :

```
let triangle1 n =
  for i = 0 to n do
    for j = 0 to i do
      print_int (binom1(i,j));
      print_string " ";
    done;
    print_newline ();
  done;;
```

2. Calculez C_{30}^{10}mmmmmmmm....c'est un peu long. Une autre idée est d'utiliser un tableau a priori rempli de 0. On va remplacer certaines cellules en place en utilisant encore la relation de PASCAL.

Avec notre chameau, on utilise la fonction `Array.make_matrix n m` qui construit une matrice de taille $n \times m$. On accède à l'élément situé sur la ligne i et la colonne j avec `t.(i).(j)` si on a nommé `t` la matrice.

On rappelle qu'on peut substituer la valeur initiale de la cellule t_{ij} par la valeur v en place à l'aide de la commande :

```
t.(i).(j) <- v
```

On créera donc une fonction `triangle2 n` de signature :

```
val triangle2 : int -> int array array = <fun>
```

puis une fonction `binom(n, p)` :

```
let binom(n,p) =
  let b = triangle2 n in
  b.(n).(p);;
```

Calculez alors C_{30}^{10} avec cette nouvelle fonction....c'est mieux!

3. On voudrait donc comparer la complexité en temps de ces fonctions en prenant comme unité une addition de coefficients binomiaux (on négligera les décréments d'indices). Faites-le !...

4. Une petite touche artistique à présent... Dans le triangle de PASCAL, on va remplacer les coefficients pairs par une espace et les impairs par une étoile.

Créez une fonction `triangle4` : `int -> unit = <fun>` à cet effet. Que donne `triangle4 n` avec `n` égal successivement à 31, 63, 127, 255 ?

Exercice 2 - 17 Formules binomiales

On voudrait calculer $\sum_{k=0}^n \binom{n}{k}^2$ à l'aide du calcul d'un seul coefficient binomial (i.e. déterminer i et j tels que

la somme précédente soit égale à C_i^j).

Plusieurs méthodes existent bien sûr.

Une consiste à avoir un raisonnement ensembliste similaire à celui utilisé pour prouver la relation de PASCAL.

Vous déterminerez d'abord le lien entre $\binom{n}{k}^2$ et $C_n^k C_n^{n-k}$ puis vous ferez le lien entre notre problème et la situation suivante : on a n filles et n garçons ; combien de groupes de n étudiants peut-on faire ?...

Une autre idée consiste à « penser polynôme » :

$$(1 + X)^n (1 + X)^n = (1 + X)^{2n}$$

Des idées ?...

Exercice 2 - 18 Suissesses multinomiales

Combien existe-il d'anagrammes au mot « suissesses » ?

On définit un nouveau type de coefficient :

Coefficient multinomial

Soit k_1, k_2, \dots, k_p des entiers naturels tels que $\sum_{i=1}^p k_i = n$. On note alors :

$$C_n^{k_1, k_2, \dots, k_p} = \frac{n!}{k_1! k_2! \dots k_p!}$$

Définition 2 - 13



Quel est le coefficient multinomial égal à C_n^p ?

Comment interpréter ces coefficients en termes combinatoires ?

Quel est le rapport avec les suissesses ?

Il existe un pendant multinomial à la formule du binôme :

Formule du multinôme

$$(x_1 + x_2 + \dots + x_p)^n = \sum_{k_1+k_2+\dots+k_p=n} C_n^{k_1, k_2, \dots, k_p} x_1^{k_1} x_2^{k_2} \dots x_p^{k_p}$$

Théorème 2 - 13

Par exemple, quel est le coefficient de $x^2 y^3 z^4$ dans $(x + y + z)^8$?

Exercice 2 - 19 Somme d'entiers

Prouvez la formule : $C_{n+1}^{p+1} = \sum_{i=0}^{n-p} C_{p+i}^p$ Déduisez-en $\sum_{k=1}^n k$, $\sum_{k=2}^n k(k-1)$, $\sum_{k=1}^n k^2$, $\sum_{k=1}^n k^3$.

Exercice 2 - 20 Encore une formule binomiale

On considère n boules numérotées de 1 à n .

1. Soit k un entier vérifiant $p \leq k \leq n$. On tire simultanément p boules de l'urne. Combien y a-t-il de tirages dont le plus grand numéro est k ?
2. En déduire une expression de $\sum_{k=p}^n C_{k-1}^{p-1}$ sous forme d'un unique coefficient binomial.

Exercice 2 - 21 Les boules

1. Démontrez que pour tous entiers naturels n et k tels que $2 \leq k < n - 1$, on a :

$$C_{n-2}^{k-2} + 2C_{n-2}^{k-1} + C_{n-2}^k = C_n^k$$

2. On considère deux entiers naturels n et k tels que $2 \leq k < n - 1$. On dispose d'une urne contenant n boules indiscernables au toucher. Deux des boules sont rouges, les autres sont blanches. On tire au hasard et simultanément k boules de l'urne. On appelle A l'évènement « au moins une boule rouge a été tirée ».
- Calculez de deux manières différentes le cardinal de A .

Exercice 2 - 22 Pesées



On dispose de 28 pièces de monnaie dont une est fautive et a une masse légèrement supérieure aux autres. Nous disposons uniquement d'une balance de Roberval et des 28 pièces. En combien de pesées peut-on déterminer la fautive pièce ?

Exercice 2 - 23 Sommes et cardinaux

Soit E un ensemble fini de cardinal n . Déterminer les sommes suivantes en fonction de n :

a) $\sum_{A \in \mathcal{P}(E)} \text{Card} A.$ b) $\sum_{(A,B) \in \mathcal{P}(E)^2} \text{Card}(A \cap B).$ c) $\sum_{(A,B) \in \mathcal{P}(E)^2} \text{Card}(A \cup B).$

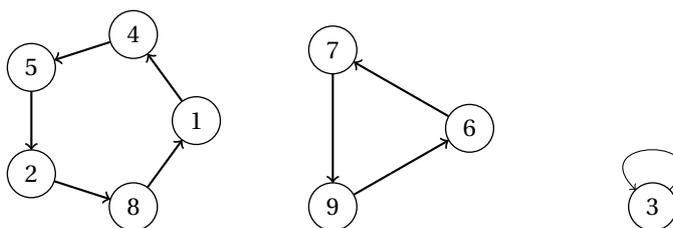
Exercice 2 - 24 Permutations

On peut être amené à rechercher les *cycles* d'une permutations. Il est aisé de les représenter à l'aide de *graphes* orientés...cycliques.

Par exemple, si l'on considère la permutation :

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 8 & 3 & 5 & 2 & 9 & 6 & 1 & 7 \end{pmatrix}$$

alors on trouve trois cycles :



- Décomposez $\pi = (2\ 1\ 3\ 5\ 4\ 7\ 9\ 6\ 8)$ en cycles.
- La *longueur* d'un cycle est le nombre de ses sommets. Soit p une permutation. On note $p^k = p \circ p^{k-1}$ avec $p^0 = \text{id}$, la permutation identité. Soit $p \in \mathfrak{S}_n$. On définit une relation \mathcal{R} sur $E = \{1, 2, \dots, n\}$ par :

$$i \mathcal{R} j \Leftrightarrow \exists k \in \mathbb{N}^*, p^k(i) = j$$

Est-ce que \mathcal{R} est une relation d'équivalence ? Quels sont ses classes ?

- L'*ordre* d'une permutation est le plus petit entier naturel strictement positif k tel que $p^k = \text{id}$. Déterminez l'ordre de π . Est-ce que toute permutation a un ordre fini ? Comment le calculer ?

Exercice 2 - 25 Permutations de bosses

Voici une série de fonctions CAML : pourriez-vous les spécifier ?

Sachez que la fonction `int(n)` du module `Random` renvoie un entier aléatoire compris entre 0 et $n - 1$ compris, que `Array.make n 1` crée un tableau de longueur n rempli de 1.

```
let f1(n) =
  let t = Array.make n 1 in
  for k = 0 to n-1 do
    t.(k) <- k+1;
  done;
  t;;

let f2(t, a, b) =
  let c = t.(a) in
  t.(a) <- t.(b);
  t.(b) <- c;
```

```

t;;

open Random;;

let hasard(a,b) = a + int(b-a+1);;

let f3(n) =
  let t = ref (f1(n)) in
  for k = 0 to n-1 do
    t := f2(!t,k,hasard(k,n-1));
  done;
  !t;;

```

Voici une autre série de fonctions à spécifier. Sachez que `Array.map` f t applique la fonction f aux éléments du tableau t , que `Array.to_list` transforme un tableau en liste.

```

let f4 p r =
  let c = ref [r] in
  let i = ref r in
  while p.(!i-1) <> r do
    i := p.(!i-1);
    c := !c @ [!i];
  done;
  !c;;

let f5 p =
  let n = (Array.length p) in
  let l = ref (f4 p 1) in
  let lc = ref [] !l [] in
  let j = ref 1 in
  while !j < n do
    while (List.mem (p.(!j-1)) !l) && (!j < n) do
      j := !j + 1;
    done;
    if not (List.mem !j !l) then
      lc := Array.append !lc [] f4 p !j [];
      l := (f4 p !j) @ !l;
      j := !j + 1;
    done;
  !lc;;

let f6 p =
  Array.map List.length (f5 p);;

let rec f7(a,b) =
  if a >= b then
    if b = 0 then
      a
    else
      f7(b,a mod b)
  else
    f7(b,a);;

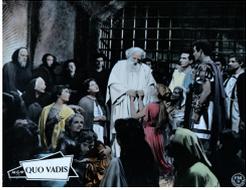
let f8(a,b) =
  (a*b)/f7(a,b);;

let rec f9 = function
  | [] -> 0
  | tete::[] -> tete
  | a::[b] -> f8(a,b)
  | tete::queue -> f8(tete,f9 queue);;

let f10 p =
  f9 (Array.to_list(f6 p));;

```

Exercice 2 - 26 Flavius Josèphe



Flavius Josèphe (37 - 100) ou plutôt Yossef ben Matityahou HaCohen est un historien romain d'origine juive et de langue grecque et qui ne devait pas être trop mauvais en mathématiques si on en croit la sinistre anecdote suivante. Lors de la première guerre judéo-romaine, Yossef fut piégé dans une grotte avec 39 autres de ses compagnons en juillet 67. Ne voulant pas devenir esclaves, ils mirent au point un algorithme d'auto-destruction : il s'agissait de se mettre en cercle et de se numéroter de 1 à 40. Chaque septième devait être tué jusqu'à ce qu'il n'en reste plus qu'un qui devait alors se suicider. Ce dernier fut Yossef lui-même...et il ne se suicida pas ! Après deux ans de prison, il fut libéré, il entra au service des romains comme interprète et acquies la citoyenneté romaine deux ans plus tard. Quel numéro portait Yossef ? Créez un programme CAML qui donne l'ordre des exécutions quelque soit le nombre de prisonniers et le chiffre sinistre.

3

Espace probabilisé discret

« Le nom seul de calcul des probabilités est un paradoxe : la probabilité, opposée à la certitude, c'est ce qu'on ne sait pas, et comment peut-on calculer ce que l'on ne connaît pas ? »

Henri POINCARÉ - La Science et l'hypothèse (1908)

« [La théorie des probabilités] donne les aperçus les plus sûrs qui puissent nous guider dans nos jugements. [...] Il n'est point de science plus digne de nos méditations, et qu'il soit plus utile de faire entrer dans le système de l'instruction publique. »

Pierre Simon de LAPLACE - Essai philosophique sur les probabilités (rédigé de 1795 à 1825)

« The true logic of this world lies in the calculus of probabilities »

James C. MAXWELL - The Scientific Letters and Papers of James C. Maxwell, Vol. 1, 1846-1862

« We cannot predict whether a given photon will arrive at A or B. All we can predict is that out of 100 photons that come down, an average of 4 will be reflected by the front surface. Does this mean that physics, a science of great exactitude, has been reduced to calculating only the probability of an event, and not predicting exactly what will happen? Yes. That's a retreat, but that's the way it is : Nature permits us to calculate only probabilities. Yet

INCROYABLE!

RECALÉ ALI BAC,
IL JOUE SES NOTES
ALI LOTO ET GAGNE
37 MILLIONS!



1 Probabilités ?

Le calcul des probabilités est très important en informatique et dans toute science en général. Il nous permet de résoudre des problèmes tout en étant confrontés à l'incertitude ou à des systèmes très complexes.

Contrairement aux clichés, la théorie des probabilités, en tant que discipline mathématique, est dotée de toute la rigueur et de la précision possibles. Elle puise ses sources dans tous les champs des mathématiques. Il n'est dès lors pas étonnant de constater que les prix mathématiques les plus prestigieux sont attribués depuis quelque temps à des travaux liés aux probabilités.

En informatique, leur usage est proportionnel à la complexité des problèmes abordés.

Un informaticien dirait qu'en probabilité le modèle est totalement spécifié, permet et impose la plus grande rigueur suite à un raisonnement déductif.

Il ne faut pas confondre les probabilités avec la statistique qui complète un modèle inconnu à l'aide d'observations (raisonnement inductif) et fait plutôt appel à l'intuition qui n'est pas la principale qualité d'un ordinateur...

Enfin, s'il est tout à fait possible de calculer des probabilités sans faire de statistique, le contraire est quasiment impossible c'est pourquoi nous aborderons la statistique inférentielle après avoir étudié les probabilités.

2 Avant la formalisation

Essayons, grâce à notre petite expérience probabiliste acquise au lycée, de dégager les grandes lignes de ce que pourrait être une théorie formalisée, nécessaire au « dialogue » avec une machine.

Lors de l'étude des probabilités, on est amené à effectuer une **expérience aléatoire**, c'est-à-dire un processus dont on ne peut prévoir à l'avance le résultat qui peut donner des résultats différents même si l'on répète l'expérience dans des conditions qui semblent identiques.

Nous allons considérer par la suite le lancement d'un dé cubique équilibré et tout et tout.

On doit d'abord connaître l'ensemble des résultats possibles : c'est l'**univers** que nous noterons par la suite Ω .

Par exemple, si je lance une fois le dé, $\Omega_1 = \{1, 2, 3, 4, 5, 6\}$.

Si je lance deux fois de suite le dé et que je note dans l'ordre le numéro des faces, j'obtiens dans ce cas $\Omega_2 = \{11, 12, 13, \dots, 66\}$

Si je lance les dés jusqu'à obtenir deux 6 à la suite, le nombre de lancers n'est maintenant pas fixé à l'avance et l'univers est a priori infini mais en demeurant discret : chaque évènement peut en effet être « numéroté ». Plus rigoureusement, Ω_3 peut être mis en bijection avec \mathbb{N} : $\Omega_3 = \{1366, 3546213466, 55236221466, \dots\}$.

Il faut maintenant étudier la notion d'« évènement ^a ». Disons dans un premier temps qu'il s'agit d'une assertion vérifiable relativement au résultat d'une expérience aléatoire. Il est donc « associé » à un sous-ensemble (une partie) de Ω . Un élément de Ω sera associé à un **évènement élémentaire**.

Reprenons le deuxième exemple. Voici quelques évènements :

- « Obtenir deux fois 3 » : $E_a = \{33\}$. C'est un évènement élémentaire ;
- « Ne pas obtenir deux fois 3 » : $E_b = \Omega_2 \setminus E_a \subseteq \Omega_2$;
- « Obtenir deux faces dont la somme des numéros fait 11 » : $E_c = \{56, 65\} \subseteq \Omega_2$;
- « Obtenir deux faces dont le numéro est inférieur à 7 » : $E_d = \Omega_2$. Il s'agit donc de l'**évènement certain** ;
- « Obtenir deux faces dont la somme des numéros fait 37 » : $E_e = \emptyset$. Il s'agit cette fois de l'**évènement impossible**.

On comprend que l'univers lui-même et l'ensemble vide sont des évènements. On voudrait que la réunion et l'intersection d'évènements soit encore un évènement et que le complémentaire de la partie associée à un évènement définisse encore un évènement : les mathématiciens (français...) disent alors que l'ensemble de tous ces évènements forment une **tribu**. Dans le reste du monde, on parle plutôt de σ -algèbre de Boole : les ensembles concernés vérifient en effet certaines propriétés *algébriques*.

Il reste enfin à définir un moyen de **mesurer** les probabilités : c'est une sorte de degré de certitude correspondant à une échelle graduée de 0 à 1, 0 correspondant au degré de l'évènement impossible et 1 au degré de l'évènement certain. Il est important de comprendre que ce degré de certitude est donné **a priori** et permet donc de prévoir l'issue de l'expérience avant de la réaliser.

Il faudrait que cette mesure assure une cohérence entre les évènements et les degrés.

Par exemple, si $A \subseteq B$, on aimerait que $\mathbb{P}(A) \leq \mathbb{P}(B)$. On aimerait également que l'union de deux parties de l'univers se « transforme » en somme de leur degré de certitude, mais on « sent » bien que cela ne pourra être

a. Depuis 1990, l'Académie française préconise de mettre un accent grave au deuxième e d'évènement qui doit donc s'écrire évènement...

vrai que si les évènements sont totalement **incompatibles** : $A \cap B = \emptyset \implies \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$. On retrouve ici cette notion algébrique de « transport » des opérations : la notion d'*isomorphisme*.

Considérons pour illustrer ces dernières notions le premier exemple du lancer unique d'un dé.

Notre univers Ω_1 contient 6 éléments et prenons comme tribu naturellement associée les $2^6 = 64$ parties de Ω_1 qui forment l'ensemble $\mathcal{P}(\Omega_1)$.

Imposons pour notre modèle que la mesure du degré de certitude ne dépende pas des numéros des faces mais uniquement du cardinal de la partie considérée : chacun des six évènements élémentaires aura donc la même probabilité : $\mathbb{P}(1) = \mathbb{P}(2) = \dots = \mathbb{P}(6) = \frac{1}{6}$.

La probabilité d'obtenir un numéro pair sera donc, sachant que l'évènement est associé à $A = \{2, 4, 6\}$: $\mathbb{P}(A) = \mathbb{P}(2) + \mathbb{P}(4) + \mathbb{P}(6) = \frac{1}{2}$.

Mais si le dé est truqué ? Si l'univers est infini ? S'il n'est pas dénombrable ?

Il est temps de définir clairement et rigoureusement les notions afin de ne pas dire de bêtises...

3 Espace probabilisable - Espace probabilisé

Dans ce chapitre, Ω est un ensemble dénombrable. Nous le considérerons parfois fini.

Tribu

Une tribu \mathcal{T}_Ω définie sur un univers Ω est une partie de $\mathcal{P}(\Omega)$ qui vérifie :

1. $\Omega \in \mathcal{T}_\Omega$;
2. $A \in \mathcal{T}_\Omega \implies \complement_\Omega A \in \mathcal{T}_\Omega$;
3. $\forall I \subseteq \mathbb{N}, (A_i)_{i \in I} \in \mathcal{T}_\Omega^{\text{II}} \implies \bigcup_{i \in I} A_i \in \mathcal{T}_\Omega$.

Définition 3 - 1

Attention ! Une tribu est une partie de $\mathcal{P}(\Omega)$: c'est donc un ensemble de sous-ensembles de Ω !

Espace probabilisable

Soit \mathcal{T}_Ω une tribu définie sur l'univers Ω . On dit que $(\Omega, \mathcal{T}_\Omega)$ est un espace probabilisable (ou espace d'évènements). Les éléments de la tribu sont appelés **évènements**.

Définition 3 - 2

Attention ! Pour une certaine expérience aléatoire, il peut y avoir de nombreux espaces probabilisables différents.

Par exemple, en reprenant un des exemples introduits dans la section précédente :

- $(\Omega_1, \{\emptyset, \Omega_1\})$ est un espace probabilisable : on ne pourra s'intéresser qu'aux évènements « obtenir un numéro entre 1 et 6 » et « ne pas obtenir de numéro entre 1 et 6 » ce qui est un peu sommaire ;
- $(\Omega_1, \{\emptyset, A, \complement_\Omega A, \Omega_1\})$ avec A associé à « obtenir un 6 » ;
- $(\Omega_1, \mathcal{P}(\Omega_1))$ est ce qu'on peut faire de plus précis.

Vérifiez qu'il s'agit bien d'espaces probabilisables !

Axiomes de КОЛМОГОРОВО

Soit (Ω, \mathcal{T}) un espace probabilisable. On appelle **mesure de probabilité** (ou probabilité tout court) une application \mathbb{P} définie sur \mathcal{T} vérifiant :

1. $\forall A \in \mathcal{T}, \mathbb{P}(A) \in [0, 1]$;
2. $\mathbb{P}(\Omega) = 1$;
3. $\forall (A, B) \in \mathcal{T} \times \mathcal{T}, A \cap B = \emptyset \implies \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

On dit alors que $(\Omega, \mathcal{T}, \mathbb{P})$ est un **espace probabilisé**.

Définition 3 - 3



Андрей Николаевич
КОЛМОГОРОВ
(1903 - 1987)

Vous pouvez vérifier que l'espace probabilisable $(\Omega, \mathcal{P}(\Omega))$ avec $\Omega = \{1, 2, 3, 4, 5, 6\}$ muni de l'application :

$$\mathbb{P}: \begin{array}{ll} \mathcal{P}(\Omega) & \rightarrow [0, 1] \\ A & \mapsto \frac{|A|}{6} \end{array}$$

définit un espace probabilisé.

Pour l'axiome 3, on se souviendra de la formule du crible (cf théorème 2 - 6 page 17).

Comme Ω est dénombrable, il existe une certaine famille I d'entiers telle que $\Omega = \bigcup_{i \in I} \{\omega_i\}$ et pour tout événement A de \mathcal{T} , il existe une famille $J \subseteq I$ telle que $A = \bigcup_{j \in J} \{\omega_j\}$.

Alors, par application itérée du troisième axiome :

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{j \in J} \{\omega_j\}\right) = \sum_{j \in J} \mathbb{P}(\{\omega_j\})$$

Ainsi la donnée des valeurs prises par \mathbb{P} sur les événements élémentaires suffit à caractériser \mathbb{P} sur \mathcal{T} .

Dans le cas où Ω est fini et si tous les événements élémentaires ont la même probabilité (on dit qu'il y a **équiprobabilité**), alors :

$$\mathbb{P}(\omega_1) = \mathbb{P}(\omega_2) = \dots = \mathbb{P}(\omega_{|\Omega|}) = \frac{1}{|\Omega|}$$

Dans ce cas, \mathbb{P} est appelée **probabilité uniforme** sur Ω .



Attention ! Toute probabilité n'est pas uniforme.. Votre portable sonne : quelle est la probabilité que cela soit Monica B. qui vous fasse une déclaration d'amour ? La tribu contient deux événements en plus de l'événement certain et de l'événement impossible : « c'est Monica » et « ce n'est pas Monica ». Le rapport entre cas favorables et cas possibles est bien $\frac{1}{2}$ et pourtant...

Vous noterez enfin que lorsque l'univers est infini dénombrable, il ne peut pas y avoir équiprobabilité : Pourquoi?...

Principales propriétés

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé. La probabilité \mathbb{P} vérifie :

1. $\mathbb{P}(\complement_{\Omega} A) = 1 - \mathbb{P}(A)$;
2. $\mathbb{P}(\emptyset) = 0$;
3. $\forall (A, B) \in \mathcal{T} \times \mathcal{T}, A \subset B \implies \mathbb{P}(A) \leq \mathbb{P}(B)$;
4. $\forall (A, B) \in \mathcal{T} \times \mathcal{T}, \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ (cas particulier de la formule du crible de POINCARÉ).

Théorème 3 - 1

Dans la suite, quand il n'y aura pas ambiguïté, on notera $\complement_{\Omega} A$ par \bar{A} .

Démonstration?...

Système complet d'événements

Une famille A_1, A_2, \dots, A_n d'événements non vides de \mathcal{T} forme un système complet d'événements si, et seulement si :

- $\forall i \neq j, A_i \cap A_j = \emptyset$ (événements incompatibles deux à deux) ;
- $\bigcup_{i=1}^n A_i = \Omega$.

On dit aussi que la famille (A_i) forme une **partition** de Ω .

Définition 3 - 4

Théorème des probabilités totales (version 1)

Soit $(A_i)_{1 \leq i \leq n}$ un système complet d'événements, alors :

$$\forall E \in \mathcal{T}, \mathbb{P}(E) = \sum_{i=1}^n \mathbb{P}(E \cap A_i)$$

Théorème 3 - 2

Preuve ?

4 Probabilités conditionnelles

Dans toute cette section on travail dans l'espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$.

4 1 Un exemple pour comprendre

Considérons l'expérience simplissime consistant à lancer deux fois un dé à six faces. L'univers Ω est donc constitué de l'ensemble des couples (i, j) , avec i et j appartenant à l'ensemble $[[1, 6]]$: il y a donc 36 éléments dans Ω . Intéressons nous à la somme des deux chiffres et soit A l'événement « le total fait neuf ». On prendra $\mathcal{T} = \mathcal{P}(\Omega)$.

$$A = \{(3, 6), (4, 5), (5, 4), (6, 3)\}$$

donc

$$\mathbb{P}(A) = \frac{4}{36} = \frac{1}{9}$$

Soit B l'événement : « on obtient 3 au premier lancer », alors

$$B = \{(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6)\}$$

donc

$$\mathbb{P}(B) = \frac{6}{36} = \frac{1}{6}$$

Jusqu'ici, tout roule. Posons-nous maintenant le problème suivant : sachant que B est réalisé, i.e. que l'on obtient 3 au premier lancer, quelle est la probabilité que le total fasse 9 ? On ne considère plus tous les éléments de Ω . Il semble alors nécessaire de définir un nouvel univers (notre nouvel ensemble des possibles) :

$$\Omega' = \{(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6)\} = \Omega \cap B$$

Puisque l'univers change, la probabilité aussi. L'événement A' « le total fait neuf » dans ce nouveau modèle s'écrit

$$A' = \{(3, 6)\} = A \cap B$$

et donc

$$\mathbb{P}'(A') = \frac{1}{6} = \frac{\text{nombre d'éléments de } A \cap B}{\text{nombre d'éléments de } B} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

4 2 Définition

Cela nous incite à poser la définition suivante :

Probabilité conditionnelle

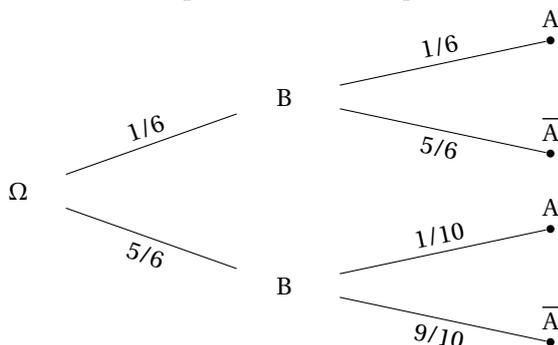
Soient A et B deux événements d'un espace probabilisé muni d'une probabilité \mathbb{P} , avec $\mathbb{P}(B) \neq 0$. La **probabilité de A sachant B** est définie par

$$\mathbb{P}(A | B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

On parle de probabilité conditionnelle. Mais est-ce bien une probabilité ? C'est-à-dire est-ce que $(\Omega, \mathcal{T}, \mathbb{P}(\cdot | B))$ est un espace probabilisé ? D'après la définition 3 - 3 page 30, il y a trois conditions à vérifier : faites-le !

4 3 Arbre

Reprenons l'exemple de la section 3.4.1 en répondant aux mêmes questions à l'aide d'un arbre *pondéré*, c'est à dire un arbre dont chaque branche est marquée de la probabilité (du *poids*) correspondant. Alors la somme des probabilités de chaque « ramification » est égale à 1.



À l'aide de la formule des probabilités totales, il est aisé d'obtenir les résultats suivant :

Définition 3 - 5

1. Calcul de $\mathbb{P}(A)$

B et \bar{B} forment une partition. D'après la formule des probabilités totales,

$$\mathbb{P}(A) = \mathbb{P}(B \cap A) + \mathbb{P}(\bar{B} \cap A) = \mathbb{P}(B) \times \mathbb{P}(A | B) + \mathbb{P}(\bar{B}) \times \mathbb{P}(A | \bar{B})$$

$$\mathbb{P}(A) = \frac{1}{6} \times \frac{1}{6} + \frac{5}{6} \times \frac{1}{10} = \frac{1}{9}$$

2. Calcul de $\mathbb{P}(\bar{A})$

B et \bar{B} forment une partition. D'après la formule des probabilités totales,

$$\mathbb{P}(\bar{A}) = \mathbb{P}(B \cap \bar{A}) + \mathbb{P}(\bar{B} \cap \bar{A}) = \mathbb{P}(B) \times \mathbb{P}(\bar{A} | B) + \mathbb{P}(\bar{B}) \times \mathbb{P}(\bar{A} | \bar{B})$$

$$\mathbb{P}(\bar{A}) = \frac{1}{6} \times \frac{5}{6} + \frac{5}{6} \times \frac{9}{10} = \frac{8}{9}$$

3. Probabilités conditionnelles

$$\mathbb{P}(B | A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} = \frac{\frac{1}{36}}{\frac{1}{9}} = \frac{1}{4}$$

$$\mathbb{P}(\bar{B} | A) = \frac{\mathbb{P}(\bar{B} \cap A)}{\mathbb{P}(A)} = \frac{\frac{1}{12}}{\frac{1}{9}} = \frac{3}{4}$$

$$\mathbb{P}(B | \bar{A}) = \frac{\mathbb{P}(B \cap \bar{A})}{\mathbb{P}(\bar{A})} = \frac{\frac{5}{36}}{\frac{8}{9}} = \frac{5}{32}$$

$$\mathbb{P}(\bar{B} | \bar{A}) = \frac{\mathbb{P}(\bar{B} \cap \bar{A})}{\mathbb{P}(\bar{A})} = \frac{\frac{3}{4}}{\frac{8}{9}} = \frac{27}{32}$$

4 4 Formule de BAYES

Cette formule correspond au « retournement » d'un arbre.

Deuxième formule de BAYES (probabilité des causes)

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé et $(A_i)_{i \in I \subseteq \mathbb{N}}$ un système complet d'événements.

$$\forall B \in \mathcal{T}, \quad \mathbb{P}(A_k | B) = \frac{\mathbb{P}(B | A_k) \times \mathbb{P}(A_k)}{\sum_{i \in I} \mathbb{P}(B | A_i) \times \mathbb{P}(A_i)}$$

Théorème 3 - 3

Preuve ?

On utilisera souvent ce théorème dans le cas particulier de deux événements :

Première formule de BAYES

$$\mathbb{P}(A | B) = \frac{\mathbb{P}(B | A) \times \mathbb{P}(A)}{\mathbb{P}(B)}$$

Théorème 3 - 4

On aurait pu écrire :

$$\mathbb{P}(\text{cause} | \text{effet}) = \frac{\mathbb{P}(\text{effet} | \text{cause}) \times \mathbb{P}(\text{cause})}{\mathbb{P}(\text{effet})}$$

Statistiquement, on a accès aux effets connaissant les causes (c'est le déterminisme...) mais dans la réalité, on subit les effets et on voudrait en connaître la cause.

Le cas le plus célèbre d'utilisation de la *statistique bayésienne* est le test de grossesse. Les médecins « disposent » d'un échantillon de femmes enceintes ou non (et donc de $\mathbb{P}(E)$ et $\mathbb{P}(\bar{E})$) et d'un test de grossesse. Ils font subir le test à tout l'échantillon et disposent donc de statistiques sur $\mathbb{P}(T | E)$, $\mathbb{P}(\bar{T} | \bar{E})$, $\mathbb{P}(T | \bar{E})$ et $\mathbb{P}(\bar{T} | E)$ mais ce qui intéresse le médecin et surtout la femme qui utilise le test c'est $\mathbb{P}(E | T)$ et $\mathbb{P}(\bar{E} | \bar{T})$...

4 5 Indépendance

La notions d'événements indépendants est l'une des difficultés du calcul des probabilités. Comme souvent, cette notion purement mathématique renvoie, par son appellation, à une notion intuitive utilisée dans le langage courant. Il faut bien garder en mémoire que le mot du vocabulaire courant est souvent un « faux ami ». ^b Donnons tout d'abord sa définition.

Définition 3 - 6

Événements indépendants

Les événements A et B sont dits **indépendants** si $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$

Ainsi, dans notre exemple de dés, on a $\mathbb{P}(A) \times \mathbb{P}(B) = \frac{1}{6} \times \frac{6}{36} = \frac{1}{6} = \mathbb{P}(A \cap B)$, donc le lancer du 2^{ème} dé est indépendant du premier, ce qui est rassurant.

b. On aurait aussi bien parler de szjwrtrpklance, mais cela aurait été plus difficile à prononcer.

Cette notion est purement abstraite et ne renvoie qu'à des propriétés mathématiques dont la principale est :

Théorème 3 - 5

Indépendance et probabilités conditionnelles

Soient A et B deux évènements tels que $\mathbb{P}(B) \neq 0$. Les évènements A et B sont indépendants si et seulement si $\mathbb{P}(A | B) = \mathbb{P}(A)$

On peut faire dire ce que l'on veut à des probabilités selon le modèle choisi.

Supposons que sur un groupe de 100 personnes, 20 portent des sous-vêtements en polystyrène expansé (pour avoir chaud l'hiver), 50 se grattent la tête avec l'index gauche et 10 font les deux à la fois. On met ces 100 personnes dans une boîte et on en tire une au hasard. Vérifiez que les évènements « la personne tirée porte des sous-vêtements en polystyrène expansé » et « la personne tirée se gratte la tête avec l'index gauche » sont indépendants.

Étudiez le même problème en considérant cette fois-ci que 15 personnes se grattent la tête avec l'index gauche (pourquoi pas).^c

Veillez à ne pas confondre évènements *indépendants* et évènements *incompatibles*.

On peut montrer d'ailleurs que deux évènements incompatibles de probabilité non nulle ne sont jamais indépendants.

En effet, A et B sont incompatibles si, et seulement si, $A \cap B = \emptyset$, donc $\mathbb{P}(A \cap B) = \mathbb{P}(\emptyset) = 0$. Si $\mathbb{P}(A) \neq 0$ et $\mathbb{P}(B) \neq 0$, on a forcément $\mathbb{P}(A \cap B) \neq \mathbb{P}(A) \times \mathbb{P}(B)$.

La seule idée à retenir est que, si A et B sont indépendants, **avoir observé la réalisation de A ne modifie pas la probabilité d'une éventuelle réalisation de B.**

Ainsi, en supposant que la Française des Jeux n'utilise pas de boules truquées, on peut considérer que deux tirages successifs du loto sont indépendants.

Indépendance mutuelle

Soit $(A_i)_{i \in I \subseteq \mathbb{N}}$ une suite d'évènements de \mathcal{T} .

- 1. On dit que ces évènements sont **indépendants deux à deux** si, et seulement si :

$$\forall (i, j) \in I^2, i \neq j, \mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i) \times \mathbb{P}(A_j)$$

- 2. On dit que ces évènements sont **mutuellement indépendants** si, et seulement si, pour toute partie J de I on a :

$$\mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}(A_j)$$

Théorème 3 - 6

5 Variables aléatoires réelles finies

On travaillera *en général* dans l'espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ avec $\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$ un univers dénombrable, voire le plus souvent fini de cardinal n .

5 1 Définition

L'ensemble $\mathcal{P}(\mathbb{R})$ est bien une tribu mais il pose quelques problèmes lors d'une étude plus approfondie des probabilités. Dans \mathbb{R} , on utilisera donc une tribu particulière qui correspond en fait à la tribu à laquelle on pense « naturellement », celle créée à partir des intervalles :

Définition 3 - 7

Tribu borélienne

On note $\mathbf{B}(\mathbb{R})$ l'ensemble des parties *engendrées* (par intersections et réunions dénombrables) par les intervalles de \mathbb{R} .

Vous vérifierez qu'il s'agit d'une tribu.

Définition 3 - 8

Variable aléatoire réelle

On appelle variable aléatoire réelle finie toute application de Ω dans \mathbb{R} vérifiant :

$$\forall I \subseteq \mathbb{R}, X^{-1}(I) \in \mathcal{T}$$

Cette définition nous assure que $X^{-1}(I)$ est un évènement dont on peut calculer la probabilité.

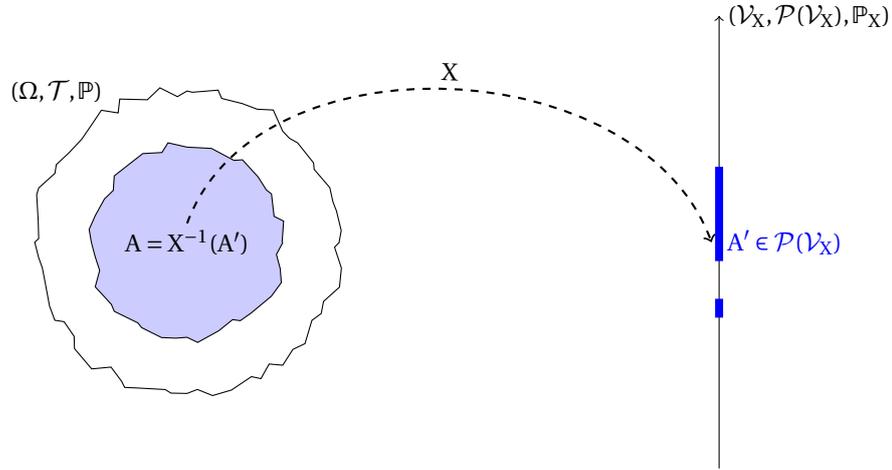
c. De façon générale, la définition probabiliste de l'indépendance est plus large que la notion intuitive.

Une v.a. est réelle car X est à valeurs dans \mathbb{R} et finie car l'univers est fini donc l'ensemble image de X aussi.

Danger

Vous noterez qu'une variable aléatoire réelle (v.a.r. pour les intimes) est en fait une fonction !...

Nous noterons $\mathcal{V}_X = X(\Omega)$ l'ensemble des valeurs prises par la v.a. X .



Soit x_1, \dots, x_k les différentes valeurs prises par la fonction X .

On note $\{X = x_i\}$ l'événement « la variable aléatoire prend la valeur x_i ». Il se note rigoureusement $\{\omega \in \Omega \mid X(\omega) = x_i\}$, ce qui se lit « l'ensemble des ω tels que $X(\omega) = x_i$ » ou encore $X^{-1}(\{x_i\})$.

La v.a.r. X induit une nouvelle probabilité (ou plutôt loi de probabilité comme nous allons le voir) \mathbb{P}_X sur $(\mathbb{R}, \mathbf{B}(\mathbb{R}))$. Par exemple, $\mathbb{P}_X(\{2\}) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) = 2\})$.

La définition précédente n'est pas très parlante, il faut essentiellement retenir que l'on va s'intéresser à des événements décrits par des nombres, les variables aléatoires nous permettent de décrire, et donc de noter, facilement des événements à l'aide de nombres.

Si \mathcal{V}_X est réduit à un singleton, $\mathcal{V}_X = \{\alpha\}$, on dit que X est une variable certaine, on note alors $X = \alpha$.

5 2 Loi de probabilité

Loi de probabilité d'une v.a.r.

Soit X une variable aléatoire réelle sur Ω . On appelle **loi de probabilité** de X la fonction \mathbb{P}_X de $\mathbf{B}(\mathbb{R})$ dans $[0, 1]$ définie par

Définition 3 - 9

$$\mathbb{P}_X: \begin{array}{ll} \mathbf{B}(\mathbb{R}) & \rightarrow [0, 1] \\ B & \mapsto \mathbb{P}(\{X^{-1}(B)\}) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in B\}) = \mathbb{P}(X \in B) = \mathbb{P}_X(B) \end{array}$$

Vérifiez que \mathbb{P}_X vérifie les axiomes de Kolmogorov donc que \mathbb{P}_X est bien une mesure de probabilité sur $(\mathbb{R}, \mathbf{B}(\mathbb{R}))$.

Dans la plupart des cas, on étudiera $\mathbb{P}(X = x)$, c'est-à-dire $\mathbb{P}(\{\omega \in \Omega \mid X(\omega) = x\})$.

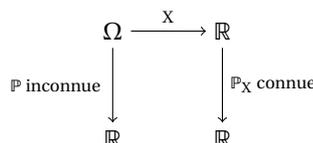
Si $x \notin X(\Omega)$, alors $\{X = x\} = \emptyset$ et donc $\mathbb{P}(\{X = x\}) = 0$.

Définir la *loi de probabilité d'une expérience aléatoire* reviendra donc à :

- déterminer toutes les valeurs possibles x_1, \dots, x_n prises par X ;
- calculer les probabilités p_1, \dots, p_n des événements correspondants ;
- regrouper les résultats dans un tableau du type :

Valeurs prises par X	x_1	x_2	\dots	x_n
Probabilité correspondante $\mathbb{P}(\{X = x_i\})$	p_1	p_2	\dots	p_n

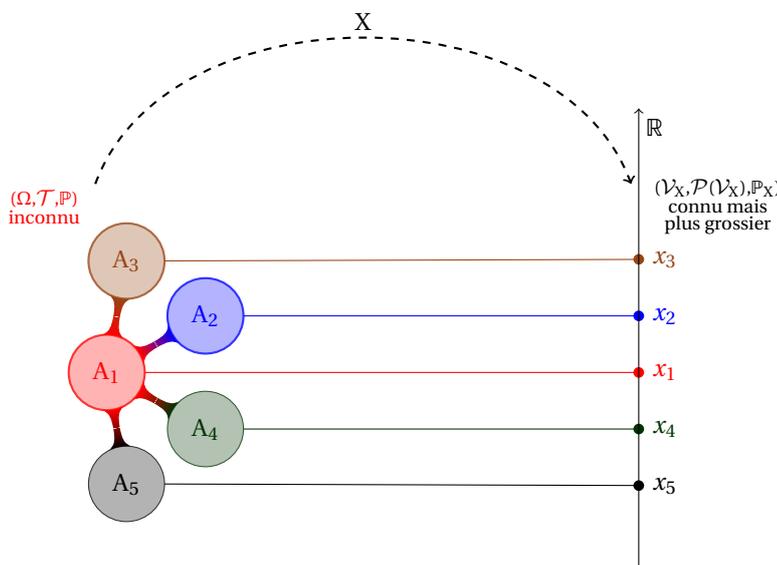
Vous n'oublierez pas de vérifier que $p_1 + \dots + p_n = 1$ d'après la loi des probabilités totales. En général, on se trouve dans la situation suivante :



À retenir

Dans la vraie vie, on connaît les valeurs prises par la v.a.r. mais on ne connaît pas bien Ω , voire pas du tout. On réfléchit donc sur les lois et non pas sur des ensembles comme au lycée.
 Par exemple, on peut vouloir réfléchir au nombre de plantages d'un réseau. On estime statistiquement les probabilités de tomber 0, 1, 2, etc. fois en panne dans une journée. On ne sait pas trop ce qu'est Ω mais on peut parler de la loi de X .
 En fait, une variable aléatoire « transporte » les calculs de probabilités d'un univers inconnu vers des valeurs réelles connues.
 La v.a.r. a pour effet de remplacer la tribu \mathcal{T} par une tribu $X^{-1}(\mathcal{P}(\mathcal{V}_X))$ en général plus petite : elle condense l'information sur l'univers en une information plus grossière (numérique) mais c'est déjà bien car en général on ne connaît de toute manière pas ou peu l'univers de départ!...
On retiendra donc que le « praticien » travaille la plupart du temps avec des lois de probabilités et non pas des ensembles : cela simplifie son travail car il peut travailler sur des espaces complexes (l'univers de départ est souvent trop complexe voire inconnu) sans les connaître.

L'information donnée par la loi est souvent plus grossière : par exemple, on lance deux dés et on regarde la somme des deux faces. La tribu de départ contient 2^{36} évènements alors que la tribu de la loi n'en contient que 2^{11} (il y a 11 sommes possibles).



À partir de la loi, on induit un système complet d'évènements sur Ω : la connaissance de la loi nous donne des renseignements sur l'univers mais souvent de manière grossière (par « paquets »).

Danger

Deux v.a.r. peuvent avoir la même loi sans être égales! On lance n fois une pièce non truquée. Considérez X le nombre de pile tombés et $n - X$ le nombre de face. Ces deux v.a.r. suivent la même loi binomiale de mêmes paramètres n et $1/2$ et pourtant elles ne sont pas égales (ni même égales presque partout ce qui signifie sauf sur un évènement de probabilité nulle mais ceci est une autre histoire).

5 3 Variables aléatoires indépendantes

La notation $\mathbb{P}(X = x_i, Y = y_j)$ signifie usuellement $\mathbb{P}(\{X = x_i\} \cap \{Y = y_j\}) = \mathbb{P}(X = x_i \text{ et } Y = y_j)$.
 On peut donc définir un **vecteur aléatoire** $Z = (X, Y)$ tel que

$$\mathbb{P}(Z = (x_i, y_j)) = \mathbb{P}(X = x_i, Y = y_j)$$

On peut se représenter l'espace produit à l'aide d'une représentation cartésienne : on parle alors d'évènements « parallèles » à des « axes » (les Ω_i), ces « axes » étant « orthogonaux ». Un système complet d'évènements est alors une « base ». Tous ces ponts de vocabulaire reliant plusieurs théories est un plaisir évident pour le mathématicien qui pourra faciliter la vie de l'informaticien mais vous pouvez laisser cela de côté dans un premier temps...

Couple de variables indépendantes

Les v.a.r. X et Y sont dites indépendantes si, et seulement si, pour toute partie $(A, B) \in \mathcal{V}_X \times \mathcal{V}_Y$,

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A) \cdot \mathbb{P}(Y \in B)$$

Définition 3 - 10

Définition 3 - 11

Variables aléatoires mutuellement indépendantes

Si $X_1, X_2, X_3, \dots, X_k$ sont k v.a. définies sur Ω , on dit qu'elles sont mutuellement indépendantes si, et seulement si, quels que soit $(I_1, I_2, \dots, I_k) \in \mathcal{V}_{X_1} \times \mathcal{V}_{X_2} \times \dots \times \mathcal{V}_{X_k}$ on a :

$$\mathbb{P}(X_1 \in I_1, X_2 \in I_2, \dots, X_k \in I_k) = \mathbb{P}(X_1 \in I_1) \times \mathbb{P}(X_2 \in I_2) \times \dots \times \mathbb{P}(X_k \in I_k)$$

Attention! L'indépendance deux à deux de variables aléatoires n'entraîne pas leur indépendance mutuelle, comme nous le verrons en TD.

5 4 Fonctions de répartition**Fonction de répartition**

Soit X une v.a.r. On appelle fonction de répartition de X la fonction F de \mathbb{R} dans \mathbb{R} définie par :

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}_X(]-\infty, x])$$

On note souvent cette fonction F_X .

Voici quelques propriétés que vous démontrerez :

Propriétés de F_X

1. $\forall x \in \mathbb{R}, F_X(x) \in [0, 1]$
2. F_X est croissante (sens large).
3. F_X est continue à droite en tout point.
4. $\lim_{x \rightarrow -\infty} F_X(x) = 0$
5. $\lim_{x \rightarrow +\infty} F_X(x) = 1$
6. $\mathbb{P}(a < X \leq b) = \mathbb{P}_X(]a, b])F_X(b) - F_X(a)$
7. $F_X(a) = \mathbb{P}(X \leq a) = \mathbb{P}_X(]-\infty, a]) = 1 - \mathbb{P}(X > a)$
8. $\mathbb{P}(X = a) = \lim_{x \rightarrow a^+} F_X(x) - \lim_{x \rightarrow a^-} F_X(x) = F_X(a) - \lim_{x \rightarrow a^-} F_X(x)$. Par conséquent en tout point a où F_X est continue on a $\mathbb{P}(X = a) = 0$.

Théorème 3 - 7

5 5 Espérance mathématique**Espérance mathématique**

On appelle espérance mathématique de la variable aléatoire discrète X le nombre noté $\mathbb{E}(X)$ défini par

$$\mathbb{E}(X) = x_1 \cdot \mathbb{P}_X(x_1) + x_2 \cdot \mathbb{P}_X(x_2) + \dots + x_n \cdot \mathbb{P}_X(x_n) = \sum_{i=1}^n x_i \cdot \mathbb{P}_X(x_i)$$

Exemple : on lance un dé honnête. On définit la variable aléatoire X qui prend la valeur 2 si le numéro du dé est pair et 1 sinon.

Notons ω_i l'événement « le numéro de la face est i », alors $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$ et

i	1	2	3	4	5	6
$\mathbb{P}(\omega_i)$	1/6	1/6	1/6	1/6	1/6	1/6
$X(\omega_i)$	1	2	1	2	1	2

D'après le théorème précédent, on a $\mathbb{E}(X) = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 1 \times \frac{1}{6} + 2 \times \frac{1}{6} = \frac{3}{2}$

Nous aurions pu procéder autrement pour définir la loi de probabilité :

Valeurs x_i prises par X	1	2
$\mathbb{P}_X(x_i)$	1/2	1/2

alors, d'après la définition, $\mathbb{E}(X) = 1 \times \frac{1}{2} + 2 \times \frac{1}{2} = \frac{3}{2}$

Espérance de $f \circ X$

Soit X une v.a.r. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ telle que $f \circ X$ soit une v.a.r. Alors

$$\mathbb{E}(f \circ X) = \sum_{n \in \mathbb{N}} f(x_n) \cdot \mathbb{P}(X = x_n) = \sum_{n \in \mathbb{N}} f(x_n) \cdot \mathbb{P}_X(x_n)$$

Théorème 3 - 8

Cette propriété sera admise dans un premier temps.

5 6 Variance

La variance mesure l'« écart » par rapport à l'espérance en faisant la somme des carrés des « distances » entre chaque valeur prise par la variable aléatoire et l'espérance pondérés par la probabilité correspondante, ce qui donne :

Définition 3 - 14

Variance

$$V(X) = \sum_{i=1}^n (x_i - \mathbb{E}(X))^2 \cdot \mathbb{P}(X = x_i) = \sum_{i=1}^n (x_i - \mathbb{E}(X))^2 \cdot \mathbb{P}_X(x_i)$$

On a choisi d'utiliser les carrés de manière arbitraire pour ne pas avoir de problèmes de signes ; on aurait pu choisir une autre méthode mais celle-ci a l'avantage de rappeler la distance euclidienne bien connue. La variance est en ce sens homogène au carré d'une distance. On va donc définir une distance proprement dite en en prenant la racine carrée : c'est ce qu'on appelle l'**écart-type**.

Définition 3 - 15

Écart-type

$$\sigma(X) = \sqrt{V(X)}$$

Vous pouvez obtenir espérance, variance et écart-type très simplement à l'aide des modules statistiques de vos calculatrices. Il suffit de rentrer les valeurs prises par la variable aléatoire en liste 1, les probabilités correspondantes en liste 2.

5 7 Linéarité de l'espérance

À partir de variables aléatoires existantes, on peut en créer de nouvelles. Avec des notations usuelles on obtient

- $aX + b : \omega_i \mapsto aX(\omega_i) + b$ avec a et b des réels.
- $X + Y : \omega_i \mapsto X(\omega_i) + Y(\omega_i)$

On peut alors démontrer les propriétés suivantes :

Théorème 3 - 9

Linéarité de l'espérance

- $\mathbb{E}(aX + b) = \sum_{i=1}^n (ax_i + b) \cdot \mathbb{P}_X(x_i) = a\mathbb{E}(X) + b$
- $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$

Les démonstrations sont similaires. Notons $Y = aX + b$.

Pour tout $\omega_i \in \Omega$ on a $Y(\omega_i) = aX(\omega_i) + b$ et donc $y_i = ax_i + b$ avec les notations usuelles donc $\{\omega \in \Omega | Y(\omega) = y_i\} = \{\omega \in \Omega | aX(\omega) + b = ax_i + b\} = \{\omega \in \Omega | X(\omega) = x_i\}$

Alors $\mathbb{E}(Y) = \sum_{i=1}^n y_i \mathbb{P}(Y = y_i) = \sum_{i=1}^n (ax_i + b) \mathbb{P}(X = x_i) = a \sum_{i=1}^n x_i \mathbb{P}(X = x_i) + b \sum_{i=1}^n \mathbb{P}(X = x_i)$

On en déduit que $\mathbb{E}(Y) = a\mathbb{E}(X) + b$.

5 8 Théorème de König-Huygens

Johann KÖNIG (1712-1757) fut un mathématicien allemand, élève de Jean BERNOULLI et Christian HUYGENS (1629-1695) était lui néerlandais.

Théorème 3 - 10

Théorème de König-Huygens

$$V(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

À démontrer en TD...

5 9 V.a.r. centrée réduite

Soit X une v.a.r. qui admet une variance non nulle. On appelle v.a.r. centrée réduite associée à X la v.a.r. :

Définition 3 - 16

$$X^* = \frac{X - \mathbb{E}(X)}{\sigma(X)}$$

Cette définition en est-elle une ? Que valent $\mathbb{E}(X^*)$ et $V(X^*)$?

5 10 Fonction indicatrice d'un évènement

Soit A un évènement aléatoire appartenant à \mathcal{T}_Ω . Définissons la variable aléatoire X :

$$X(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases}$$

On note cette variable aléatoire $X = \mathbb{1}_A$ et on l'appelle **fonction indicatrice** de A .

Théorème 3 - 11

Interprétation d'une probabilité en terme d'espérance d'un v.a.

$$\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$$

Preuve ?

6 Quelques lois discrètes classiques**6 1 Loi uniforme (discrète)**

C'est la loi qui correspond au tirage équiprobable d'un jeton blanc dont on note le numéro dans une urne qui contient n jetons blancs numérotés.

Définition 3 - 17

Loi uniforme

On dit que $X : \Omega \rightarrow [1, n]$ suit une **loi uniforme** sur $[1, n]$ et on note $X \rightsquigarrow \mathcal{U}([1, n])$ si :

$$X(\Omega) = [1, n] \quad \text{et} \quad \forall k \in [1, n], \mathbb{P}(X = k) = \frac{1}{n}$$

Calculons l'espérance et la variance de $X \rightsquigarrow \mathcal{U}([1, n])$.

$$\mathbb{E}(X) = \sum_{k=1}^n k \mathbb{P}(X = k) = \sum_{k=1}^n k \frac{1}{n} = \frac{1}{n} \sum_{k=1}^n k = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}.$$

De manière analogue, on obtient que $\mathbb{V}(X) = \frac{n^2-1}{12}$ sachant que $\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$.

6 2 Loi de Bernoulli : to be or not to be

On place des jetons blancs et des noirs dans une urne. On suppose que toutes les conditions sont réunies pour effectuer le tirage d'un jeton dans le cadre d'équiprobabilité et que la proportion de jetons blancs est p . Que dire de la variable aléatoire qui vaut 1 si le jeton tiré est blanc et 0 s'il est noir ?

Loi de Bernoulli

On dit que la v.a.r. finie X suit la **loi de Bernoulli de paramètre p** si :

- $p \in]0; 1[$, $X(\Omega) = \{0, 1\}$;
- $\mathbb{P}(X = 1) = p$.

On note alors $X \rightsquigarrow \mathcal{B}(1, p)$

Définition 3 - 18

On dit aussi que X est une **variable aléatoire de Bernoulli**. Sa loi de probabilité est simple à décrire :

x_i	0	1
$\mathbb{P}(X = x_i)$	$1-p$	p

Il est simple de déterminer son espérance et sa variance :

- $\mathbb{E}(X) = 0 \times \mathbb{P}(X = 0) + 1 \times \mathbb{P}(X = 1) = p$;
- $\mathbb{V}(X) = (0 - p)^2 \times \mathbb{P}(X = 0) + (1 - p)^2 \times \mathbb{P}(X = 1) = p^2(1 - p) + (1 - p)^2 p = p(1 - p)$.

6 3 Loi binomiale

Dans les mêmes conditions que précédemment, on effectue n tirages d'un jeton avec remise. À chaque tirage on associe une v.a.r. indépendante des autres qui suit la loi de Bernoulli de paramètre p . Les issues élémentaires ω de ces n tirages sont des « mots » de n lettres, chaque lettre étant un B (pour blanc) ou un \bar{B} . On définit alors la variable aléatoire X à valeurs dans $\{0, 1, 2, \dots, n\}$, donnant le nombre de jetons blancs de ces issues élémentaires, c'est à dire que X est la somme des v.a.r. finies élémentaires suivant la loi de Bernoulli de manière indépendante.

Loi binomiale

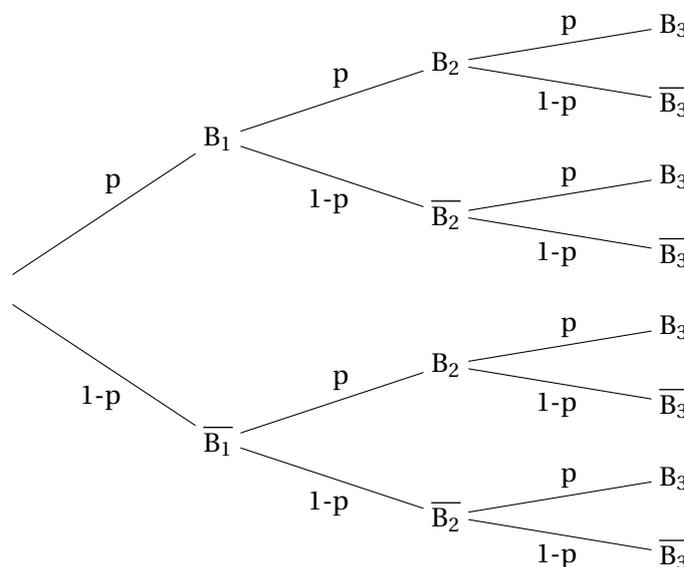
On dit que la v.a.r. finie X suit la **loi binomiale de paramètres n et p** si $n \in \mathbb{N} \setminus \{0\}$, $p \in]0; 1[$, $X(\Omega) = \llbracket 1, n \rrbracket$ et si l'expérience est une suite de n expériences de Bernoulli de paramètre p menées indépendamment. On note alors $X \rightsquigarrow \mathcal{B}(n, p)$

Définition 3 - 19

On s'intéresse maintenant à l'évènement $(X = k)$, c'est à dire l'ensemble des mots de n lettres écrits avec k B et $n - k$ \bar{B} . Notons B_j l'évènement « obtenir un jeton blanc au $j^{\text{ème}}$ tirage ». Chaque tirage d'une boule étant indépendant des autres, on a

$$\mathbb{P}(\omega) = \mathbb{P}(B_1) \times \dots \times \mathbb{P}(B_k) \times \mathbb{P}(\overline{B_{k+1}}) \times \dots \times \mathbb{P}(\overline{B_n}) = p^k (1 - p)^{n-k}$$

comme l'illustre cet arbre :



Il reste à déterminer combien il y a de tels mots. Ce sont les anagrammes (permutations distinctes) du mot de n lettres $BB\dots B\bar{B}\bar{B}\dots\bar{B}$. Il y en a $n!$ qu'il faut diviser par les $k!$ permutations des B et les $(n - k)!$ permutations des \bar{B} , c'est-à-dire $\binom{n}{k}$. Finalement, on obtient :

Pour toute v.a.r. finie $X \rightsquigarrow \mathcal{B}(n, p)$, on a, pour tout $k \in X(\Omega)$,

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Théorème 3 - 12

Vous démontrerez que $\mathbb{E}(X) = np$ et $\mathbb{V}(X) = np(1 - p)$.

6 4 Loi hypergéométrique

Modèle : on considère une urne contenant N boules ($N \in \mathbb{N}^*$), la proportion de boules blanches dans l'urne étant p . On tire simultanément n boules (les boules sont choisies au hasard) et X est la v.a. qui indique le nombre de boules blanches tirées. On dit que X suit la loi hypergéométrique de paramètres N , n et p notée $\mathcal{H}(N, n, p)$.

On retiendra :

- X peut prendre toute valeur entière de l'intervalle $\{\text{Max}\{0, n - N(1 - p)\}, \text{Min}\{n, Np\}\}$
- $k \in \mathcal{V}_X, \mathbb{P}(\{X = k\}) = \frac{\binom{n}{k} \binom{N-n}{n-k}}{\binom{N}{n}}$
- $\mathbb{E}(X) = np, \mathbb{V}(X) = np(1 - p) \frac{N - n}{N - 1}$

6 5 Loi géométrique

Modèle : on répète indéfiniment une épreuve de Bernoulli dont la probabilité de succès est p .

Les répétitions étant mutuellement indépendantes, X désigne le rang de l'essai donnant pour la première fois un succès. On dit que X suit la loi géométrique de paramètre p et on écrit $X \rightsquigarrow \mathcal{G}(p)$. On retiendra :

$$\begin{aligned}\mathcal{V}_X &= \mathbb{N}^* \\ k \in \mathbb{N}^*, \mathbb{P}(\{X = k\}) &= p(1-p)^{k-1} \\ \mathbb{E}(X) &= \frac{1}{p}, \mathbb{V}(X) = \frac{1-p}{p^2}\end{aligned}$$

La loi géométrique modélise donc le **temps d'attente jusqu'au premier succès**.

Problème du collectionneur



Voici une application classique de la loi géométrique : en Syldavie, le petit Ivan collectionne les images du chocolat Wonka. Il y a n images différentes. Lorsqu'il aura obtenu les n images, il aura droit de visiter les usines Wonka. À chaque fois qu'il achète une tablette, la probabilité d'obtenir une image quelconque est $1/n$. Soit T_n le temps d'attente (en nombre de tablettes achetées) pour obtenir les n images. Calculez $\mathbb{E}(T_n)$.

Soit X_i la variable aléatoire correspondant au temps d'attente pour passer de $i-1$ images différentes à i images différentes. Quand on est dans l'état $i-1$, on dispose de $i-1$ images différentes. Si on tombe sur l'une des $i-1$ images dont on dispose (avec la probabilité $\frac{i-1}{n}$), on reste dans l'état $i-1$, sinon, on passe à l'état i (avec la probabilité $\frac{n-(i-1)}{n}$).

Ainsi, $X_i \rightsquigarrow \mathcal{G}(p_i)$ avec $p_i = \frac{n-(i-1)}{n}$. On en déduit que $\mathbb{E}(X_i) = \frac{1}{p_i} = \frac{n}{n+1-i}$.

Or, $T_n = X_1 + X_2 + \dots + X_n$, donc $\mathbb{E}(T_n) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \dots + \mathbb{E}(X_n)$:

$$\mathbb{E}(X_n) = \sum_{i=1}^n \frac{1}{p_i} = n \sum_{i=1}^n \frac{1}{n+1-i} = n \left(\frac{1}{n} + \frac{1}{n-1} + \dots + \frac{1}{2} + 1 \right)$$

Par exemple, avec $n = 10$:

```
-->k=[1:10];
-->10*sum(1 ./k)
ans =
    29.289683
```

Il faut en moyenne acheter 30 tablettes pour obtenir les dix images.

6 6 Loi de Pascal

Modèle : on répète indéfiniment une épreuve \mathcal{E} de Bernoulli de probabilité de succès p .

Les répétitions étant mutuellement indépendantes, X désigne le rang de l'essai amenant le $r^{\text{ème}}$ succès. On dit que X suit la loi de Pascal de paramètres r et p et on écrit $X \rightsquigarrow \mathcal{Pa}(r, p)$. On retiendra (démonstration en TD) :

$$\begin{aligned}\mathcal{V}_X &= \{r, r+1, r+2, r+3, \dots\} \\ k \in \mathcal{V}_X, \mathbb{P}(\{X = k\}) &= \binom{k-r}{k-1} p^r (1-p)^{k-r} \\ \mathbb{E}(x) &= \frac{r}{p}, \mathbb{V}(X) = \frac{r(1-p)}{p^2}\end{aligned}$$

Si $r = 1$ on retrouve la loi géométrique.

6 7 Loi binomiale négative

Modèle : on répète indéfiniment une épreuve \mathcal{E} de Bernoulli de probabilité de succès p .

Les répétitions étant mutuellement indépendantes, X désigne le nombre d'échecs précédant le $r^{\text{ème}}$ succès. On dit que X suit la loi binomiale négative de paramètres r et p et on écrit $X \rightsquigarrow \mathcal{J}(r, p)$. On retiendra (démonstration en TD) :

$$\begin{aligned}\mathcal{V}_X &= \mathbb{N} \\ k \in \mathbb{N}, \mathbb{P}(\{X = k\}) &= \binom{k}{k+r-1} p^r (1-p)^k \\ \mathbb{E}(x) &= \frac{r(1-p)}{p}, \mathbb{V}(X) = \frac{r(1-p)}{p^2}\end{aligned}$$

6 8 Loi de Poisson

La variable aléatoire X suit la loi de Poisson de paramètre λ si X peut prendre toute valeur entière k avec la probabilité $\mathbb{P}(\{X = k\}) = \frac{\lambda^k e^{-\lambda}}{k!}$. On retiendra aussi : $\mathbb{E}(X) = \lambda$ et $\mathbb{V}(X) = \lambda$.



Siméon-Denis POISSON (1781 - 1840)

6 8 1 Que modélise la loi de Poisson ?

Considérons une probabilité définie par :

$$p_j(\pi_n, n) = \begin{cases} \binom{n}{j} \pi_n^j (1 - \pi_n)^{n-j} & \text{si } j \leq n \\ 0 & \text{si } j > n \end{cases}$$

Ainsi, cette suite de probabilités est une extension sur \mathbb{N} tout entier de la loi binomiale $\mathcal{B}(n, \pi_n)$. Supposons que la suite (π_n) tende vers 0 quand n tend vers l'infini de telle sorte que

$$\lim_{n \rightarrow +\infty} n\pi_n = \lambda \in \mathbb{R}_+^*$$

Il faut également se souvenir que :

$$\forall t \in \mathbb{R}, \quad \lim_{n \rightarrow +\infty} \left(1 + \frac{t}{n}\right)^n = e^t$$

Alors on montre (au tableau pendant l'amphi...) que :

$$\lim_{n \rightarrow +\infty} p_j(n, \pi_n) = \frac{\lambda^j e^{-\lambda}}{j!}$$

Ainsi, la loi de Poisson modélise bien le nombre d'apparitions de phénomènes rares (le nombre de pannes sur un systèmes GNU, le nombre de sessions sans panne d'un système Windows, le nombre d'étudiants d'info travaillant leur cours, etc.) dans une suite infinie d'évènements ($\pi_n \sim \frac{\lambda}{n}$ est « petit » et n est « grand »).

7 Somme de v.a.r. indépendantes - Stabilité

7 1 Convolution

Le produit de convolution intervient dans de nombreux domaines, particulièrement en probabilités et en théorie du signal. On l'utilise sous sa forme continue (à l'aide d'une intégrale qui est une somme...continue) ou discrète (à l'aide d'une somme...discrète).

On peut l'exprimer de manière universelle avec un symbole « \int » bien pratique car souvent les propriétés des sommes discrètes et continues sont les mêmes (par exemple, l'espérance d'une loi continue ou discrète peut s'exprimer de la même manière).

$$f(x) \otimes g(x) = \int f(t) \cdot g(x-t) = \int f(x-t) \cdot g(t)$$

signifie que cet opérateur \otimes signifie à la fois :

$$f(x) \otimes g(x) = \int_{\mathbb{R}} f(t) \cdot g(x-t) dt = \int_{\mathbb{R}} f(x-t) \cdot g(t) dt$$

et

$$f(n) \otimes g(n) = \sum_{k \in \mathbb{Z}} f(k) \cdot g(n-k) = \sum_{k \in \mathbb{Z}} f(n-k) \cdot g(k)$$

En particulier, avec l'écriture usuelle des suites :

$$u_n \otimes v_n = \sum_{k \in \mathbb{Z}} u_k \cdot v_{n-k} = \sum_{k \in \mathbb{Z}} u_{n-k} \cdot v_k$$

7 2 Somme de v.a.r. indépendantes

Somme de v.a.r. indépendantes

Si X et Y sont deux v.a.r. indépendantes

$$\mathbb{P}(X + Y = k) = \sum_{i \in \mathbb{Z}} \mathbb{P}(X = i) \cdot \mathbb{P}(Y = k - i) = \sum_{i \in \mathbb{Z}} \mathbb{P}(X = k - i) \cdot \mathbb{P}(Y = i) = \mathbb{P}(X = k) \otimes \mathbb{P}(Y = k)$$

Ceci reste à démontrer...en TD of course !

Théorème 3 - 13

7 3 Applications**7 3 1 Loi binomiale comme somme de lois de Bernoulli**

Commencez par considérer la somme de 1 variable aléatoire indépendante de Bernoulli de même paramètre p ... puis généralisez par récurrence au cas de n variables aléatoires de Bernoulli de même paramètre p . Cela nous donnera également un moyen d'obtenir l'espérance de la loi binomiale $\mathcal{B}(n, p)$.

Remarque

Une loi de Bernoulli de paramètre p est une loi binomiale de paramètre $(1, p)$...

7 3 2 Stabilité de la loi binomiale

Montrez que si $X \rightsquigarrow \mathcal{B}(n, p)$ et $Y \rightsquigarrow \mathcal{B}(m, p)$ sont deux v.a.r. indépendantes alors $X + Y \rightsquigarrow \mathcal{B}(n + m, p)$.

7 3 3 Stabilité de la loi de Poisson

$$\left. \begin{array}{l} X \rightsquigarrow \mathcal{P}(\lambda) \\ Y \rightsquigarrow \mathcal{P}(\mu) \\ X \text{ et } Y \text{ indépendantes} \end{array} \right\} \Rightarrow Z = X + Y \rightsquigarrow \mathcal{P}(\lambda + \mu)$$

Démonstration : cherchons la loi de probabilité de Z . Comme les v.a. sont indépendantes

$$\begin{aligned} \mathbb{P}(Z = k) &= \sum_{i=1}^k \mathbb{P}(X = i) \cdot \mathbb{P}(Y = k - i) \\ \mathbb{P}(Z = k) &= \sum_{i=1}^k \frac{\lambda^i e^{-\lambda}}{i!} \cdot \frac{\mu^{k-i} e^{-\mu}}{(k-i)!} = e^{-(\lambda+\mu)} \sum_{i=1}^k \frac{1}{k!} \frac{k!}{i!(k-i)!} \lambda^i \mu^{k-i} \\ \mathbb{P}(Z = k) &= \frac{e^{-(\lambda+\mu)}}{k!} \sum_{i=1}^k \binom{k}{i} \lambda^i \mu^{k-i} = \frac{(\lambda + \mu)^k}{k!} \cdot e^{-(\lambda+\mu)} \end{aligned}$$

Preuve de $\lim_{n \rightarrow +\infty} \left(1 + \frac{x}{n}\right)^n = e^x$

Montrez que $\lim_{u \rightarrow 0} \frac{\ln(1+u)}{u} = 1$ en reconnaissant un taux de variation. Alors

$$\begin{aligned} (1 + x/n)^n &= e^{n \ln(1+x/n)} \\ &= e^{n \frac{\ln(1+x/n)}{x/n} \cdot x/n} \\ &= e^{x \cdot \frac{\ln(1+x/n)}{x/n}} \end{aligned}$$

Conclusion ?

Aparté

8

EXERCICES

Exercice 3 - 1 Formule du crible de Poincaré : cas de trois évènements

Déterminez $\mathbb{P}(A \cup B \cup C)$ en fonction des probabilités de $A, B, C, A \cap B, A \cap C, B \cap C$ et $A \cap B \cap C$.

Exercice 3 - 2 Un peu de technique

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé. Déterminez n tel que, pour tous évènements A et B de \mathcal{T} , $\mathbb{P}(A \cup B) + \mathbb{P}(A \cup \bar{B}) + \mathbb{P}(\bar{A} \cup B) + \mathbb{P}(\bar{A} \cup \bar{B}) = n$

Déterminez également les signes de $\mathbb{P}(A \cap B) - \mathbb{P}(A) + \mathbb{P}(\bar{B})$ et de $\mathbb{P}(A \cap B \cap C) - 1 + \mathbb{P}(\bar{A}) + \mathbb{P}(\bar{B}) + \mathbb{P}(\bar{C})$.

Exercice 3 - 3 Définition d'une probabilité

Soit $\Omega = \{\omega_1, \omega_2, \dots\}$ un ensemble infini dénombrable. On suppose que, pour tout entier naturel non nul k , $\mathbb{P}(\{\omega_k\}) = \frac{\lambda}{37^k}$ avec λ un réel.

Existe-il une valeur de λ telle que $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ soit un espace probabilisé ?

Exercice 3 - 4 Paradoxe des anniversaires

Quelle est la probabilité qu'au moins deux personnes dans cette salle aient leur anniversaire le même jour (on négligera les années bissextiles...)?

Exercice 3 - 5 Problèmes historiques

Tout a commencé au XVII^e siècle, lorsque les Princes demandèrent aux mathématiciens de les aider à gagner au jeu...

1. Le Prince de Toscane demande à Galilée pourquoi, en lançant trois dés, on obtient plus souvent un total de 10 qu'un total de 9, alors qu'il y a dans les deux cas exactement 6 façons d'obtenir ces résultats.
2. Le Chevalier de Méré soutient à Pascal que les deux jeux suivants sont favorables au joueur : obtenir au moins un 6 en lançant 4 fois de suite un dé, et obtenir au moins un double 6 en lançant 24 fois de suite 2 dés.

Vous simulerez ces expériences à l'aide de Scilab et vous démontrerez les résultats observés.

Exercice 3 - 6 Simulations

1. On lance trois fois de suite une pièce de monnaie. On compte combien de fois on obtient pile : simuler à l'aide de Scilab et démontrez dans le cas général.
2. Faites de même avec l'expérience aléatoire suivante : on dispose de trois urnes, la première contenant 7 boules blanches et 4 noires, la seconde 5 blanches et 2 noires et la troisième 6 blanches et 3 noires. On tire une boule dans chaque urne et on note le nombre de boules blanches obtenues.
3. On tire successivement et avec remise quatre boules dans une urne qui contient 7 boules blanches et 3 roses. On compte le nombre de tirages contenant exactement deux boules blanches - au moins une boule blanche.
4. On lance dix fois une pièce de monnaie et on s'intéresse au nombre maximal de résultats consécutifs égaux.

Exercice 3 - 7 Indépendance mutuelle

On lance un dé deux fois. On considère les évènements :

- A : « Le premier lancer a donné un numéro pair » ;
- B : « Le second lancer a donné un numéro impair » ;
- C : « La somme des deux numéros est un chiffre pair ».

Les évènements sont-ils deux à deux indépendants ? Mutuellement indépendants ?

Exercice 3 - 8 Roulette

Le rouge est sorti dix fois de suite à la roulette : quelle est la probabilité qu'il sorte une onzième fois ? On se placera dans le cas très improbable que la roulette n'est pas truquée et que le casino n'est pas dirigé par un mafieux ou un ex-entraîneur d'une équipe de France et ex-ministre.

Exercice 3 - 9 Formule de B. et Monica B.

En Syldavie, 70% des habitants sont de sexe masculin et 60% des hommes et 25% des femmes ont une photo de Monica B. dans leur portefeuille. Un extra-terrestre enlève au hasard une personne syldave et découvre une photo de Monica B. dans son portefeuille : quelle est la probabilité que la personne enlevée soit une femme ?

Exercice 3 - 10 Marylin et les chèvres

Dans le numéro du 9 septembre 1990 de *Parade*, Marilyn VOS SAVANT connue pour avoir un QI de 186, soit l'un des plus élevés au monde, répondit au courrier suivant : *Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say number 1, and the host, who knows what's behind the doors, opens another door, say number 3, which has a goat. He says to you, "Do you want to pick door number 2?" Is it to your advantage to switch your choice of doors ?*

Craig. F. Whitaker Columbia, MD

Cette lettre décrit en fait le déroulement d'un jeu télévisé des années 1970, *Let's make a deal*, animé par Carol MERRIL et **Monty HALL** : c'est en référence à ce dernier qu'on donna son nom à ce problème célèbre (plus tard, un haut représentant de la culture française reprit ce jeu : ce fut Lagaf et son *Bigdil...*).

Marilyn répondit qu'il valait mieux pour le candidat changer de porte quand le présentateur l'y invitait. Qu'en pensez-vous ?

Vous pourrez également proposer une simulation avec Scilab. Vous aurez alors peut-être besoin de la fonction `modulo`.

Exercice 3 - 11 Duel

Alexandre et Évariste s'affrontent en duel au pistolet suivant les règles suivantes :

- Ils tirent chacun leur tour. Le premier qui atteint la cible a gagné...
- Lorsqu'il tire, Alexandre atteint la cible avec la probabilité a ($0 < a < 1$) et il la rate avec la probabilité $\bar{a} = 1 - a$.
- Lorsqu'il tire, Évariste atteint la cible avec la probabilité b ($0 < b < 1$) et il la rate avec la probabilité $\bar{b} = 1 - b$.
- Alexandre tire le premier. Ainsi, Alexandre (resp. Évariste) n'effectue que des tirs de rang impair (resp. pair).

On considère, pour tout entier $n \geq 1$, les événements A_{2n-1} : « Alexandre gagne à l'issue du tir numéro $2n-1$ », B_{2n} : « Évariste gagne à l'issue du tir numéro $2n$ ».

1. Calculez, en fonction de a et b , les probabilités des événements A_1 , B_2 et A_3 . Plus généralement, calculez $\mathbb{P}(A_{2n-1})$ et $\mathbb{P}(B_{2n})$.
2. Pour $n \geq 1$, on note C_n (resp. D_n) l'évènement : « Alexandre (resp. Évariste) gagne à un tir dont le numéro est entre 1 et $2n-1$ (resp. entre 2 et $2n$) ». Calculer $\mathbb{P}(C_n)$ et $\mathbb{P}(D_n)$.
3. Calculer $\alpha = \lim_{n \rightarrow +\infty} \mathbb{P}(C_n)$ et $\beta = \lim_{n \rightarrow +\infty} \mathbb{P}(D_n)$. Vérifier que $\alpha + \beta = 1$. Comment interpréter ce résultat ?

Exercice 3 - 12 QCM

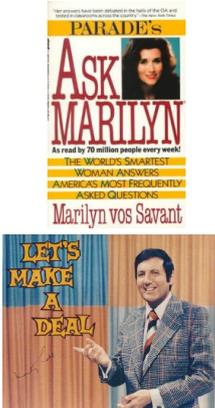
Dites si les variables aléatoires X définies ci-dessous suivent une loi binomiale ? Donnez quand c'est possible $X(\Omega)$, les paramètres de la loi ainsi que l'espérance.

1. Dans une classe on tire au sort et sans remise 5 élèves. X est le nombre d'élèves travaillant pour les services secrets bordures dans le lot tiré au sort.
2. Dans un sac de 20 billes contenant 7 noires et 13 blanches, on tire avec remise 3 d'entre elles, X étant le nombre de billes noires obtenues.
3. On lance 10 dés, X est le nombre de « 5 » obtenus.
4. Un circuit comprend 32 lampes en série. Pour chacune d'elle, la probabilité qu'elle fonctionne est de $3/100$. X est le nombre de lampes qui s'allument lorsqu'on appuie sur l'interrupteur.
5. Même question avec cette fois des lampes en parallèle.

Exercice 3 - 13 Lois usuelles

Dans chacune des expériences aléatoires ci-dessous, reconnaître une loi usuelle pour la variable aléatoire X . Donner alors son espérance.

1. On choisit une carte au hasard dans un jeu de 32 cartes numérotées. X = « hauteur de la carte choisie »



2. Un lion mange à chaque repas une gazelle (avec probabilité $2/3$) ou un zèbre (avec probabilité $1/3$). On suppose que la composition d'un repas est indépendante de celle des autres. X = « nombre de zèbres mangés sur 10 repas consécutifs ».
3. Dans un enclos se trouvent 15 gazelles et 30 zèbres. Un lion mange à chaque repas une gazelle ou un zèbre, pris parmi les animaux de l'enclos. X = « nombre de zèbres mangés sur 10 repas consécutifs ».
4. On tire une à une les cartes d'un jeu de 32 cartes, jusqu'à obtenir le valet de pique. X = « nombre de tirage effectuées ».
5. On lance un dé à 6 faces. X = « chiffre obtenu ».
6. Monica a dans sa poche 7 allumettes et 12 pièces de monnaie. Il met sa main dans sa poche et en sort 5 objets. X = « nombre de pièces obtenues ».
7. Monica fait passer une audition pour trouver des partenaires pour son prochain film. 3000 candidats passent l'audition. Chaque candidat réussit son entretien avec probabilité $2/3$. X = « nombre de candidats sélectionnés par Monica ».
8. James dort chez son amie Monica qui lui a prêté son appartement. Il dispose d'un trousseau de 7 clés, et ne sait pas laquelle ouvre la porte de l'appartement. Il les essaye une par une, en mettant de côté chaque clé essayée. X = « nombre d'essais nécessaires pour trouver la bonne clé ».

Exercice 3 - 14 Loi hypergéométrique

Justifiez la valeur donnée dans le cours de $\mathbb{P}(\{X = k\})$ lorsque $X \rightsquigarrow \mathcal{H}(N, n, p)$.

Exercice 3 - 15 Espérance géométrique

Nous avons prouvé au module précédent que :

$$\forall q \in \mathbb{R}, |q| < 1 \implies S = \sum_{n \geq 0} q^n = \frac{1}{1-q}$$

Pour $|q| < 1$, notons $T = \sum_{n \geq 1} nq^{n-1}$ en supposant que cette écriture a un sens.

Calculer $(1-q)T$ en fonction de S et déduisez-en T puis l'espérance d'une loi géométrique.

Exercice 3 - 16 Attentes

Combien faut-il en moyenne de lancers de dés pour obtenir les six faces ?

Combien faut-il en moyenne de tirages d'un entier choisi au hasard entre 0 et 9 pour les obtenir tous ?

Exercice 3 - 17 Roulette



Une roulette contient 36 cases numérotées de 1 à 36 dont 18 sont rouges et 18 sont noires, plus une case numérotée 0 de couleur verte.

Un joueur qui mise sur la couleur rouge ou noire gagne deux fois sa mise si la couleur choisie sort.

Un joueur qui mise sur un numéro de 1 à 36 gagne 36 fois sa mise si le numéro sort.

Il est interdit de miser sur le zéro.

1. Un joueur mise a € sur une couleur. Soit C la variable aléatoire correspondant au gain associé. Trouvez la loi de C puis calculez $\mathbb{E}(C)$ et $\sigma(C)$.
2. Un joueur mise a € sur un numéro. Soit N la variable aléatoire correspondant au gain associé. Trouvez la loi de N puis calculez $\mathbb{E}(N)$ et $\sigma(N)$.
3. Vaut-il mieux miser sur une couleur ou un numéro ?

Exercice 3 - 18 Dé truqué

On considère 6 dés, cinq étant équilibrés. Le dernier est pipé de manière à ce que lorsque l'on lance ce dé, chacun des chiffres apparaît avec une probabilité proportionnelle à ce chiffre.

- a) Donner la loi, l'espérance et la variance de la variable aléatoire égale au chiffre donné par le dé truqué lorsqu'on le lance.

On effectue n tirages successifs et indépendants d'un dé parmi les six.

- b) Quelle est la loi suivie par la variable aléatoire égale au nombre de fois où on a tiré le dé truqué ? Combien de tirages doit-on effectuer pour que la probabilité d'avoir obtenu le dé truqué parmi ceux tirés soit supérieure ou égale à $1/2$?

On effectue n tirages ($n \leq 6$) successifs sans remise d'un dé parmi les six.

- c) Quelle est la loi suivie par la variable aléatoire égale au nombre de fois où est tiré le dé truqué ? Combien de tirages doit-on effectuer pour que la probabilité d'avoir obtenu le dé truqué parmi ceux tirés soit supérieure ou égale à $1/2$?

Exercice 3 - 19 « C'est sans danger... »



Afin d'effectuer des économies dans le budget de l'État, le premier ministre syldave a supprimé la faculté de médecine. Dans le but de promouvoir les liens sociaux, les dentistes sont maintenant choisis parmi les anciens boxeurs aveugles et parkinsoniens. Ils arrachent les dents de leurs patients au hasard.

Suite à quelques plaintes, une étude statistique a été menée dans le cabinet d'un de ces nouveaux dentistes. Les syldaves venant le consulter ont toujours une seule dent de malade parmi les trente-deux qu'ils possèdent encore avant l'intervention des tenailles ou des poings, c'est selon (une étude précédente a en effet montré que les syldaves ne retournent jamais chez leur dentiste après une première consultation).

L'étude a été effectuée sur n patients. On note X la variable aléatoire égale au nombre de dents malades extraites à bon escient.

- Déterminez la loi de probabilité de la variable aléatoire X . Calculez la probabilité pour qu'aucune dent malade n'ait été extraite.
- Combien de patients doit-il traiter pour extraire au moins une dent malade avec une probabilité supérieure à $0,9$?
- Le dernier patient est prêt à se sacrifier pour la gloire des statistiques syldaves et se laisse arracher les dents une à une tant que la dent malade n'a pas été extraite. On note Y le nombre de dents saines que ce vaillant patriote voit tomber de ses mâchoires endolories.

Calculez la probabilité pour qu'il reparte complètement édenté, puis $\mathbb{E}(Y)$ et $\sigma(Y)$.

Exercice 3 - 20 Indépendance de v.a.

L'expérience aléatoire est le lancer de deux dés parfaits. X_1 est la variable de (Bernoulli) qui prend la valeur 1 si le premier dé donne un résultat pair 0 sinon, X_2 est la variable (de Bernoulli) qui prend la valeur 1 si le deuxième dé donne un résultat pair et 0 sinon et X_3 est la variable (de Bernoulli) qui prend la valeur 1 si la somme des numéros tirés est paire et 0 sinon.

Les variables aléatoires sont-elles mutuellement indépendantes ? Deux à deux indépendantes ?

Exercice 3 - 21 Théorème de König-Huygens

Démontrez ce théorème...

Exercice 3 - 22 Tirages sans remise avec condition d'arrêt

Une urne contient n boules. On y effectue des tirages sans remise.

- Quelle est la probabilité d'obtenir p boules données en k tirages sans remise ?
- L'urne contient p boules rouges et $n - p$ boules blanches. On note T la variable aléatoire égale au nombre de tirages nécessaires à l'obtention des p boules rouges.
 - Déterminer $T(\Omega)$ puis $\mathbb{P}(T \leq k)$ pour $k \in T(\Omega)$. Vérifier que cette probabilité s'exprime en fonction de $\frac{\mathbb{C}_k^p}{\mathbb{C}_n^p}$.
 - Déterminer la loi de T .
 - On rappelle que, pour tout $r \in \mathbb{N}$, pour tout $q \in \{0, \dots, r\}$: $\sum_{k=q}^r \mathbb{C}_k^q = \mathbb{C}_{r+1}^{q+1}$. Calculer $\mathbb{E}(T)$.

Exercice 3 - 23 Ordinateur syldave

Piotr a récupéré un vieil ordinateur de bord muni d'un processeur 4 bits sur un avion bordure abattu par les forces du peuple syldave. Il crée un compilateur qui pour l'instant ne renvoie qu'un quartet (une chaîne de 4 bits) au hasard, chaque bit valant 0 avec une probabilité p . Soit X la variable aléatoire égale au quartet renvoyé exprimé en base 10.

- Établissez la loi de probabilité de X .
- Calculez son espérance et sa variance.
- Calculez la probabilité d'obtenir un nombre pair.
- Calculez la probabilité que le quartet soit un palindrome.

Exercice 3 - 24 **Produit de variables aléatoires indépendantes**

Soit X_1 et X_2 deux v.a.r. discrètes finies prenant leurs valeurs respectives dans I_1 et I_2 . On rappelle que X_1 et X_2 sont indépendantes si, et seulement si, pour tout couple (x_1, x_2) de $I_1 \times I_2$:

$$\mathbb{P}(\{X_1 = x_1\} \cap \{X_2 = x_2\}) = \mathbb{P}(X_1 = x_1) \times \mathbb{P}(X_2 = x_2)$$

Montrez alors le théorème suivant :

Théorème 3 - 14

Espérance d'un couple de v.a.r. indépendantes

Si X et Y sont indépendantes alors $\mathbb{E}(XY) = \mathbb{E}(X) \times \mathbb{E}(Y)$

Exercice 3 - 25 **Opérations sur la variance**

Comment écrire $\mathbb{V}(aX)$, $\mathbb{V}(X + b)$ et $\mathbb{V}(X + Y)$ en fonction de $\mathbb{V}(X)$ et $\mathbb{V}(Y)$?

Exercice 3 - 26 **Approximation de la loi binomiale par la loi de Poisson**

On considère une suite de v.a.r. X_n ayant une loi de probabilité binomiale de paramètres n et λ/n . Montrez que :

$$\lim_{n \rightarrow +\infty} \mathbb{P}_{X_n}(k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

On dit alors que la suite de variables aléatoires X_n converge en loi (c.a.d. $\lim_{n \rightarrow +\infty} \mathbb{P}_{X_n} = \mathbb{P}_X$: la loi de X_n converge vers la loi de X). On étudie en fait un cas particulier de convergence de *suite de fonctions...* vers une v.a.r. X de loi de Poisson de paramètre λ .

Déduisez-en une approximation de $\binom{n}{k} p^k (1-p)^{n-k}$ lorsque $n \gg 1$.

Application : la fusion amoindrit les risques



Deux sociétés assurent des biens très chers mais dont le risque d'accident est très faible : vaut-il mieux fusionner ou bien les risques s'additionnent-ils ?

Prenons le cas d'une société d'assurance sicilienne ayant pour client la société des croisières Costa.

1. Supposons que la société Costa possède 500 navires dont la valeur est de 5 millions de Schpżřřřř. L'assureur rembourse la perte totale d'un navire, évènement dont la probabilité est estimée à 0,001 pour une année. On suppose que les risques de pertes des navires sont indépendants. Donner la loi de X égale au nombre de navires perdus au cours de l'année.
2. L'assureur rembourse son client une fois par an. Combien doit-il avoir dans son coffre pour qu'il puisse rembourser son client avec une probabilité de 0,999. On utilisera l'approximation du préambule et on utilisera une table de la loi de Poisson.
3. Une autre compagnie assure les 500 navires de la White Star dans les mêmes conditions. Les deux compagnies décident de fusionner suite au suicide du président de la première compagnie dont on a retrouvé le corps percé de 52 coups de couteaux. La compagnie d'assurance prend donc en charge à présent 1000 navires. Refaire les calculs de la question précédente dans ces conditions. Commentaires ?

Exercice 3 - 27 **Opération sur une v.a.r.**

Soit X une v.a.r. telle que $X(\Omega) = \mathbb{N}^*$ et $\mathbb{P}(\{X = n\}) = \frac{\lambda}{n!}$.

1. Déterminez λ .
2. Calculer $\mathbb{E}(X)$.
3. Calculer $\mathbb{E}\left(\frac{1}{1+X}\right)$

Exercice 3 - 28 **Espérance et variance de la loi de Poisson**

Démontrer les formules donnant l'espérance et la variance de la loi de Poisson.

On se souviendra de la série exponentielle évoquée au module précédent :

$$\forall x \in \mathbb{R}, \quad \sum_{k \geq 0} \frac{x^k}{k!} = e^x$$



Exercice 3 - 29 Les étudiants et le poisson

Une étude sérieuse vient de prouver que le nombre d'étudiants arrivant en salle F22 entre 12h30 et 12h35 les jeudis suit une loi de Poisson de paramètre $\lambda = 3$.

1. Quel est, en moyenne, le nombre d'étudiants qui arrivent entre 12h30 et 12h35 ?
2. Calculer la probabilité, qu'entre 12h30 et 12h35 aucun étudiant ne se présente en F22.
3. Calculer la probabilité, qu'entre 12h30 et 12h35 trois étudiants se présentent en F22.
4. Calculer la probabilité, qu'entre 12h30 et 12h35 plus de 4 étudiants se présentent en F22.
5. Calculer la probabilité, qu'entre 12h30 et 12h35 au moins un étudiant se présente en F22.

Exercice 3 - 30 Détection d'erreur par répétition

On considère une source binaire d'information constituée d'une suite de 1 et de 0 indépendants et équiprobables. La source émet à destination d'un récepteur à travers un canal de transmission bruyant entraînant des erreurs éventuelles en réception.

Soit p la probabilité, commune, qu'un bit émis soit détecté comme son complémentaire. On suppose que les erreurs commises sur chaque bit sont indépendantes mutuellement.

Pour tenter de lutter contre ce bruit, on utilise le protocole simple suivant :

- à l'émission, on répète une fois chaque bit ;
- à la réception, la réception de (b, b) entraîne la détection du symbole b ; la réception de (b, \bar{b}) donne lieu à une demande de retransmission d'un nouveau couple de bits identiques. On continue tant qu'un couple (b, \bar{b}) est reçu.

1. En quoi consiste l'erreur résiduelle d'une telle stratégie ? Donnez des exemples de chaînes transmises aboutissant à une erreur sur le bit détecté.
2. Montrez que $\mathbb{P}(\text{erreur}) = \mathbb{P}(\text{erreur} \mid 0 \text{ émis}) = \frac{p^2}{1-2p+2p^2}$.
3. Soit N la v.a.r. donnant le nombre de symboles émis par l'émetteur pour un symbole utile émis. Étudiez la loi de N et donnez son espérance et sa variance. Vous pourrez montrer que la loi de $N/2$ est géométrique. Vous tracerez l'allure des courbes de la loi, de l'espérance et de la variance en fonction de p . Commentez. Étudiez le cas particulier où $p \ll 1$.

Exercice 3 - 31 Loi faible des grands nombres

Cet exercice est en fait du cours et représente même un résultat plus que primordial...

Souvenez-vous : il y a bien longtemps, vous étiez en classe de seconde et vous découvriez les probabilités. Vous lanciez un dé et notiez la fréquence d'obtention de tel ou tel résultat. Déjà fort perspicace, vous aviez remarqué que plus vous lanciez le dé, plus la fréquence d'obtention de 4, par exemple, tendait vers $1/6$ qui, vous l'avez appris depuis, est la probabilité d'obtenir 4 en lançant un dé équilibré.

Pour résumer, plus vous répétez une expérience aléatoires dans des conditions identiques et avec indépendance des résultats, plus la *fréquence observée* de succès tend vers la probabilité de succès : c'est ce qu'on appelle la loi faible des grands nombres et nous allons essayer de prouver ce résultat dans le cas particulier de la loi binomiale.

Inégalité de Markov

Soit X une variable aléatoire suivant une loi de probabilité \mathbb{P}_X et ne prenant que des valeurs positives (rangées comme d'habitude dans l'ordre croissant) x_1, x_2, \dots, x_n .

Soit a un nombre (qui sera notre seuil d'erreur fréquence/probabilité) strictement positif fixé.

Le monde se sépare en deux catégories : les x_i strictement inférieurs à a et ceux qui lui sont supérieurs.

Supposons par exemple que $x_1 \leq x_2 \leq \dots \leq x_{k-1} < a \leq x_k \leq \dots \leq x_n$

1. Rappelez la définition de l'espérance $\mathbb{E}(X)$.
2. Montrez que $\mathbb{E}(X) \geq a \sum_{i=k}^n \mathbb{P}_X(x_i)$.
3. Déduez-en l'inégalité suivante $\mathbb{P}_X([a, +\infty[) \leq \frac{\mathbb{E}(X)}{a}$.
4. Généralisez afin d'obtenir l'inégalité de MARKOV : $\mathbb{P}_X([a, +\infty[) \leq \frac{\mathbb{E}(X^p)}{a^p}$ pour tout entier naturel non nul p .
5. Déduez-en cette autre formulation : $\mathbb{P}_X([0, a]) \geq 1 - \frac{\mathbb{E}(X^p)}{a^p}$
6. **Cas particulier de la loi binomiale**

On suppose que X suit la loi binomiale $\mathcal{B}(n, p)$. Donnez une majoration de $\mathbb{P}\left(\frac{X}{n} \geq a\right)$ et une minoration de $\mathbb{P}\left(\frac{X}{n} < a\right)$ à l'aide de l'inégalité de MARKOV.

Inégalité de Bienaymé-Tchebychev

1. Variance

La variance d’une variable aléatoire est la quantité qui mesure la *dispersion* de X autour de sa moyenne, c’est à dire son espérance. Plus les valeurs sont dispersées, plus sa variance augmente. On attend donc que si X est une variable aléatoire constante, sa variance sera nulle.

On aura remarqué que la variance est « homogène » à X².

2. Appliquez l’inégalité de Markov astucieusement pour obtenir l’inégalité de BIENAYMÉ-TCHEBYCHEV

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{\mathbb{V}(X)}{a^2}$$

Majorez alors la probabilité qu’une v.a.r. s’éloigne de 10σ(X) de son espérance.

3. Une application de ces inégalités

Les inégalités de nos amis franco-russes permettent d’avoir une estimation de certaines probabilités sans qu’on connaisse la loi de probabilité. il faut toutefois être conscient que la probabilité exacte peut être assez éloignée de la borne proposée.

Par exemple, le nombre de caleçons molletonnés fabriqués dans une usine syldave en une semaine est une variable aléatoire d’espérance 50 et de variance 25.

Estimez, grâce à l’inégalité de Markov, la probabilité que la production de la semaine à venir dépasse 75 caleçons molletonnés.

Estimez, grâce à l’inégalité de Bienaymé-Tchebychev, la probabilité que la production de la semaine à venir soit strictement comprise entre 40 et 60 caleçons molletonnés.



Loi faible des grands nombres dans le cas de la loi binomiale

1. Soit λ un réel strictement positif et X une variable aléatoire quelconque. Montrez que $\mathbb{E}(\lambda X) = \lambda \mathbb{E}(X)$ puis que $\mathbb{V}(\lambda X) = \lambda^2 \mathbb{V}(X)$.

2. Considérons maintenant le cas où X obéit à la loi binomiale $\mathcal{B}(n, p)$.

Intéressons-nous à la probabilité d’obtenir k succès au cours des n répétitions de l’épreuve.

Le rapport k/n est alors le nombre relatif (la fréquence) de succès.

Nous allons essayer de confirmer notre étude expérimentale, à savoir que, plus n est grand, plus le rapport k/n est proche de la probabilité p de succès à chacune des épreuves.

On a démontré que la variance de X vérifie $\mathbb{V}(X) = np(1 - p)$

Montrez que

$$\mathbb{P}\left(\left|\frac{X}{n} - p\right| > a\right) \leq \frac{p(1-p)}{na^2} \quad \text{puis que} \quad \lim_{n \rightarrow +\infty} \mathbb{P}\left(\left|\frac{X}{n} - p\right| \leq a\right) = 1$$

Qu’en déduisez-vous ?

Exercice 3 - 32 Norme et Espace \mathcal{L}^1

Nous étudierons l’an prochain la notion d’*espace vectoriel*. Disons pour faire (très) court que c’est un espace qui est stable par combinaisons linéaires : si j’additionne deux vecteurs ou si je multiplie un vecteur par un scalaire, je reste dans l’espace et il existe un vecteur nul.

Une *norme* est une application définie sur un espace vectoriel \mathcal{E} , à valeurs dans \mathbb{R}_+ , notée le plus souvent $\|\cdot\|$ et vérifiant :

- 1.** $\forall x \in \mathcal{E}, \|x\| = 0 \Leftrightarrow x = 0$;
- 2.** $\forall x \in \mathcal{E}, \forall \lambda \in \mathbb{R}, \|\lambda \cdot x\| = |\lambda| \times \|x\|$;
- 3.** $\forall (x, y) \in \mathcal{E}^2, \|x + y\| \leq \|x\| + \|y\|$.

Vérifiez que l’application définie sur l’ensemble des v.a.r. possédant une espérance finie par $\|X\| = \mathbb{E}(|X|)$ est une norme. On note cet espace $\mathcal{L}^1(\Omega, \mathcal{T}_\Omega, \mathbb{P})$ ou plus simplement \mathcal{L}^1 .

Exercice 3 - 33 Produit scalaire et espace \mathcal{L}^2

Montrer que l’espace des v.a.r. possédant une variance finie est un espace vectoriel normé si on le muni de la norme $\|X\|_2 = \sqrt{\mathbb{E}(X^2)}$. Cet espace est noté $\mathcal{L}^2(\Omega, \mathcal{T}_\Omega, \mathbb{P})$ ou plus simplement \mathcal{L}^2 .

En fait, cette norme est déduite du *produit scalaire* défini par $\langle X, Y \rangle = \mathbb{E}(XY)$.

Un produit scalaire défini sur un espace vectoriel \mathcal{E} est une application à valeurs dans \mathbb{R} vérifiant les propriétés suivantes :

- 1.** $\forall (x, y) \in \mathcal{E}^2, \langle x, y \rangle = \langle y, x \rangle$;
- 2.** $\forall (x, y, z) \in \mathcal{E}^3, \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$;

$$3. \forall (\lambda, (x, y)) \in \mathbb{R} \times \mathcal{E}^2, \langle \lambda \cdot x, y \rangle = \lambda \times \langle x, y \rangle;$$

$$4. \forall x \in \mathcal{E}, \langle x, x \rangle \geq 0;$$

$$5. \forall x \in \mathcal{E}, \langle x, x \rangle = 0 \Leftrightarrow x = 0.$$

Vérifiez que $\langle X, Y \rangle = E(XY)$ définit un produit scalaire sur l'espace des v.a.r. possédant une variance finie.

Exercice 3 - 34 Covariance, corrélation et cosinus

Toujours plus fort dans l'abstraction des notions mathématiques : nous allons pouvoir parler d'angles de v.a.r., de v.a.r. orthogonales, de cosinus... Tout est géométrique et algébrique en ce bas monde ? Il est amusant de noter que si ces propriétés géométriques et algébriques semblent se retrouver dans de très nombreux domaines, elles ont également tendance à disparaître de l'enseignement secondaire, mais ceci est une autre histoire...

Aparté

Les statistiques et moi...

J'ai longtemps été hermétique à toutes ces formules quand elles m'avaient été présentées dans un cadre purement statistique car je ne voyais pas la motivation de créer de telles fonctions. Lorsque j'ai pu faire le lien avec l'intuition géométrique, tout s'est éclairé...

Souvenez-vous... En géométrie, vous aviez travaillé sur la produit scalaire, avec des définitions du type : $\langle \vec{u}, \vec{v} \rangle = \|\vec{u}\| \times \|\vec{v}\| \times \cos(\widehat{\vec{u}, \vec{v}})$.

On va retrouver ce genre de relation en probabilités.

Covariance

Soit X et Y deux v.a.r. \mathcal{L}^2 . On appelle *covariance* de X et Y et on note $\text{cov}(X, Y)$ le réel :

$$\text{cov}(X, Y) = E((X - E(X))(Y - E(Y)))$$

Vérifiez que :

Covariance et produit de v.a.r.

$$\text{cov}(X, Y) = E(XY) - E(X)E(Y)$$

Vérifiez que la covariance définit un produit scalaire et démontrez que :

Variance de la somme de deux v.a.r.

$$\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2 \text{cov}(X, Y)$$

Vérifiez également que :

Covariance et v.a.r. indépendantes

$$\text{cov}(X, Y) = 0 \iff X \perp\!\!\!\perp Y$$

la notation $X \perp\!\!\!\perp Y$ indiquant que X et Y sont indépendantes (il y a effectivement un lien entre indépendance et orthogonalité...).

On remarque toutefois que la covariance dépend de l'ordre de grandeur des v.a.r. : on va donc les « normer », c'est-à-dire « neutraliser » leur ordre de grandeur en divisant par leur norme 2 :

Coefficient de corrélation

Soit X et Y deux v.a.r. \mathcal{L}^2 . On appelle *coefficient de corrélation* de X et Y le réel

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma(X)\sigma(Y)} = \frac{\langle X, Y \rangle}{\|X\|_2 \|Y\|_2}$$

en prenant comme produit scalaire la covariance.

On retrouve alors la définition du cosinus de deux vecteurs du plan. Vérifier que ce nouveau « cosinus » est compris entre -1 et 1.

Comment interpréter cet « angle » entre deux v.a.r. ? En fait, en géométrie, plus le cosinus se rapproche de 1 en valeur absolue, plus les vecteurs sont « liés » linéairement (ils sont presque colinéaires quoi...).

Nous allons donc chercher l'interprétation du coefficient de corrélation en tant que « mesure » des « liens linéaires » entre deux variables.

Nous allons donc chercher la meilleure approximation linéaire de Y par une fonction affine de X de la forme $aX+b$ au sens habituel des **moindres carrés**, c'est-à-dire pour minimiser la quantité $\mathbb{E}((Y - (aX+b))^2)$. Notons $f(a, b)$ cette quantité. Un résultat classique de l'étude des fonctions de deux variables montre qu'il faut que la dérivée par rapport à a et celle par rapport à b de $f(a, b)$ doivent être nulles.

« Développez » $f(a, b)$, c'est-à-dire utilisez la linéarité de l'espérance pour « sortir » a et b des \mathbb{E} .

Dérivez alors $f(a, b)$ par rapport à a puis par rapport à b : ces deux quantités devant s'annuler, on obtient un système de deux équations linéaires d'inconnues a et b que vous résoudrez.

Vérifiez que :

$$a = \rho(X, Y) \frac{\sigma(Y)}{\sigma(X)}, \quad b = \mathbb{E}(Y) - a\mathbb{E}(X)$$

Montrez alors que le carré de l'erreur vaut $\sigma^2(Y)(1 - \rho^2(X, Y))$. On prouve ainsi que **plus $|\rho(X, Y)|$ est proche de 1, meilleure est l'approximation affine.**

Définition 3 - 22

Droite des moindres carrés

La droite d'équation $y = ax + b$ avec a et b les coefficients calculés précédemment est appelée droite des moindres carrés de Y sur X .

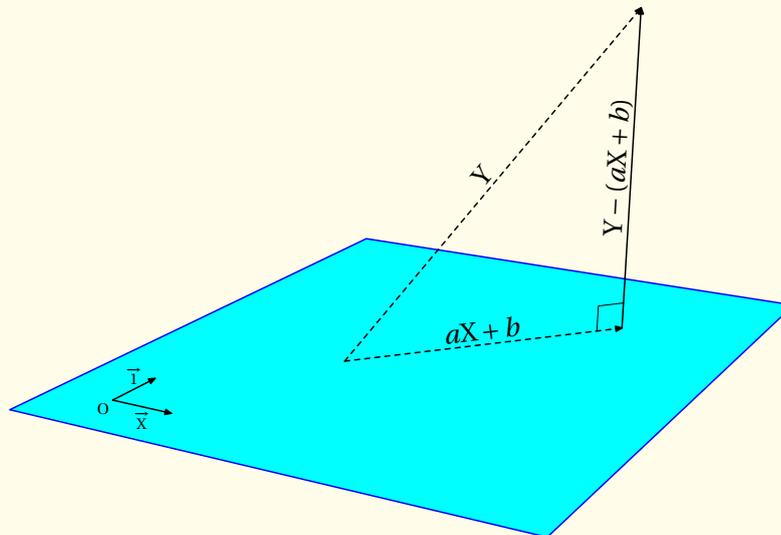
Vision géométrique

Un mathématicien a bien en tête que nous travaillons depuis un moment dans l'espace de Hilbert \mathcal{L}^2 : en gros, quelque chose qui ressemble à notre espace géométrique du lycée, vecteurs compris.

« L'espace des $aX + b$ » est en fait un plan de base les « vecteurs » X et $\mathbb{1}$.

Trouver le vecteur qui minimise $\mathbb{E}((Y - (aX + b))^2)$ c'est chercher la projection orthogonale de Y sur le « plan » engendré par X et $\mathbb{1}$.

Pour aller plus loin



On doit avoir :

1. $Y - (aX + b) \perp X \implies \langle Y - aX - b, X \rangle = \mathbb{E}((Y - aX - b)X) = 0$;
2. $Y - (aX + b) \perp \mathbb{1} \implies \langle Y - aX - b, \mathbb{1} \rangle = \mathbb{E}((Y - aX - b)1) = 0$;

et on retrouve les mêmes résultats...

Exercice 3 - 35 Table de Poisson et Scilab

Expliquez ce qui se passe :

```
function [tabp] = tab_poisson(precision)
    tabp=[];
    if precision == 1 then rang = 0.1:0.1:1;
    else if precision == 2 then rang = 1.1:0.1:2;
    else if precision == 3 then rang = 2.5:0.5:7
    else if precision == 4 then rang = 7.5:0.5:12
        end
    end
end
```

```

end
end
ligne=['k\l']
for l = rang do
    ligne = [ligne,string(l)]
end
tabp=[tabp;ligne]
for k = 0:7 do
    ligne=[string(k)];
    for l = rang do
        ligne = [ligne,string(exp(-1)*1^k/prod(1:k))];
    end
    tabp = [tabp;ligne]
end
endfunction

```

qui donne :

```

-->format(7)
-->tab_poisson(1)
ans =

!k\l  0.1    0.2    0.3    0.4    0.5    0.6    0.7    0.8    0.9    1      !
!
!0    0.9048  0.8187  0.7408  0.6703  0.6065  0.5488  0.4966  0.4493  0.4066  0.3679 !
!
!1    0.0905  0.1637  0.2222  0.2681  0.3033  0.3293  0.3476  0.3595  0.3659  0.3679 !
!
!2    0.0045  0.0164  0.0333  0.0536  0.0758  0.0988  0.1217  0.1438  0.1647  0.1839 !
!
!3    0.0002  0.0011  0.0033  0.0072  0.0126  0.0198  0.0284  0.0383  0.0494  0.0613 !
!
!4    4.D-06  5.D-05  0.0003  0.0007  0.0016  0.0030  0.0050  0.0077  0.0111  0.0153 !
!
!5    8.D-08  2.D-06  2.D-05  6.D-05  0.0002  0.0004  0.0007  0.0012  0.0020  0.0031 !
!
!6    1.D-09  7.D-08  8.D-07  4.D-06  1.D-05  4.D-05  8.D-05  0.0002  0.0003  0.0005 !
!
!7    2.D-11  2.D-09  3.D-08  2.D-07  9.D-07  3.D-06  8.D-06  2.D-05  4.D-05  7.D-05 !

-->tab_poisson(2)
ans =

!k\l  1.1    1.2    1.3    1.4    1.5    1.6    1.7    1.8    1.9    2      !
!
!0    0.3329  0.3012  0.2725  0.2466  0.2231  0.2019  0.1827  0.1653  0.1496  0.1353 !
!
!1    0.3662  0.3614  0.3543  0.3452  0.3347  0.3230  0.3106  0.2975  0.2842  0.2707 !
!
!2    0.2014  0.2169  0.2303  0.2417  0.2510  0.2584  0.2640  0.2678  0.2700  0.2707 !
!
!3    0.0738  0.0867  0.0998  0.1128  0.1255  0.1378  0.1496  0.1607  0.1710  0.1804 !
!
!4    0.0203  0.0260  0.0324  0.0395  0.0471  0.0551  0.0636  0.0723  0.0812  0.0902 !
!
!5    0.0045  0.0062  0.0084  0.0111  0.0141  0.0176  0.0216  0.0260  0.0309  0.0361 !
!
!6    0.0008  0.0012  0.0018  0.0026  0.0035  0.0047  0.0061  0.0078  0.0098  0.0120 !
!
!7    0.0001  0.0002  0.0003  0.0005  0.0008  0.0011  0.0015  0.0020  0.0027  0.0034 !

-->tab_poisson(3)
ans =

!k\l  2.5    3      3.5    4      4.5    5      5.5    6      6.5    7      !
!
!0    0.0821  0.0498  0.0302  0.0183  0.0111  0.0067  0.0041  0.0025  0.0015  0.0009 !
!
!1    0.2052  0.1494  0.1057  0.0733  0.0500  0.0337  0.0225  0.0149  0.0098  0.0064 !
!
!2    0.2565  0.2240  0.1850  0.1465  0.1125  0.0842  0.0618  0.0446  0.0318  0.0223 !
!
!3    0.2138  0.2240  0.2158  0.1954  0.1687  0.1404  0.1133  0.0892  0.0688  0.0521 !
!
!4    0.1336  0.1680  0.1888  0.1954  0.1898  0.1755  0.1558  0.1339  0.1118  0.0912 !
!
!5    0.0668  0.1008  0.1322  0.1563  0.1708  0.1755  0.1714  0.1606  0.1454  0.1277 !
!
!6    0.0278  0.0504  0.0771  0.1042  0.1281  0.1462  0.1571  0.1606  0.1575  0.149 !
!
!7    0.0099  0.0216  0.0385  0.0595  0.0824  0.1044  0.1234  0.1377  0.1462  0.149 !

```

```

-->tab_poisson(4)
ans =
!k\l 7.5    8      8.5    9      9.5    10     10.5   11     11.5   12     !
!
!0    0.0006 0.0003 0.0002 0.0001 7.D-05 5.D-05 3.D-05 2.D-05 1.D-05 6.D-06 !
!
!1    0.0041 0.0027 0.0017 0.0011 0.0007 0.0005 0.0003 0.0002 0.0001 7.D-05 !
!
!2    0.0156 0.0107 0.0074 0.0050 0.0034 0.0023 0.0015 0.0010 0.0007 0.0004 !
!
!3    0.0389 0.0286 0.0208 0.0150 0.0107 0.0076 0.0053 0.0037 0.0026 0.0018 !
!
!4    0.0729 0.0573 0.0443 0.0337 0.0254 0.0189 0.0139 0.0102 0.0074 0.0053 !
!
!5    0.1094 0.0916 0.0752 0.0607 0.0483 0.0378 0.0293 0.0224 0.0170 0.0127 !
!
!6    0.1367 0.1221 0.1066 0.0911 0.0764 0.0631 0.0513 0.0411 0.0325 0.0255 !
!
!7    0.1465 0.1396 0.1294 0.1171 0.1037 0.0901 0.0769 0.0646 0.0535 0.0437 !

```

4

Initiation aux processus aléatoires



Андрей Андреевич МАРКОВ (1856 - 1922), grand mathématicien russe, initiateur des probabilités modernes, grand amateur de poésie, n'hésita pourtant pas à s'impliquer dans la tumultueuse histoire de son pays. Il n'eut pas peur d'affirmer publiquement son anti-tsarisme et fut renvoyé de l'Université. Il fut réhabilité après la Révolution.

Nous étudierons de manière élémentaire quelques résultats de ses travaux sur les processus aléatoires (ou stochastiques), c'est-à-dire l'évolution dans le temps d'une variable aléatoire appelés *chaînes de MARKOV* dont les applications sont très nombreuses, notamment en informatique (reconnaissance de la langue d'un texte, techniques de compression, l'algorithme PageRank de Google, files d'attente, bioinformatique,...)

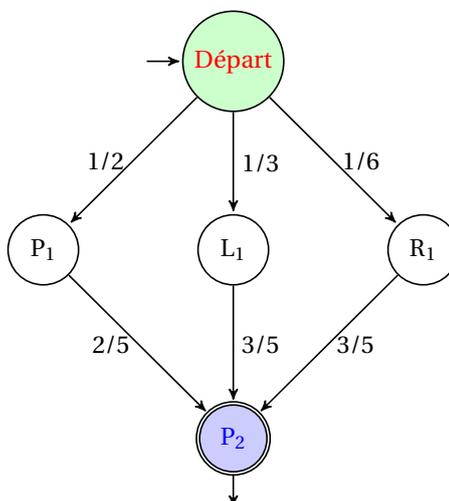
La présentation proposée ici est celle du célèbre mathématicien et pédagogue allemand Arthur ENGEL.

D'un point de vue plus probabiliste, les chaînes de MARKOV vont nous permettre d'étudier des v.a.r. non indépendantes ce que nous ne pouvions pas envisager jusqu'à maintenant.

1 Vision dynamique des probabilités - Automates

1 1 Découverte

Considérons l'habituelle urne et ses boules : on y trouve 3 boules pervenche, 2 lila et 1 rose. On tire successivement et sans remise deux boules de l'urne : quelle est la probabilité d'obtenir une boule pervenche ? On peut représenter ce jeu par l'automate suivant :

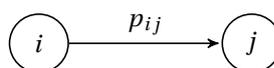


Les résultats habituels conduisent au résultat attendu :

$$\mathbb{P}(P_2) = \frac{1}{2} \times \frac{2}{5} + \frac{1}{3} \times \frac{3}{5} + \frac{1}{6} \times \frac{3}{5} = \frac{1}{2}$$

L'orientation de l'automate nous permet de dire que la probabilité de A à B est égale à la somme de tous les chemins allant de A vers B.

On parle souvent dans ce cas de *processus aléatoire* ou de *marche aléatoire* : il s'agit d'une succession d'expériences aléatoires simples. Une dimension temporelle vient donc s'ajouter aux outils statiques habituels et induit une vision *dynamique* des probabilités. L'utilisation d'automates apparaît alors naturelle car elle est associée aux notions d'états et de changements d'états :



Le nombre p_{ij} représente la probabilité de passer de l'état i à l'état j .

Parmi tous les processus aléatoires, un processus de MARKOV est caractérisé par la connaissance :

- de l'état initial;
- de tous les états;
- de tous les p_{ij} .

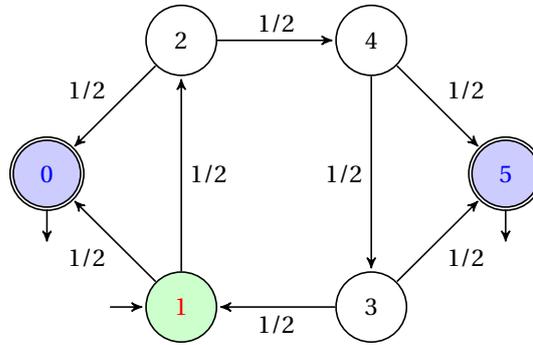
La particularité des chaînes de MARKOV est de représenter des processus où la probabilité permettant de passer à un instant donné d'un état présent à un état futur ne dépend du passé qu'à travers le présent soit : « la probabilité du futur sachant présent et passé est égale à la probabilité du futur sachant le présent ». ou encore : le présent contient toutes les données du passé...

1 2 Premier exemple - Règles de parcours

Voyons à présent un exemple plus précis pour introduire quelques définitions.

Considérons le jeu suivant : je dispose d'un Zwjrsch (monnaie syldave) mais je dois absolument disposer de 5 Zwjrsch pour inviter Pavlova à la patinoire. Je propose un jeu à mon riche voisin Ivan Ivanovitch : je lance ma pièce équilibrée de 1 Zwjrsch. Je mise tout l'argent que je possède si j'ai 1 ou 2 Zwjrsch ou uniquement le complément à 5 si j'ai 3 ou 4 Zwjrsch (je ne suis pas totalement stupide ayant suivi l'enseignement du célèbre mathématicien français Konanovitch au contraire de mon voisin).

On peut représenter ce jeu par l'automate suivant :



Un état i est dit *absorbant* si, et seulement si, $p_{ii} = 1$. Ici, les états 5 et 0 sont absorbants (j'y suis, j'y reste). L'ensemble $\{0, 5\}$ des états absorbants constituent le *bord* de l'ensemble des états. Les autres états sont dits *intérieurs*.

Une chaîne de MARKOV est *absorbante* lorsque son bord est non vide.

On peut alors se demander quelle est la probabilité que j'accompagne Pavlova à la patinoire et quelle sera la durée moyenne du jeu.

Nous allons donner trois règles qui vont nous guider tout au long de notre étude et dont les deux premières reprennent les résultats sur le conditionnement. Elles concernent les *parcours* ou *promenades aléatoires* le long de l'automate. La *longueur* d'un parcours désigne le nombre de transitions séparant les deux extrémités.

1. la probabilité de réaliser un parcours donné à partir d'un point donné est égale au produit de toutes les probabilités de transition le long de ce parcours ;
2. La probabilité d'atteindre un sous-ensemble du bord à partir d'un état donné est égale à la somme des probabilités de tous les parcours allant de cet élément aux éléments du sous-ensemble ;
3. la durée moyenne des parcours allant d'un état au bord (le *temps moyen d'absorption*) est la moyenne des longueurs des parcours de ce sommet au bord pondérée par les probabilités de chaque transition. Si on introduit une variable aléatoire égale au temps de parcours, il s'agit donc de l'espérance de cette v.a.r.

Le problème, c'est qu'il peut y avoir une infinité de parcours absorbants pour un sommet donné comme dans l'exemple proposé que nous allons pouvoir à présent étudier.

Nous allons étudier tous les parcours absorbants partant de 1, c'est-à-dire tous les parcours allant de 1 à 5. Il y en a bien une infinité du fait de la présence du cycle $1 \rightarrow 2 \rightarrow 4 \rightarrow 3 \rightarrow 1$. La dernière partie du parcours peut être $1 \rightarrow 2 \rightarrow 4 \rightarrow 5$ ou $1 \rightarrow 2 \rightarrow 4 \rightarrow 3 \rightarrow 5$ de longueurs respectives 3 et 4. Ainsi, la probabilité p_1 cherchée vaut :

$$\begin{aligned}
 p_1 &= \underbrace{\left(\frac{1}{2^3} + \frac{1}{2^4}\right)}_{0 \text{ boucle}} + \underbrace{\left(\frac{1}{2^{3+4}} + \frac{1}{2^{4+4}}\right)}_{1 \text{ boucle}} + \underbrace{\left(\frac{1}{2^{3+2 \times 4}} + \frac{1}{2^{4+2 \times 4}}\right)}_{2 \text{ boucles}} + \underbrace{\left(\frac{1}{2^{3+3 \times 4}} + \frac{1}{2^{4+3 \times 4}}\right)}_{3 \text{ boucles}} + \dots \\
 &= \left(\frac{3}{2} \times \frac{1}{2^3}\right) + \left(\frac{3}{2} \times \frac{1}{2^7}\right) + \left(\frac{3}{2} \times \frac{1}{2^{11}}\right) + \left(\frac{3}{2} \times \frac{1}{2^{15}}\right) + \dots \\
 &= 3 \sum_{k=0}^{+\infty} \left(\frac{1}{2^4}\right)^k \\
 &= 3 \frac{1}{1 - \frac{1}{16}} \\
 &= \frac{1}{5}
 \end{aligned}$$

Vous voici une nouvelle fois confronté au problème d'une somme infinie à valeur finie...

Recherche

Faites un calcul similaire pour déterminer la probabilité de perdre : vous connaissez déjà le résultat donc vous pourrez contrôler vos calculs...

Introduisons à présent la v.a.r. égale au temps de parcours. Il y a quatre « familles » de parcours :

1. $1 \rightarrow 0$ précédé de k cycles : $X = 4k + 1$, $\mathbb{P}([X = 4k + 1]) = \left(\frac{1}{2^4}\right)^k \times \frac{1}{2} = \frac{1}{2^{4k+1}}$;
2. $1 \rightarrow 2 \rightarrow 0$ précédé de k cycles : $X = 4k + 2$, $\mathbb{P}([X = 4k + 2]) = \left(\frac{1}{2^4}\right)^k \times \left(\frac{1}{2}\right)^2 = \frac{1}{2^{4k+2}}$;
3. $1 \rightarrow 2 \rightarrow 4 \rightarrow 5$ précédé de k cycles : $X = 4k + 3$, $\mathbb{P}([X = 4k + 3]) = \left(\frac{1}{2^4}\right)^k \times \left(\frac{1}{2}\right)^3 = \frac{1}{2^{4k+3}}$;
4. $1 \rightarrow 2 \rightarrow 4 \rightarrow 3 \rightarrow 5$ précédé de k cycles : $X = 4k + 4$, $\mathbb{P}([X = 4k + 4]) = \left(\frac{1}{2^4}\right)^k \times \left(\frac{1}{2}\right)^4 = \frac{1}{2^{4k+4}}$.

Dans tous les cas, $\mathbb{P}([X = n]) = \frac{1}{2^n}$ donc $E(X) = \sum_{n=1}^{+\infty} \frac{n}{2^n}$.

Nous avons prouvé au module précédent que :

$$\forall q \in \mathbb{R}, |q| < 1 \implies S = \sum_{n \geq 0} q^n = \frac{1}{1-q}$$

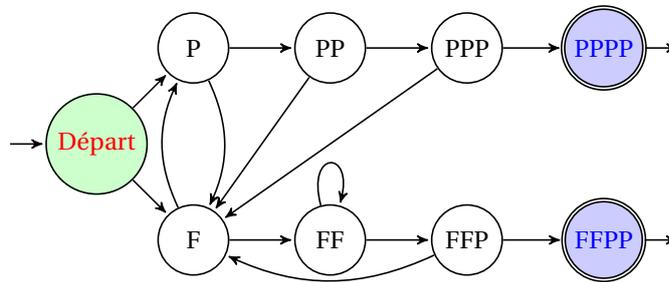
Pour $|q| < 1$, notons $T = \sum_{n \geq 1} nq^{n-1}$ en supposant que cette écriture a un sens.

Calculer $(1-q)T$ en fonction de S et déduisez-en que $T = \frac{1}{(1-q)^2}$ puis que $E(X) = 2$.

Recherche

1 3 Deuxième exemple - Valeur moyenne

Voyons à présent un nouveau jeu : je lance une pièce. Si j'obtiens quatre fois de suite « Pile », mon voisin Ivan Ivanitch a gagné. Si j'obtiens deux fois « Face » puis deux fois « Pile », j'ai gagné. Je présente ce jeu comme équitable à Ivan car la probabilité de gagner à chaque fois est $(\frac{1}{2})^4 = \frac{1}{16}$. Suis-je malhonnête ? Représentons le jeu à l'aide d'un automate. On ne représentera pas les probabilités de transitions car elles valent toutes 1/2 : ma malhonnêteté éventuelle n'est pas due à l'emploi d'une pièce truquée...



Le problème, c'est que cette fois, il y a trop de cycles pour pouvoir utiliser les simples outils de parcours vus précédemment mais nous allons partir d'eux pour trouver des résultats plus puissants en les considérant par « paquets ».

Soit une chaîne de MARKOV dont l'ensemble des états est $S = \{1, 2, \dots, n\}$.

Nous allons tout d'abord nous intéresser à la probabilité p_i pour un état i d'être « absorbé » dans un sous-ensemble SB du bord.

Sur le bord, $p_i = 1$ si $i \in SB$ et 0 sinon.

À l'intérieur, en utilisant la règle des probabilités totales ou bien la deuxième règle du parcours pour obtenir que :

$$p_i = \sum_{k \in S} p_{ik} p_k$$

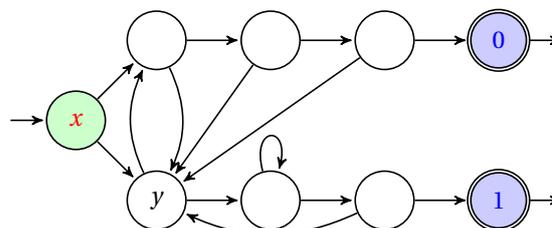
Or les p_{ik} sont nuls s'il n'existe pas de transition de i vers k . Notons \mathcal{E}_i l'ensemble des extrémités des transitions d'origine i , alors :

$$p_i = \sum_{k \in \mathcal{E}_i} p_{ik} p_k$$

Appliquons ce résultat à notre jeu.

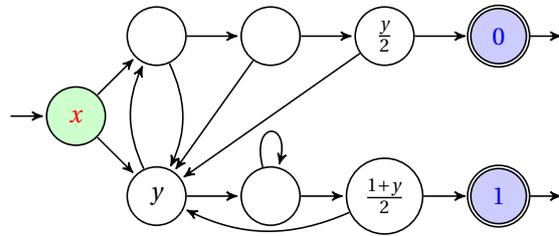
Appelons x la probabilité que je gagne au départ. Il me manque une donnée pour appliquer la règle précédente donc j'appelle y la probabilité que je gagne depuis l'état F.

Je connais également $p_{PPPP} = 0$ et $p_{FFPP} = 1$ puisque je m'intéresse au sous-ensemble du bord qui me fait gagner.

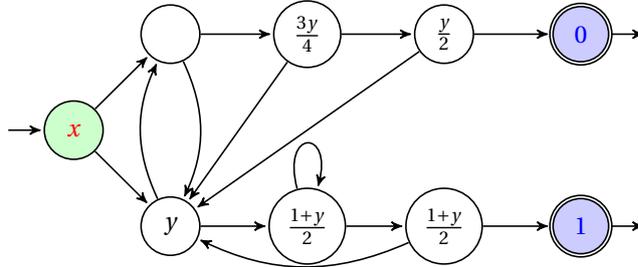


Remplissons alors de proche en proche l'automate en remplissant les cercles avec les p_i .

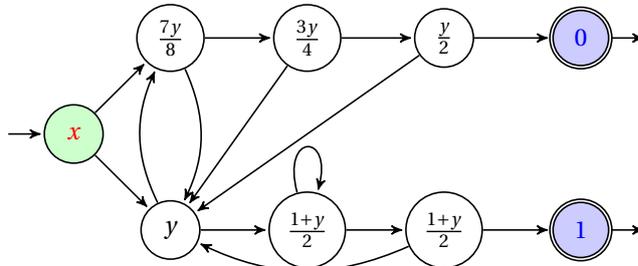
Je peux remplir les deux états précédents les bords : $p_{PPP} = \frac{1}{2} \times 0 + \frac{1}{2} \times y$ et $p_{FFP} = \frac{1}{2} \times 1 + \frac{1}{2} \times y$



Alors $p_{PP} = \frac{1}{2} \times \frac{y}{2} + \frac{1}{2} \times y = \frac{3y}{4}$ et $p_{FF} = \frac{1+y}{4} + \frac{1}{2} p_{FF} = \frac{1+y}{2}$



Ici, $p_P = \frac{3y}{8} + \frac{y}{2} = \frac{7y}{8}$

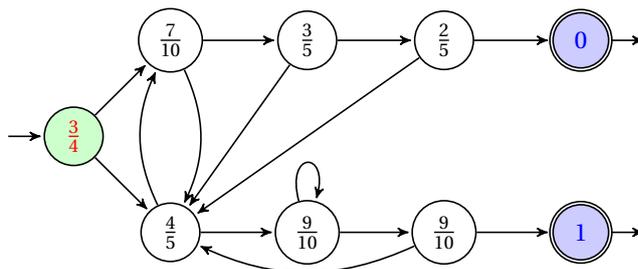


Il nous reste à déterminer y et x mais nous n'avons pas exploité l'état F :

$$y = p_F = \frac{1+y}{4} + \frac{7y}{16} = \frac{11y+4}{16} \text{ d'où } y = \frac{4}{5}.$$

$$\text{Finalement, } x = \frac{1}{2}y + \frac{7y}{16} = \frac{15}{16} \times \frac{4}{5} = \frac{3}{4}.$$

Je gagne donc avec une probabilité de $\frac{3}{4}$: avec un peu de mathématiques, je peux donc embrouiller mon voisin Ivan Ivanitch.



Je peux maintenant me demander combien de fois en moyenne je vais lancer la pièce avant d'atteindre un état absorbant.

Recherche

Démontrez que $m_i = \sum_{k \in \mathcal{E}_i} p_{ik}(1 + m_k)$.

Déduisez-en que la durée moyenne d'une partie avec Ivan Ivanitch est de 12 lancers.

2

Comportement asymptotique des chaînes de Markov

Zlot, Brzxxx et Morzgniouf sont trois villes situées respectivement en Syldavie, Bordurie et Bouzoukstan. Des trafiquants de photos dédicacées du groupe ABBA prennent leur marchandise le matin dans n'importe laquelle de ces villes pour l'apporter le soir dans n'importe quelle autre. On notera pour simplifier V_1, V_2 et

V_3 ces villes et p_{ij} la probabilité qu'une marchandise prise le matin dans la ville V_i soit rendue le soir dans la ville V_j . La matrice $P = (p_{ij})_{\substack{1 \leq i \leq 3 \\ 1 \leq j \leq 3}}$ est appelée *matrice de transition* de la chaîne de MARKOV.

Définition 4 - 1

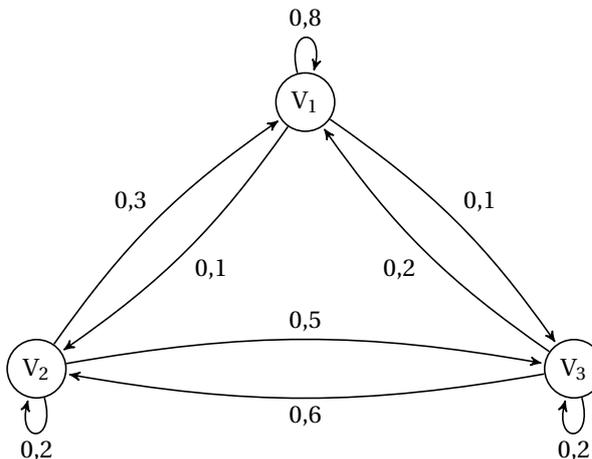
Matrice stochastique

Une matrice dont tous les coefficients sont positifs ou nuls et dont la somme des coefficients de chaque ligne vaut 1 est appelée **matrice stochastique**.

Si, de plus, la somme des colonnes vaut 1, la matrice est dite **bistochastique**.

Que peut-on dire d'une matrice de transition ?

Voici le diagramme de transition des trafiquants suédophiles :



Il n'y a pas d'état absorbant a priori. Il faut donc changer notre fusil d'épaule. Nous allons en fait utiliser le calcul matriciel. La matrice de transition est :

$$\begin{pmatrix} 0,8 & 0,1 & 0,1 \\ 0,3 & 0,2 & 0,5 \\ 0,2 & 0,2 & 0,6 \end{pmatrix}$$

On notera $x_i^{(k)}$ la proportion de trafiquants qui se trouvent au matin du jour k dans la ville V_i .

Définition 4 - 2

Vecteur d'état

On appelle *vecteur d'état* tout élément (x_1, \dots, x_n) de \mathbb{R}^n tel que $x_1 + \dots + x_n = 1$.

Ainsi, $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, x_3^{(k)})$ est un vecteur d'état (pourquoi?).

Recherche

Démontrez que :

$$\forall k \in [1, n], \quad x^{(k)} = x^{(k-1)} \cdot P$$

puis que :

$$\forall k \in [1, n], \quad x^{(k)} = x^{(0)} \cdot P^k$$

Supposons que le chef de la mafia locale dispose de 1000 trafiquants qui partent tous le matin du jour 0 de la ville de Zlot. Quelle sera la proportion de trafiquants dans chacune des villes au bout d'une semaine ? d'un an ?

Il s'agit donc de calculer des puissances successives de P.

On obtient les proportions au bout d'une semaine en calculant $x^{(0)} \cdot P^7$ avec $x^{(0)} = (1000, 0, 0)$.

Nous allons utiliser SCILAB pour les longs calculs :

```
--> A = [0.8,0.1,0.1;
-->      0.3,0.2,0.5;
-->      0.2,0.2,0.6]
A =
    0.8    0.1    0.1
    0.3    0.2    0.5
    0.2    0.2    0.6
```

```

-->X = [1000,0,0]
X =

    1000.    0.    0.

-->X*A
ans =

    800.    100.    100.

-->X*A^7
ans =

    546.9038    144.5711    308.5251

-->X*A^(365)
ans =

    536.58537    146.34146    317.07317

```

Il semble que la proportion reste stable au bout d'un certain temps :

```

-->A^(1000)
ans =

    0.5365854    0.1463415    0.3170732
    0.5365854    0.1463415    0.3170732
    0.5365854    0.1463415    0.3170732

-->A^(10000)
ans =

    0.5365854    0.1463415    0.3170732
    0.5365854    0.1463415    0.3170732
    0.5365854    0.1463415    0.3170732

```

On observe que cette matrice « à l'infini » a trois lignes identiques. Étudions le produit d'un vecteur ligne de « A_∞ » par A :

```

-->Ainfini = A^(10000)
Ainfini =

    0.5365854    0.1463415    0.3170732
    0.5365854    0.1463415    0.3170732
    0.5365854    0.1463415    0.3170732

-->v = Ainfini(1,:)
v =

    0.5365854    0.1463415    0.3170732

-->v*A
ans =

    0.5365854    0.1463415    0.3170732

```

Troublant....En fait, ce résultat se généralise.

Définition 4 - 3

Matrice de transition régulière

Une matrice de transition A est dite régulière si, et seulement si, il existe un entier naturel r tel que A^r a ses coefficients tous **strictement** positifs.

Théorème 4 - 1

Théorème des chaînes de MARKOV régulières

Soit A une matrice régulière. Il existe une matrice stochastique :

$$A_{\infty} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_1 & a_2 & \cdots & a_n \\ \vdots & \vdots & \vdots & \vdots \\ a_1 & a_2 & \cdots & a_n \end{pmatrix}$$

telles que $\lim_{n \rightarrow +\infty} A^n = A_{\infty}$.

De plus, $v = (a_1, a_2, \dots, a_n)$ est l'unique vecteur d'état stationnaire (tel que $v \cdot A = v$).

Nous démontrerons ce résultat après un peu de pratique du calcul matriciel.

3 Exercices

3 1 Chaînes absorbantes - automates

Exercice 4 - 1 Ascension



Vaclav veut faire l'ascension du mont Plexy-Gladz, plus haut sommet de Bordurie (1808 m). Il décide de passer une nuit au refuge du Guide Suprême (1010 m) et également au refuge de La Grande Moustache (1507 m). Chaque matin il écoute le bulletin météo. S'il est bon, il poursuit son ascension jusqu'à l'étape suivante, sinon, il redescend d'une étape. On suppose que Vaclav est initialement au refuge du Guide Suprême et que s'il est obligé de redescendre au refuge de Szprinkoth (778 m), il abandonne son projet d'ascension. Enfin, il faut savoir qu'en Bordurie, la météo est donnée en fonction d'un lancer de dé : on annonce du beau temps si la face marque V ou VI, et de la pluie si la face marque I, II, III ou IV.

Calculez la probabilité que Vaclav arrive au sommet et la durée moyenne de son ascension en jours.

Exercice 4 - 2 Tennis

Deux joueurs s'affrontent dans une partie de tennis. Chaque point joué est gagné par le joueur A avec une probabilité de $3/5$, sinon c'est B qui gagne. On suppose que l'issue de chaque échange est indépendante des autres.

Initialement, les joueurs sont à 40 partout. Calculez la probabilité que A gagne le jeu et la durée moyenne du jeu en nombre d'échanges.

Exercice 4 - 3 Palindrome

On dispose d'une roue de foire non truquée qui affiche 1 avec la probabilité p et 0 sinon. On lance la roue et on note le résultat en continuant tant qu'on n'a pas obtenu un mot symétrique (un palindrome) comme 11, 101, 110010011, etc. Attention ! 01010 ne convient pas car la roue s'arrête à 010.

On supposera que 1 et 0 ne sont pas des palindromes. Quelle est la durée moyenne d'attente pour obtenir un palindrome ?

Exercice 4 - 4 Jeu

Vladimir choisit un mot constitué de 0 et de 1 et Anastasia un autre mot. On lance ensuite une pièce de monnaie équilibrée ($P=1$, $F=0$) jusqu'à ce qu'un des deux mots choisis apparaisse. Le gagnant est celui dont le mot apparaît en premier. Ce jeu est-il équilibré dans les cas suivants :

- | | | |
|----------------------------|---------------------------|--------------------------|
| 1. $v = 00, a = 111$; | 4. $v = 1101, a = 0101$; | 7. $v = 111, a = 0011$; |
| 2. $v = 000, a = 111111$; | 5. $v = 1010, a = 0011$; | 8. $v = 010, a = 1001$; |
| 3. $v = 00, a = 101$; | 6. $v = 101, a = 1001$; | 9. $v = 010, a = 111$. |

Exercice 4 - 5 Contrôle des naissances

Pour lutter contre la surpopulation, le gouvernement bordure teste plusieurs mesures. Il est décrété que les parents auront des enfants jusqu'à ce que :

- ils aient une fille ;
- les deux sexes soient représentés ;
- les deux sexes soient représentés deux fois ;
- les deux sexes soient représentés et que le premier sexe réapparaisse (5 filles, 8 garçons puis une fille par exemple).

Exercice 4 - 6 Craps



Le craps, ou la passe anglaise, est le jeu de dés le plus populaire des États-Unis. On jette deux dés et on additionne les numéros sortis. Si on obtient 7 ou 11 on gagne, si on obtient 2, 3 ou 12 on perd.

Dans les autres cas, on rejoue jusqu'à obtenir à nouveau la même somme sans obtenir 7.

Dressez l'automate du craps en distinguant 6 états : « défaite », « victoire », « début », « 5 ou 9 », « 4 ou 10 », « 6 ou 8 ».

Calculez la probabilité de gagner et la durée moyenne d'une partie.

3 2 Avec le calcul matriciel**Exercice 4 - 7 Matrice bistochastique**

Déterminez un vecteur d'état simple stationnaire pour toutes les distributions bistochastiques.

Exercice 4 - 8 LMD

Voici la matrice de transition des passages $S_1 \leftrightarrow S_2 \leftrightarrow S_3$ des promos de l'IUT d'informatique de Klow :

$$\begin{pmatrix} 0,45 & 0,48 & 0,07 \\ 0,05 & 0,70 & 0,25 \\ 0,01 & 0,50 & 0,49 \end{pmatrix}$$

Calculez la probabilité de passer de S_3 en S_1 en 2 semestres, et celle de passer de S_1 en S_3 en deux semestres. Calculez également les probabilités stationnaires (l'IUT, c'est tellement bien, on voudrait y passer sa vie...).

Exercice 4 - 9 Reprise

Résolvez les exercices précédents à l'aide du calcul matriciel (on ne traitera pas les questions sur les durées moyennes).

5 Loi normale



La loi normale ou loi de Laplace-Gauss ou loi gaussienne est la loi la plus utilisée pour modéliser des phénomènes aléatoires. D'une part, parce qu'elle a des propriétés très pratiques et « confortables ».

D'autre part et surtout, le très important *théorème central limite* permet de prouver qu'un grand nombre de « petites » variables aléatoires de même loi centrée, réduites et indépendantes converge « en un certain sens » vers une variable aléatoire gaussienne.

En Russie, Gauss s'invite même dans les manifs : la bonne connaissance de la loi normale a permis aux citoyens de mettre en évidence la fraude électorale...

1 Densité de probabilité

Nous avons précédemment vu au point 8 du théorème 3 - 7 page 37 qu'en tout point où F_X est continue, on a $\mathbb{P}_X(a) = 0$.

Si F_X est continue sur \mathbb{R} , on va admettre qu'alors F_X est dérivable « presque partout » c'est-à-dire partout sauf sur des éléments B de $\mathbf{B}(\mathbb{R})$ tels que $\mathbb{P}(B) = 0$.

On dit alors que la v.a.r. est continue.

On appelle **densité de probabilité** d'une v.a.r. continue une fonction f_X positive telle que :

$$\mathbb{P}(X \leq x) = \mathbb{P}_X(]-\infty, x]) = F_X(x) = \int_{-\infty}^x f_X(t) dt$$

Il va donc falloir manipuler quelques intégrales, qui plus est à bornes éventuellement infinies...

Nous sommes en amphi, voici les résultats principaux à connaître...

[Saut dans le temps]

...et voilà, nous savons tout sur ces intégrales :-)

2 Tout sur la loi normale

Loi normale

Une v.a.r. suit une loi normale (ou gaussienne) si, et seulement si, sa fonction de densité est :

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

On note alors $X \rightsquigarrow \mathcal{N}(\mu, \sigma)$.

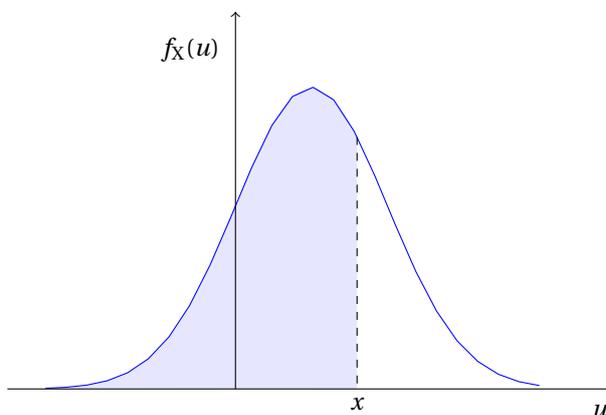
Définition 5 - 1

Nous sommes en amphi, voici les résultats principaux à connaître...

[Saut dans le temps]

...et voilà, nous savons tout sur cette loi :-)

Un petit dessin :



Un petit résumé :

t	0.0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.5279	0.53188	0.53586
0.1	0.53983	0.5438	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.6293	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.6591	0.66276	0.6664	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.7054	0.70884	0.71226	0.71566	0.71904	0.7224
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.7549
0.7	0.75804	0.76115	0.76424	0.7673	0.77035	0.77337	0.77637	0.77935	0.7823	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.8665	0.86864	0.87076	0.87286	0.87493	0.87698	0.879	0.881	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.9032	0.9049	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.9222	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.9452	0.9463	0.94738	0.94845	0.9495	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.9608	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.9732	0.97381	0.97441	0.975	0.97558	0.97615	0.9767
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.9803	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.983	0.98341	0.98382	0.98422	0.98461	0.985	0.98537	0.98574
2.2	0.9861	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.9884	0.9887	0.98899
2.3	0.98928	0.98956	0.98983	0.9901	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.9918	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.9943	0.99446	0.99461	0.99477	0.99492	0.99506	0.9952
2.6	0.99534	0.99547	0.9956	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.9972	0.99728	0.99736
2.8	0.99744	0.99752	0.9976	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861

On veut calculer $\mathbb{P}(X < 1.96)$ sachant que $X \sim \mathcal{N}(1, 0.5)$. Comme la variable réduite $T = \frac{X-1}{0.5}$ suit la loi $\mathcal{N}(0, 1)$ alors $X < 1.96$ équivaut à $T < 1.920000$.

On lit dans la table de la loi normale $\mathbb{P}(T < 1.920000) \approx \mathbb{P}(T < 1.92) \approx 0.97257$

CHAPITRE

6 Simulation informatique et probabilités

