Analyse combinatoire
Probabilités

Variables aléatoires

Couples aléatoires

Génération de nombres aléatoires

Introduction aux statistiques

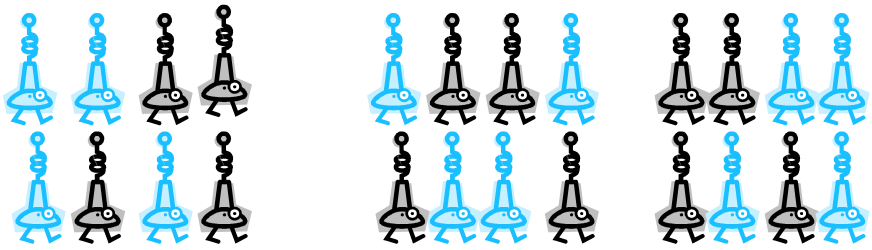
unité: MA3R11

Références bibliographiques

- Cours basé sur 2 ouvrages:
  - « Initiation aux probabilités » de Sheldon M.Ross,
    Ed : Presses polytechniques romandes
  - « Probabilités pour scientifiques et ingénieurs » de Patrick Bogaert,
    Ed: de boeck

Analyse combinatoire

- Introduction par un exemple :
  - On considère un système de communications composé de n antennes identiques alignées. Ce système n'est fonctionnel que s'il n'y a pas 2 antennes consécutives défectueuses. On suppose qu'il y a m antennes défectueuses. Quelle est la probabilité pour que le système reste fonctionnel?
  - Solution (graphique) pour n=4, m=2: p=1/2



Principe fondamental de dénombrement

- Si r expériences à réaliser avec :
  - n_1 résultats possibles pour la 1ère
  - n_2 résultats possibles pour la 2ème
  - …
  - n_r résultats possibles pour la rème
- Alors il y a $n_1 n_2 \ldots n_r$ résultats possibles pour les r expériences prises ensemble

Exercice1: nombre de plaques minéralogiques dans un département si le matricule comporte 4 chiffres et 3 lettres (en plus des 2 chiffres du département)

Exercice 2: combien de fonctions sur n points à deux valeurs possibles par point?

Permutations

- Théorème: le nombre de permutations de n objets distinguables est n!
- Théorème: le nombre de permutations de n objets parmi lesquels n_1 sont indistinguables, n_2 autres également,…, n_r autres également est: $n!/(n_1!\ n_2!\ldots\ n_r!)$
  - Exemple: pepper : 6!/(3!2!)permutations

Arrangements - Combinaisons

- n(n-1)…(n-r+1)= n!/(n-r)! représente le nombre de manières de choisir r objets parmi n. C'est le nombre d'arrangements de r objets parmi n noté A_n^r
- Ce nombre comporte toutes les permutations des r objets choisis. Si l'ordre des r objets n'a pas d'importance il faut diviser ce nombre par r!
- n(n-1)…(n-r+1)/r!= n!/(r!(n-r)!) représente le nombre de combinaisons de r objets parmi n et est noté C_n^r

Exercice des n antennes

- Identité remarquable

$$C_n^r = C_{n-1}^{r-1} + C_{n-1}^r \quad 1 \le r < n$$

11
121
1331 triangle de Pascal
14641
....

- Théorème du binôme

$$(x+y)^n = \sum_{k=0}^{n} C_n^k x^k y^{n-k}$$

Exercice : démontrer l'identité remarquable

Exercice : démontrer le théorème du binôme par récurrence

Exercice: combien y a-t-il de sous ensembles d'un ensemble à n éléments?

Coefficients multinomiaux

- Problème: un ensemble de n objets distints doit être divisé en r groupes de tailles respectives $n_1,n_2,\ldots,n_r$ (avec $n_1+\ldots+n_r=n$). De combien de manières peut on le faire?

$$C_n^{n_1} C_{n-n_1}^{n_2} \ldots C_{n-n_1-n_2\ldots-n_{r-1}}^{n_r} = \frac{n!}{n_1!n_2!\ldots n_r!}$$

- Ce nombre est appelé coefficient multinomial et noté : $C_n^{n_1,n_2,\ldots,n_r}$
- Théorème multinomial :

$$\left(x_1+x_2+\ldots+x_r\right)^n = \sum_{(n_1,n_2,\ldots,n_r):n_1+\ldots+n_r=n} C_n^{n_1,n_2,\ldots,n_r} x_1^{n_1} x_2^{n_2} \ldots x_r^{n_r}$$

- Combien de termes dans la somme? Combien de solutions entières ≥0 telles que $n_1+n_2+\ldots+n_r=n$
- Étape: solutions >0 : il faut choisir r-1 séparateurs parmi n-1 possibles : C_{n-1}^{r-1}
- Solutions ≥0 :
  même nombre que solutions >0 de:
  $n'_1+n'_2+\ldots+n'_r=n+r$ c'est-à-dire : C_{n+r-1}^{r-1}

Exercice: retour au problèmes des antennes: n antennes dont m défectueuses:

On aligne d'abord les m antennes défectueuses

n_1 antennes Ok à gauche de la 1ere défectueuse

n_2 antennes Ok à droite de la 1ere défectueuse et gauche de la 2eme défectueuse

…

n_{m+1} à droite de la meme défectueuse

Il faut $n_i>0$ pour i = 2 à m , $n_1 \geq 0$, $n_{m+1} \geq 0$ pour que le système fonctionne:
nombre de cas : sol >0 de $n'_1+n_2+\ldots+n_m+n'_{m+1} = n-m+2$: C_{n-m+1}^{m}

Axiomes des probabilités

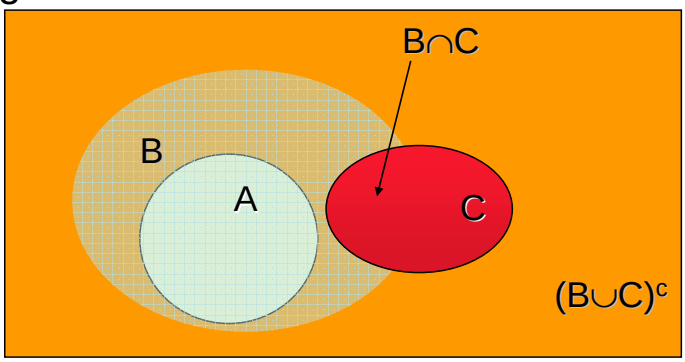
Ensemble fondamental - évènement

- Ensemble des résultats possibles d'une expérience aléatoire : c'est l'ensemble fondamental noté Ω
  - Ex: expérience = jeter 2 pièces
    Ω ={(P,P),(P,F),(F,P),(F,F)}
- Tout sous ensemble de Ω est un évènement:
  - A$\subset \Omega$ est un évènement
  - $\omega \in \Omega$ est un résultat possible ou évènement élémentaire
  - Ex: la première pièce montre pile
    A= {(P,P),(P,F)}

Relations entre évènements

- Manière la plus simple de les représenter : diagrammes de Venn

Ω



- A⊂B (exemple d'inclusion)
  - ⇒ A∩B=A
- A∩C=∅ (définition de l'incompatibilité)

Complémentarité: l'évènement complémentaire d'un évènement A est l'évènement A^c tel que:

-$A \cap A^c = \varnothing$

-$A \cup A^c = \Omega$

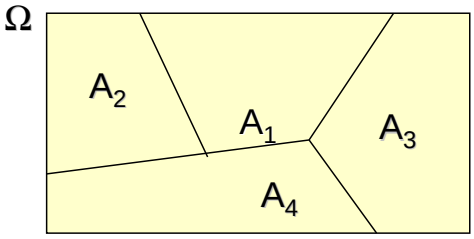
Propriétés des opérations sur les évènements

- Commutativité
  - $E \cap F = F \cap E$
  - $E \cup F = F \cup E$
- Associativité
  - $(E \cap F) \cap G = E \cap (F \cap G)$
  - $(E \cup F) \cup G = E \cup (F \cup G)$
- Distributivité
  - $(E \cap F) \cup G = (E \cup G) \cap (F \cup G)$
  - $(E \cup F) \cap G = (E \cap G) \cup (F \cap G)$
- Lois de Morgan
  - $(\cup_i E_i)^c = \cap_i E_i^c$
  - $(\cap_i E_i)^c = \cup_i E_i^c$

On peut montrer ces propriétés par des diagrammes de Venn

Système complet d'évènements ou partition

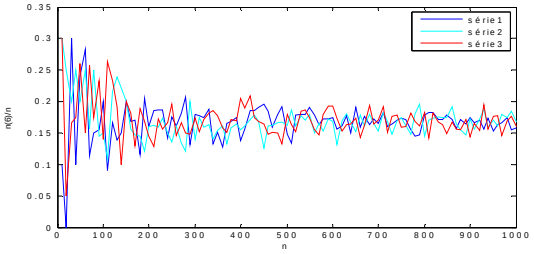
- Former une partition de Ω consiste à « couper » l'ensemble Ω en évènements incompatibles: à l'issue de l'expérience un et un seul de ces évènements sera réalisé



- Définition formelle : L'ensemble $\{A_1,A_2,\ldots,A_n\}$ est une partition de Ω si et seulement si:
  - Aucun des évènements n'est impossible: $A_i \neq \varnothing \ \forall i$
  - Les évènements sont incompatibles 2 à 2: $A_i \cap A_j = \varnothing \ \forall i \neq j$
  - L'union des évènements est l'ensemble fondamental: $(\cup_i A_i) = \Omega$

Probabilité d'un évènement

- Approche intuitive (définition en fréquence): on réalise n fois l'expérience sous les mêmes conditions. Pour chaque évènement A de Ω on définit n(A) comme le nombre de fois où A survient. Alors la probabilité de A est définie par:
  - $P(A)=\lim_{n\to\infty} n(A)/n$: fréquence limite de A
  - Exemple : lancé de dé avec A=« on obtient 6 »



  - Problème théorique : la limite existe-t-elle ?
  - Problème pratique : réaliser un très grand nombre de fois l'expérience

« Expériences » avec Matlab

- Nombreux « tutoriels » matlab disponibles sur internet
  - Un « maison »: http://intra.esiee.fr/sigtel/intro_matlab.pdf
  - D'autres
  - https://moodle.polymtl.ca/file.php/408/GuideMatlab.pdf
  - http://www.ann.jussieu.fr/~postel/matlab/
  - http://www.ann.jussieu.fr/~joly/mattex/Matlab1.pdf
  - http://www.civil.usherbrooke.ca/cours/gci101/Hiver%202010/Matlab/Tutoriel%201%20Matlab.pdf
- Je signale en fin de certaines pages de commentaires (exp matlab) les commandes matlab qui permettent de réaliser des expériences associées aux différents éléments du cours.
- Pour avoir une aide sur une commande faire:
  help nom_de_la_commande

Définition axiomatique

- La probabilité est un nombre satisfaisant les 3 axiomes de base suivants:
  - La probabilité d'un évènement A quelconque est positive ou nulle : $P(A) \geq 0$
  - La probabilité de l'évènement certain Ω est égale à 1: $P(\Omega)=1$
  - Si deux évènements A et B sont incompatibles , la probabilité de leur union est égale à la somme de leur probabilité: $P(A \cup B) = P(A)+P(B)$ si $A \cap B = \varnothing$

- Propriétés qui découlent de cette définition:
  - La probabilité de l'évènement impossible ∅ est égale à 0 : $P(\emptyset)=0$
  - La probabilité est un nombre inférieur ou égal à 1: $P(A)\leq 1$
  - La probabilité de l'évènement complémentaire est égale à 1 moins la probabilité de l'évènement: $P(A^c)=1-P(A)$
  - La probabilité de l'union d'un nombre quelconque d'évènements incompatibles deux à deux est égale à la somme de leur probabilité: $P(\cup_i A_i)=\sum_i P(A_i)$ si $A_i \cap A_j = \emptyset \ \forall i \neq j$
- La probabilité de l'union de deux évènements est égale à la somme de leur probabilité moins la probabilité de leur intersection:
  - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Généralisation:
  - $P(\cup_{i=1:n} A_i)=\sum_i P(A_i) - \sum_{i1<i2} P(A_{i1} \cap A_{i2}) + \ldots + (-1)^{r+1} \sum_{i1\leq j2<\ldots<ir} P(A_{i1} \cap A_{i2} \ldots \cap A_{ir}) + \ldots + (-1)^{n+1} P(A_1 \cap A_2 \ldots \cap A_n)$

Exercice: démonstration des propriétés

Lien avec l'analyse combinatoire

- Si les évènements élémentaires de Ω sont équiprobables et en nombre fini $\omega_1, \omega_2, \ldots, \omega_N$ avec $P(\omega_i)=1/N \ \forall i$, alors pour tout évènement A la probabilité de A est égale à la proportion dans l'ensemble Ω des évènements élémentaires contenus dans A:
  - $P(A)= Card(A)/Card(\Omega)$
- En clair cette définition possible de la probabilité ne s'applique qu'à des cas « simples »: dés, jeux de cartes,…

Exercice: Si n personnes sont présentes dans une pièce, quelle est la probabilité que leurs anniversaires tombent sur des jours tous différents? Quelle valeur faut il donner à n pour que cette probabilité descende en dessous de ½?

Exp matlab:

```
N=10000;
d1=randint(1,N,[1 6]);%d1 va contenir le résultat de 1000 lancés de dé
pd1=hist(d1,[1:6])/N;
figure;subplot(211);plot(d1);%trace les 1000 résultats
subplot(212);stem([1:6],pd1)%trace l'histogramme des résultats
%lancé d'un autre dé N fois
d2=randint(1,N,[1 6]);
pd2=hist(d2,[1:6])/N;
figure;subplot(211);plot(d2);%trace les 1000 résultats
subplot(212);stem([1:6],pd2)%trace l'histogramme des résultats
```

Probabilité conditionnelle et indépendance

Probabilité conditionnelle

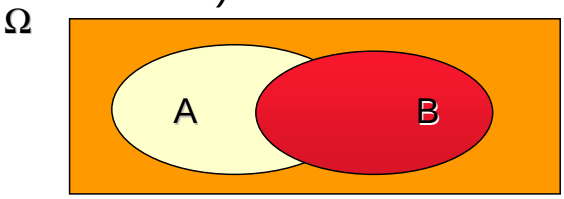
- Souvent une partie de l'information concernant le résultat de l'expérience est disponible
- Même quand ce n'est pas le cas il peut être intéressant d'utiliser le « détour » de certaines probabilités conditionnelles pour calculer la probabilité d'un évènement
- Exemple: lancé de 2 dés : $\Omega=\{(i,j),i=1\ldots6,j=1\ldots6\}$
  - Quelle est la probabilité d'avoir la somme des 2 dés égale à 8 ?
  - Si on sait que le premier dé donne 3 quelle est la probabilité d'avoir la somme des 2 dés égale à 8?
  - Si on sait que la somme est égale à 8 quelle est la probabilité pour que le premier dé donne 3?

exp Matlab: (suite de la page 20):

```
s=d1+d2;%somme des deux résultats pour chacune des N expériences
i_s_egal_8=find(s==8);%indices du vecteur s pour lesquels la valeur est égale à 8
proba_s_egal_8=length(i_s_egal_8)/N
i_d1_egal_3=find(d1==3);
i_s_egal_8_si_d1_egal_3=find(s(i_d1_egal_3)==8);
proba_s_egal_8_si_d1_egal_3=length(i_s_egal_8_si_d1_egal_3)/length(i_d1_egal_3)
i_s_egal_8=find(s==8);
i_d1_egal_3_si_s_egal_8=find(d1(i_s_egal_8)==3);
proba_d1_egal_3_si_s_egal_8=length(i_d1_egal_3_si_s_egal_8)/length(i_s_egal_8)
```

Probabilité conditionnelle

- Soient A et B deux évènements pour une expérience. Si B est réalisé, l'ensemble des résultats possibles n'est plus Ω mais B (ensemble fondamental réduit).



- La probabilité de A est alors celle de l'évènement A∩B rapportée à la probabilité de B.
- On en déduit que: La probabilité P(A/B) de A sachant que B est réalisé, appelée probabilité conditionnelle de A si B est donnée par:
  - P(A/B)=P(A∩B)/P(B)

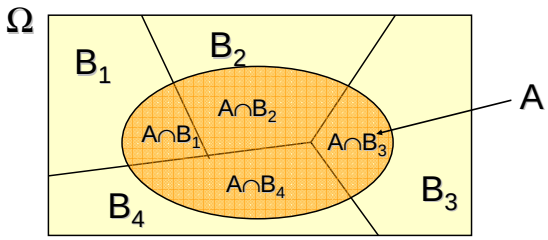
- La probabilité conditionnelle satisfait les 3 axiomes de base d'une probabilité:
  - $P(A/B) \geq 0$
  - $P(\Omega/B)=1$
  - $P(A \cup C/B) = P(A/B)+P(C/B)$ si $A \cap C=\varnothing$
- Théorème des probabilités composées
  - $P(A \cap B)=P(B)\ P(A/B) = P(A)\ P(B/A)$
  - Généralisation pour des évènements $A_1,A_2,\ldots,A_n$ quelconques:
    $P(A_1 \cap \ldots \cap A_n) = P(A_1)P(A_2/A_1)P(A_3/ A_1 \cap A_2) \ldots P(A_n/ A_1 \cap \ldots \cap A_{n-1})$
- Formule (théorème) de Bayes
  - $P(A/B)=P(A)P(B/A)\ / P(B)$

Exercice: Mathieu hésite entre suivre un cours de maths et un cours d'électronique. Il préfère l'électronique mais il estime qu'il a une probabilité 1/2 d'avoir une note A en maths et 1/3 d'avoir une note A en électronique. Il prend sa décision en tirant à pile ou face.

Quelle est la probabilité qu'il obtienne A en électronique?

Théorème des probabilités totales
Théorème de Bayes généralisé

- Si l'ensemble $\{B_1,B_2,\ldots,B_n\}$ est une partition de Ω alors (théorème des probabilités totales):
  - $P(A)=\sum_{i=1:n}P(B_i)P(A/B_i)$



- On a alors (théorème de Bayes généralisé):
  - $P(B_i/A) = P(B_i)\ P(A/B_i)\ /\ (\sum_{j=1:n}P(B_j)P(A/B_j))$

Exercice: dans une société de vente par téléphone, les vendeurs sont tellement occupés que seuls 60% des clients téléphonant pour la première fois obtiennent la communication avec l'un d'entre eux. On demande aux autres de laisser leur n° de téléphone. Trois fois sur quatre un vendeur trouve le temps de rappeler le jour même, autrement le rappel a lieu le lendemain. L'expérience a montré que la probabilité qu'un client potentiel passe effectivement commande est 0.8 s'il a pu parler directement avec un vendeur et que cette probabilité tombe à 0.5 et 0.3 respectivement s'il y a eu rappel du client le jour même ou le lendemain.

a) Quel pourcentage de clients potentiels passent commande ?

b) Quel pourcentage de ceux qui ont passé commande n'ont pas eu à attendre qu'on les rappelle?

Indépendance

- Si la probabilité conditionnelle de A si B est réalisé est égale à la probabilité de A, A est dit indépendant de B :
  - P(A/B)=P(A)
  - On a alors : $P(A \cap B)=P(B)\,P(A)$ qui implique que B est indépendant de A
- Deux évènements A et B sont indépendants (on note $A \perp B$) si et seul^{t} si:
  - $P(A \cap B)=P(B)\,P(A)$
- Propriétés: Si $A \perp B$
  - $A \perp B^{c}$
  - $A^{c} \perp B$
  - $A^{c} \perp B^{c}$

Indépendance totale de n évènements

- Un ensemble d'évènements $A_1,A_2,\ldots,A_n$ est dit totalement indépendant si pour tout sous ensemble $A_{1'},A_{2'},\ldots,A_{r'}$ ($r \leq n$):
  - $P(A_{1'} \cap \ldots \cap A_{r'}) = P(A_{1'})\ldots P(A_{r'})$
- Exemple: système à n composants en //
  - qui fonctionne si l'un des composants fonctionne
  - dont les n composants sont indépendants et fonctionnent avec une probabilité p_i (i=1 à n)
  - $P(\text{fonctionnement}) = 1-\Pi_i(1-p_i)$

Variables aléatoires

Variables aléatoires

- Dans de nombreux cas on peut associer à un événement un nombre réel. On parle alors de variable aléatoire.
- Points abordés
  - Définition
  - Variable aléatoire discrète
  - Variable aléatoire continue
  - Loi de probabilité conditionnelle

Exemples de variables aléatoires:

Le résultat d'un lancé de dé

La taille des individus

Le nombre de communications sur un réseau téléphonique en une minute

Définitions

- Une variable aléatoire (v.a.) X est une application de Ω dans $\mathbb{R}$ telle qu'à tout ω de Ω correspond une valeur $X(\omega)=x$.
- L'ensemble $R_X \subset \mathbb{R}$ des valeurs possibles de X est appelé domaine de variation ou support de X
- Remarque
  - L'application n'est pas forcément bijective (elle est surjective)
    - exemple du lancé de dé: $R_X=\{1,2,3,4,5,6\}$ ou bien $R_X=\{-1,0,1\}$
- A tout sous-ensemble F de R_X correspond un évènement sur Ω : $E=X^{-1}(F)=\{\omega / X(\omega) \in F\}$
- Cette correspondance permet de définir la loi de probabilité de la v.a. X :

$$P_X(F)=P(X^{-1}(F))=P(X \in F)$$

Exemple: jeu de fléchettes , variables aléatoires possibles:

-distance de la fléchette au centre de la cible

-coordonnées de la position de la fléchette

-nombres de lancers à effectuer avant d'atteindre le centre

Exp matlab: nombre de lancés avant d'avoir un 6

```
N=10000;%nombre d'expériences
nb_lances_pour_obtenir_6=ones(1,N);
for n=1:N
d_different_de_6=1;
while d_different_de_6
   d=randint(1,1,[1 6]);
   if d==6
      d_different_de_6=0;
   else
      nb_lances_pour_obtenir_6(n)=nb_lances_pour_obtenir_6(n)+1;
   end
end
end
maxnb=max(nb_lances_pour_obtenir_6);%valeur maximale obtenue
pnb=hist(nb_lances_pour_obtenir_6,[1:maxnb])/N
figure;stem([1:maxnb],pnb)
```

- Dans de nombreux cas l'application de Ω dans R_X est « naturelle »
  - Température
  - Nombre d'appels téléphoniques sur une durée …
- Dans d'autres cas il n'est pas évident de définir une variable aléatoire ayant du sens (respectant une notion d'ordre et de distance entre évènements)
  - Risque de défaillance d'un composant
  - Bases d'une séquence ADN

Variables aléatoires discrètes

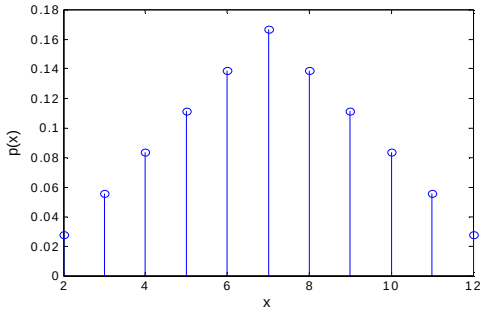
- Lorsque R_X est un ensemble fini (de taille N) ou dénombrable $\{x_1,\ldots,x_n,..\}$, X est une v.a. discrète.
- La loi de probabilité de X est définie par l'ensemble $\{p_1,\ldots,p_n,..\}$
  - $p_i = P(X^{-1}(x_i))$ (noté aussi $P(X= x_i)$ ou $p(x_i)$)
  - Propriétés:
    - $p_i \geq 0$
    - $\sum_{i=1:N} p_i = 1$
      - Démonstration: les $X^{-1}(x_i)$ forment une partition de Ω
- Pour tout $B \subset R_X$: $P_X(B) = \sum_{\{i, xi \in B\}} p_i$

Exercice: on répète le jet d'une pièce jusqu'à ce que pile apparaisse mais au plus n fois. Les jets sont indépendants et pile apparaît avec la probabilité p. X désigne le nombre de jets réalisés jusqu'à l'arrêt de l'expérience.

1) Domaine de variation de X
2) Loi de probabilité de X

3) Exp Matlab: programmez cette expérience (avec p=1/2) en vous inspirant du script donné p 30

- Représentation sous forme d'histogramme
  - Ex: somme de 2 dés
  - R_X={2,…,12}



exp matlab : dans le script donné p 20 et 22 il suffit de rajouter :

ps=hist(s,[2:12])/N;

figure;stem([2:12],ps)

Espérance et variance

- Espérance E(X) ou m
  - Moyenne statistique
  - Pour une v.a. discrète
    $E(X)=m=\sum_i p_i x_i$
- Variance Var(X) ou σ^2:
  - Mesure la dispersion autour de la moyenne, c'est l'écart carré moyen à la moyenne
  - Pour une v.a. discrète
    $Var(X)=\sigma^2=\sum_i p_i(x_i-m)^2$

Exp matlab:

La moyenne s'obtient par la fonction mean et l'écart type σ (racine carrée de la variance) s'obtient par la fonction std. Par exemple pour le nombre de lancés de dés nécessaire pour obtenir 6 (page 30) on peut obtenir ces grandeurs par les instructions suivantes:

M=mean(nb_lances_pour_obtenir_6)

sigma=std(nb_lances_pour_obtenir_6)

Lois discrètes les plus utilisées

Loi de Bernouilli : X ~ Be(p)

- Ω ={A=« réussite », A*=« échec »}
  R_X={1,0}
- P(X=1)=p et P(X=0)=q=1-p
- Exemples:
  - Pile ou face
  - Examen de fin d'unité (avoir ou non la moyenne)
  - Transmission ou non sur un slot dans un réseau
- Moyenne et variance
  - m = p , σ^2 = p(1-p)

Exercice: montrer que σ^2 = p(1-p). Pour quelle valeur de p cette variance est elle maximale?

Exp Matlab:

```
Nexp=1000;
p=0.1;
x=randsrc(Nexp,1,[0 1;1-p p]) ;
p_est=hist(x,[0:1])/Nexp;
stem([0:1],p_est)
M=mean(x)
Sigma=std(x)
```

Loi binomiale : X ~ Bi(n,p)

- Répétition (n fois) de l'épreuve de Bernouilli:
  - X : nombre de réussites parmi n épreuves indépendantes de Bernouilli de même loi (p)
  - v.a. à valeurs dans $R_X = \{0,1,2,\ldots,n\}$
- Loi de probabilité d'une variable binomiale Bi(n,p):
  - $p_i = C_n^i \, p^i (1-p)^{n-i}$ pour $i \in \{0,1,\ldots, n\}$
- Moyenne: m=np
- Variance: $\sigma^2 = np(1-p)$

Toute séquence comportant i succès et n-i échecs est de probabilité $p^i(1-p)^{n-i}$ car les épreuves sont indépendantes.

Il y a C_n^i séquences (parmi les 2^n séquences de n valeurs 0 ou 1 possibles) qui comporte i succès.

Exercice: démontrer que m=np et que $\sigma^2 = np(1-p)$

Exercice: roue de la fortune

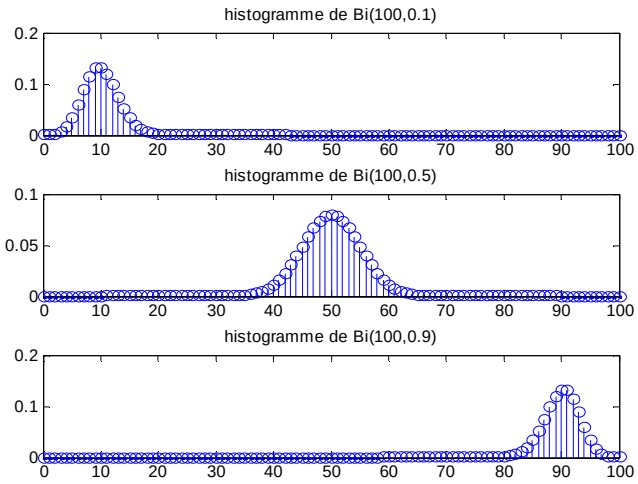
Un joueur parie sur un numéro entre 1 et 6. On jette ensuite 3 dés. Si le nombre choisi par le joueur apparait i fois (i=1,2,3) celui-ci gagne i euros, sinon il perd 1 euro. Ce jeu est il honnête?

Exp Matlab:

```
Nexp=10000;
n=100;
p=0.1;
x_ber=randsrc(Nexp,n,[0 1;1-p p]);
x_bin=sum(x_ber,2);
p_est=hist(x_bin,[0:n])/Nexp;
stem([0:n],p_est)
M=mean(x_bin)
Sigma=std(x_bin)
```

Histogrammes pour la loi Binomiale

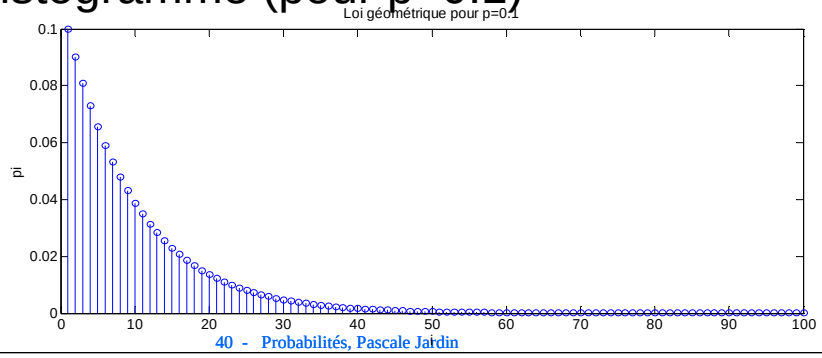
- Lorsque k croit de 0 à n, p_k croît puis décroit le pic étant atteint pour k=[(n+1)p]
- Histogrammes pour n=100
  - p=0.1
  - p=0.5
  - p=0.9



Exercice: Faire le rapport p_{k+1}/p_k

Loi géométrique

- Répétition de l'épreuve de Bernouilli
- X est le nombre d'épreuves nécessaires pour obtenir un premier succès: $R_X=\mathbb{N}^*$
- $p_i=p(1-p)^{i-1}$ pour i=1,2,…
- Histogramme (pour p=0.1)



Exercice: montrer que $\Sigma p_i=1$

Exp matlab: l'expérience proposée p 30 (nb de lancés d'un dé pour obtenir 6) correspond à une loi géométrique de paramètre p=1/6

Loi de Poisson : X ~ Po(μ)

- v.a: comptage d'évènements rares indépendants sur un espace ou une durée fixée.
- $p_i=(\mu)^i \exp(-\mu)/i!$ pour $i \in \mathbb{N}$
- μ (sans dimension) représente le produit d'une intensité λ et d'une grandeur physique (longueur, surface, volume ou intervalle de temps)
  - Ex sur le temps : nombre d'accidents de circulation mortels par jour $\mu = \lambda t$
  - Ex sur une longueur : nombre de défauts le long d'un câble en acier

Loi de Poisson: exemples: décompte des appels sur un central téléphonique, comptage des voitures à un péage. C'est la loi la plus utilisée dans les études sur les files d'attente dans les réseaux de communication. Le paramètre λ caractérise le débit moyen du flux sur la durée fixée.

Exercice: le nombre d'éruptions dangereuses du Vésuve suit une loi de Poisson: X~Po(λt) avec λ=0.0003 (l'unité de temps étant l'année)

Quelle est la probabilité d'observer au moins une éruption durant :

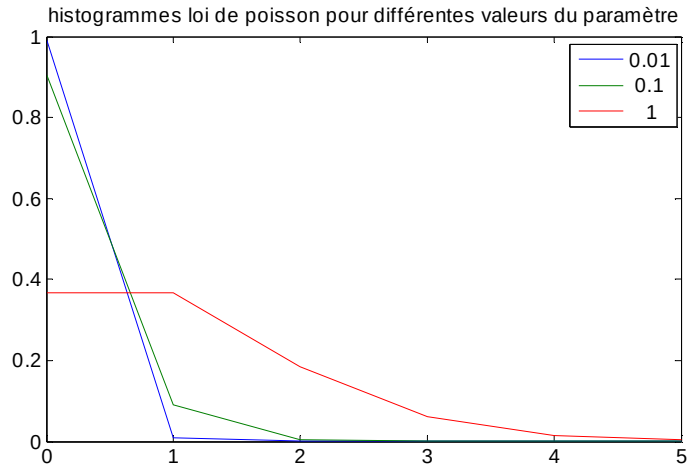
a) le siècle à venir

b) le millénaire à venir

Exp Matlab:

```
N=10000;
mu=1;
x_poisson = poissrnd(mu,[1 N]);
x_max=max(x_poisson);
p_poisson = hist(x_poisson,[0:x_max])/N;
stem([0:x_max],p_poisson)
```

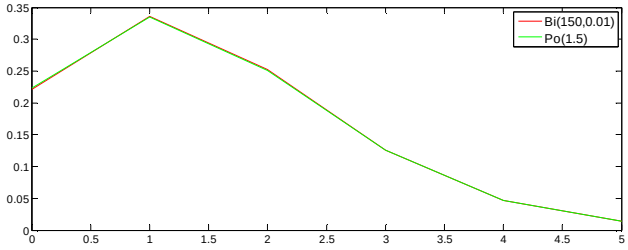
- Moyenne: $m = \mu$
- Variance: $\sigma^2 = \mu$
- Histogrammes de la loi de Poisson



Exercice: montrer que $m = \mu$ et $\sigma^2 = \mu$

Approximation de la loi binomiale

- La loi de Poisson est une approximation de la loi binomiale lorsque n est grand et p petit
- Si X ~ Bi(n,p) avec p faible et np fini
  X ~ Po(np)
- Histogrammes



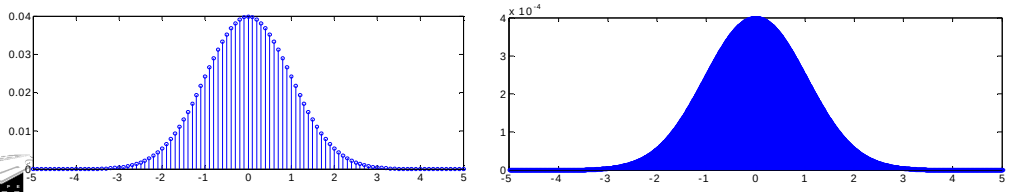
Exercice: démontrer cette propriété

Exercice: dans une société de 150 personnes si on estime à 0.01 la probabilité qu'un employé soit absent de manière ponctuelle et imprévisible, quelle est la probabilité qu'il y ait 1, 2, 3 personne(s) absente(s) un jour donné ?

Variable aléatoire continue

Variable aléatoire continue

- Le domaine de variation de X est $R_X \subset \mathbb{R}$
  - Parce que le nombre de valeurs possibles est très grand dans un intervalle limité
  - Parce que la v.a. est de nature continue (par exemple la pression, la température,…)
- Conséquence: si il y a une infinité non dénombrable de valeurs possibles de la v.a. X p(x) =P(X=x)=0 pour tout x de R_X



Remarque: La distinction v.a. discrète et v.a. continue n'est pas toujours pertinente

-En pratique la précision des appareils de mesure limite le nombre de valeurs possibles même pour les processus de nature continue

-Au fur et à mesure que le nombre de valeurs possibles augmente on a tendance à approximer la loi discrète par une loi continue

Exemple: taille des individus

Au cm près

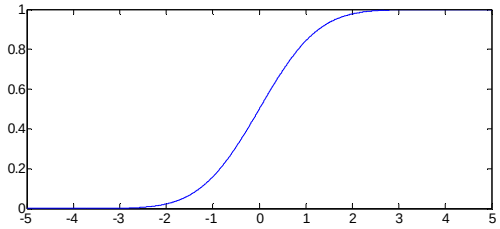
Au mm près

Au micron près

Au nanomètre près

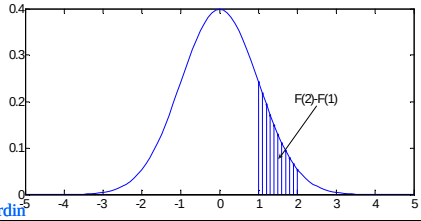
Fonction de répartition

- C'est la fonction qui associe à chaque valeur x de la v.a. X la probabilité de ne pas excéder cette valeur : $F(x)=P(X \leq x)$
- Propriétés:
  - F est croissante: si $x<y$ $F(x) \leq F(y)$
  - $P(x<X \leq y)=F(y)-F(x)$
  - F est continue sur $\mathbb{R}$



Fonction de densité de probabilité

- On appelle fonction de densité de probabilité la fonction f(x) telle que:
  - $f(x) = \lim_{\Delta x \to 0} \frac{P(x < X \leq x + \Delta x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{F(x + \Delta x) - F(x)}{\Delta x}$
  - $F(x) = \int_{-\infty}^{x} f(u) du$
- Propriétés
  - $f(x) \geq 0$
  - $\int_{-\infty}^{+\infty} f(x) dx = 1$
  - $P(a < X < b) = \int_{a}^{b} f(x) dx$



Espérance et variance pour une v.a. continue

- Espérance mathématique ou moyenne statistique
  - $$E(X) = m_X = \int_{-\infty}^{\infty} x f(x)\, dx$$
- Variance (dispersion autour de la moyenne):
  - $$Var(X) = \sigma_X^2 = E\left((X - m_X)^2\right) = \int_{-\infty}^{+\infty} (x - m_X)^2 f(x)\, dx$$
- Écart-type: σ_X

Inégalités

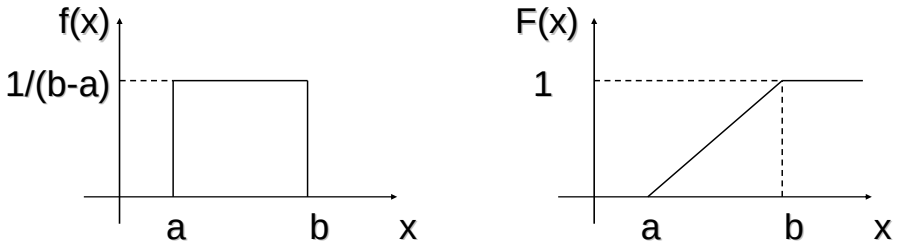
- Inégalité de Markov
  - Valable si f(x)=0(ou p(x)=0) $\forall x<0$
  - $P(X \geq a) \leq E(X)/a \qquad \forall a>0$
- Inégalité de Chebyshev
  - Valable pour toute v.a. X
  - $P(|X-m_X| \geq a) \leq Var(X)/a^2$
  - En particulier:
    - $P(|X-m_X| \geq n\sigma_X) \leq 1/n^2$

L'inégalité de Chebyshev donne une borne « lâche » majorant la probabilité d'avoir une valeur éloignée de la moyenne.

Ex: $p(|X-m|>2\sigma)<1/4$: « au moins les ¾ des valeurs de X se trouvent entre $m-2\sigma$ et $m+2\sigma$ »

Fonctions de densité les plus utilisées

- Loi uniforme sur un intervalle : X ~ Un(a,b)
  - f(x)=1/(b-a) pour x ∈ [a,b] (et =0 ailleurs)



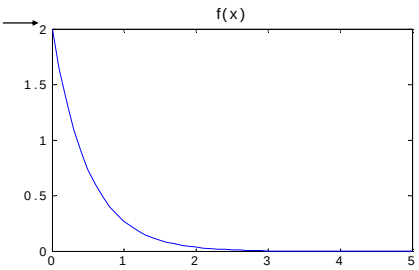
- Ex: bruit de quantification: Q ~ Un(-q/2,q/2)

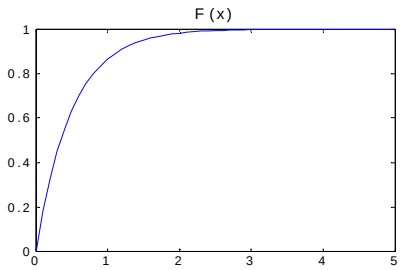
Exercice: probabilité d'attendre le métro (plus de 30s, plus de 3mn) s'il y en a un toutes les 5mn

Exercice: déterminer la moyenne et la variance d'une v.a. uniforme

- Loi exponentielle: X ~ Exp(λ)
  - f(x)= $\lambda e^{-\lambda x}$ pour x ≥0 (=0 pour x<0)
  - F(x)=1- $e^{-\lambda x}$ pour x ≥0 (=0 pour x<0)

λ →





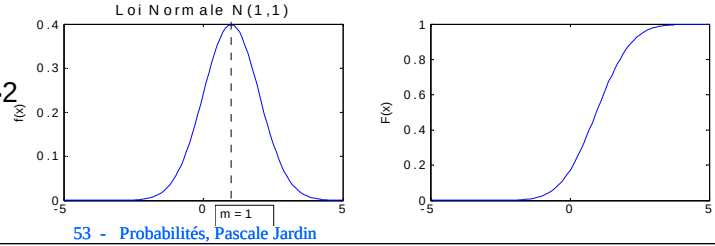
- La loi exponentielle modélise bien le temps séparant deux évènements consécutifs dans la théorie des files d'attente.

Exercice: on admet que dans la théorie des files d'attente, le nombre Y d'occurrences d'un événement sur un intervalle de temps suit une loi de Poisson Y~Po(λ). Montrer alors que le temps X séparant deux évènements consécutifs suit une loi exponentielle.

Exercice: Montrer que la moyenne vaut $1/\lambda$ et la variance vaut $1/\lambda^2$ (utiliser l'intégration par parties)

Loi Normale (ou de Gauss)

- Loi très importante en raison du théorème central limite exprimé de façon simplifiée par:
  - Si une v.a. X est la somme de n v.a. indép^{tes} Y_1, Y_2,…,Y_n de même loi , alors la loi de X tend vers une loi normale quand n grand (n ~50 suffit)
- Loi Normale : X ~$\mathcal{N}(m,\sigma^2)$
  - $f(x)= 1/(2\pi\sigma^2)^{1/2} \exp(-(x-m)^2/2\sigma^2)$
- Moyenne et variance:
  - E(X)=m
  - Var(X)= σ^2



Commentaires:

-pas de primitive pour f(x) donc pas de formule analytique pour F(x)

-F tabulée pour la loi normale $\mathcal{N}(0,1)$ et si X ~ $\mathcal{N}(m,\sigma^2)$, $Z=(X-m)/\sigma \sim \mathcal{N}(0,1)$

Loi de probabilité conditionnelle

- Si on définit un événement A tel que $P(A)\neq 0$, la fonction de répartition conditionnelle F(x/A) est définie par:
  - $F(x/A)=P(X\leq x/A)=P((X\leq x)\cap A)/P(A)$
- Pour X v.a. discrète:
  - $p(x/A)=P((X=x)\cap A)/P(A)$
- Pour X v.a. continue:
  - $f(x/A)=dF(x/A)/dx = P((x<X\leq x+dx) \cap A)/P(A)/dx = f(x)P(A/x)/P(A)$
    - avec la notation $P(A/x)=P(A/ x<X\leq x+dx)$

p(x/A) est une loi de probabilité et f(x/A) est une densité de probabilité

On peut définir leur moyenne, leur variance

- Théorème de Bayes
  - P(A/x)=P(A)f(x/A)/f(x)
- Théorème des probabilités totales
  - $$P(A) = \int_{-\infty}^{+\infty} f(x) P(A/x)\, dx$$
- Théorème de Bayes généralisé (version continue)
  - $$f(x/A) = \frac{f(x) P(A/x)}{\int_{-\infty}^{+\infty} f(x) P(A/x)\, dx}$$

Cas où A est défini par rapport à X
A≡(a<X<b)

- $F(x/(a<X<b)) = \begin{cases} 0 & \text{si } x \le a \\ (F(x)-F(b))/(F(b)-F(a)) & \text{si } a<x\le b \\ 1 & \text{si } x>b \end{cases}$
- $p(x/(a<X<b)) = \begin{cases} p(x)/(F(b)-F(a)) & \text{si } a \le x \le b \\ 0 & \text{autrement} \end{cases}$
- $f(x/(a<X<b)) = \begin{cases} f(x)/(F(b)-F(a)) & \text{si } a \le x \le b \\ 0 & \text{autrement} \end{cases}$
- Exemple: loi de Gauss $X \sim \mathcal{N}(m,\sigma^2)$ avec

  $A = (m<X<m+\sigma)$

Cas où A est défini par rapport à une autre loi que celle de X

- théorème des probabilités totales:
  - N hypothèses A_i formant une partition. $F(x)=\sum_{i=1:n}P(A_i)F(x/A_i)$
  - pour une v.a. discrète:
    - $p(x)=\sum_{i=1:n}P(A_i)p(x/A_i)$
  - pour une v.a. continue:
    - $f(x)=\sum_{i=1:n}P(A_i)f(x/A_i)$

Fonction d’une variable aléatoire

Y=h(X) avec h strictement monotone

- Cas discret: $P(Y=y)=P(h(X)=y)=P(X=h^{-1}(y))$
- Cas continu: $F_Y(y)=\begin{cases} F_X(h^{-1}(y)) & \text{si } h\uparrow \\ 1-F_X(h^{-1}(y)) & \text{si } h\downarrow \end{cases}$
- Si la fonction h^{-1} est dérivable
  - $$f_Y(y)=f_X\left(h^{-1}(y)\right)\left|\frac{dh^{-1}(y)}{dy}\right|$$
- Cas particulier: Y=aX+b
  - $$f_Y(y)=f_X\left(\frac{y-b}{a}\right)\left|\frac{1}{a}\right|$$

Il est courant que la v.a. d'intérêt ne soit pas directement observable (ou mesurable) mais que la variable observable soit liée à celle qui nous intéresse par l'intermédiaire d'une fonction connue. Le problème est de trouver la loi de la v.a. d'intérêt connaissant celle de la variable observable.

Espérance et Variance de Y=h(X)

- Espérance: E(Y)=E(h(X))
  - Cas discret

$$E(Y) = \sum_i h(x_i) p_i$$

  - Cas continu

$$E(Y) = \int h(x) f(x) dx$$

- Variance: Var(Y)=Var(h(X))
  - $Var(Y)=E(h^2(X)) - E^2(h(X))$

Exercice: Information Y=log(1/p(X)). Calcul de l'entropie qui est l'information moyenne

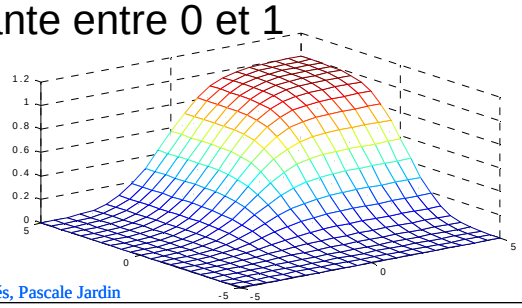
Couples aléatoires

Intérêt

- Il est courant que l'on ait à étudier deux variables aléatoires liées entre elles
  - Position (X,Y) aléatoire
  - Entrée X $\rightarrow$ Sortie Y d'un canal de transmission
  - Variable aléatoire dépendant du temps (processus aléatoire) X_t , on s'intéresse au lien entre les v.a. à deux instants différents $(X_t,X_{t'})$
- L'étude est généralisable au cas d'un nombre quelconque de v.a.
  - Vecteur aléatoire $\underline{X}=(X_1,X_2,\ldots, X_n)^T$

Domaine de variation conjoint – Fonction de répartition

- Domaine de variation R_{XY} d'un couple aléatoire (X,Y): ensemble des couples (x,y) de valeurs possibles
  - $R_{XY} \subset R_X \times R_Y$
- Fonction de répartition
  - $F(x,y)=P(X \leq x \cap Y \leq y)$
  - Fonction croissante entre 0 et 1



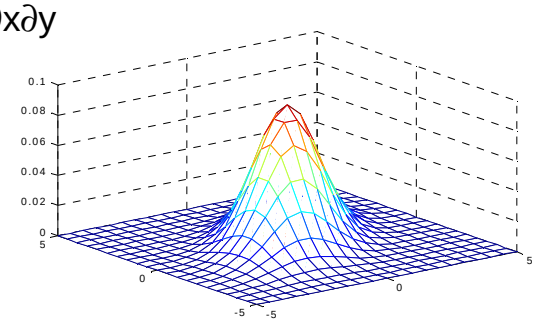
Exercice: domaine de variation R_{XY} pour X et Y à valeurs dans $R_X = R_Y = \{0,1,\ldots,N\}$ sous deux contraintes:

-$X+Y \leq N$ (ex: famille de 4 enfants au plus, N=4, X nb de filles, Y nb de garçons)

-$Y \leq X$ (ex: 2 tris successifs)

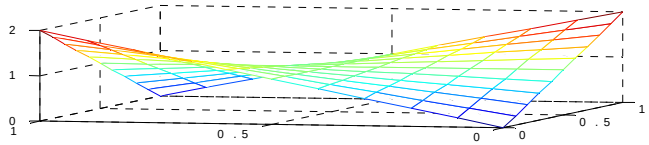
Fonction de (densité de) probabilité

- Pour un couple discret la fonction de probabilité conjointe est donnée par:
  - $p(x,y)=\begin{cases} P(X=x \cap Y=y) & \text{si } (x,y)\in R_{XY} \\ 0 & \text{si } (x,y)\notin R_{XY} \end{cases}$
- Pour un couple continu la fonction de densité de probabilité conjointe f(x,y) est donnée par:
  - $f(x,y)=\partial^2 F(x,y)/\ \partial x \partial y$



Fonctions de (densités de) probabilité marginales

- On peut obtenir les fonctions de (densités de) probabilité de chacune des variables X et Y, connaissant la loi (densité) conjointe du couple (X,Y) (l'inverse n'est pas vrai)
- Cas discret: fonctions marginales
  - $p_X(x)=\Sigma_j\ p(x,y_j)$ $\quad p_Y(y)=\Sigma_i\ p(x_i,y)$
- Cas continu: fonctions de densités marginales
  - $f_X(x)=\int_{-\infty:+\infty} f(x,y)dy \quad f_Y(y)=\int_{-\infty:+\infty} f(x,y)dx$
- Ex:



Exercice: f(x,y) =2(x+y-2xy) sur [0,1]x[0,1] et 0 ailleurs. Déterminer les marginales f(x) et f(y)

Covariance - corrélation

- Si m_X et m_Y sont les espérances respectives de X et Y, on définit la covariance de X et Y σ_{XY}=cov(X,Y) par: $\sigma_{XY} = E((X-m_X)(Y-m_Y))$
- Cas discret
  - $\sigma_{XY}=\Sigma_i \Sigma_j (x_i - m_X)(y_j - m_Y)p(x_i,y_j)$
- Cas continu
  - $\sigma_{XY}= \int_{-\infty:+\infty} \int_{-\infty:+\infty} (x - m_X)(y - m_Y)f(x,y)dx\,dy$
- Le coefficient de corrélation ρ_{XY} est défini par:
  - $\rho_{XY}= \sigma_{XY} /(\sigma_X \sigma_Y) \quad \Rightarrow \quad -1 \leq \rho_{XY} \leq 1$

Exercice: Matrice de covariance

Exercice: $Y=aX+b \Rightarrow \rho_{XY}=\pm 1$

Distributions Conditionnelles

- C'est l'un des intérêts majeurs de considérer des couples de v.a. , on peut obtenir des informations sur l'une lorsque des informations sont disponibles sur l'autre en exploitant le lien entre les deux.
- Cas discret
  - $p_X(x/y)=p(x,y)/p_Y(y)$ $p_Y(y/x)=p(x,y)/p_X(x)$
- Cas continu
  - $f_X(x/y)=f(x,y)/f_Y(y)$ $f_Y(y/x)=f(x,y)/f_X(x)$

- Théorème des probabilités composées
  - Cas discret: $p(x,y)=p_X(x/y)p_Y(y)=p_X(x)p_Y(y/x)$
  - Cas continu: $f(x,y)=f_X(x/y)f_Y(y)=f_X(x)f_Y(y/x)$
- Théorème des probabilités totales
  - Cas continu
    - $f_X(x)=\int_{-\infty;+\infty} f_X(x/y)f_Y(y)\,dy$
    - $f_Y(y)=\int_{-\infty;+\infty} f_X(x)f_Y(y/x)\,dx$
- Théorème de Bayes
  - Cas discret
    - $p_X(x/y)=\frac{p_X(x)p_Y(y/x)}{\sum_i p_X(x_i)p_Y(y/x_i)}$

Exercice: écrire le théorème des probabilités totales dans le cas discret et le théorème de Bayes dans le cas continu.

Exercice: canal binaire symétrique : calcul de la distribution de probabilité de la sortie Y

Espérance et variance conditionnelle

- Espérance $m_{X/y}=E(X/y)$
  - Cas discret: $m_{X/y}= \Sigma_i x_i\, p_X(x_i/y)$
  - Cas continu: $m_{X/y}= \int_{-\infty:+\infty} x f_X(x/y)dx$
- Variance $\sigma^2_{X/y}=Var(X/y)$
  - Cas discret: $\sigma^2_{X/y}= \Sigma_i (x_i - m_{X/y})^2\, p_X(x_i/y)$
  - Cas continu: $\sigma^2_{X/y}= \int_{-\infty:+\infty} (x-m_{X/y})^2\, f_X(x/y)dx$
- Lien entre espérance et espérance conditionnelle
  - $E(X)=\begin{cases} \Sigma_j E(X/y_j)\, p_Y(y_j) & \text{pour Y discrète} \\ \int_{-\infty:+\infty} E(X/y)\, f_Y(y)dy & \text{pour Y continue} \end{cases}$

Exercice: montrer que pour deux v.a. X et Y indépendantes:

$E(X/y)=E(X)$ et $E(Y/x)=E(Y)$

Indépendance

- Deux v.a. X et Y sont indépendantes (on note $X \perp Y$) si la connaissance de l'une ne modifie pas la distribution de probabilité de l'autre
- Cas discret:
  - $X \perp Y \Leftrightarrow \begin{cases} p_X(x/y)=p_X(x) \ \forall x \\ p_Y(y/x)=p_Y(y) \ \forall y \end{cases}$
    $\Leftrightarrow \quad p(x,y)= p_X(x)\, p_Y(y) \ \forall x,y$
- Cas continu:
  - $X \perp Y \Leftrightarrow \begin{cases} f_X(x/y)=f_X(x) \ \forall x \\ f_Y(y/x)=f_Y(y) \ \forall y \end{cases}$
    $\Leftrightarrow \quad f(x,y)= f_X(x)\, f_Y(y) \ \forall x,y$
- Propriété: Pour deux v.a. X et Y indépendantes: $\sigma_{XY}=0$

Exercice: X et Y sont les durées de vie des deux machines qui suivent des lois exponentielles X et Y ~ Exp(0.01) en heures . Quelle est la probabilité de panne (X+Y< 10) avec ces 2 machines ?

Exercice: montrer la propriété: Pour deux v.a. X et Y indépendantes: $\sigma_{XY}=0$

Somme de deux variables aléatoires indépendantes

- On considère Z=X+Y avec X⊥Y
- La variance de la somme est la somme des variances
  - $\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2$
- La distribution de probabilité de la somme est le produit de convolution des distributions
  - Cas discret
    - $p_Z(z) = \Sigma_i\, p_X(x_i)p_Y(z-x_i) = \Sigma_j\, p_Y(y_j)p_X(z-y_j)$
  - Cas continu
    - $f_Z(z) = \int_{-\infty:+\infty} f_X(x)f_Y(z-x)dx = \int_{-\infty:+\infty} f_Y(y)f_X(z-y)dy$

Exercice: démonstration des propriétés

Exercice: X et Y uniformes sur [-1,1] : loi de Z=X+Y?

Exercice: X v.a. à valeurs dans {-1,1} Y v.a. continue normale Y ~$\mathcal{N}(0,\sigma^2)$: loi de Z=X+Y?

Génération de nombres aléatoires

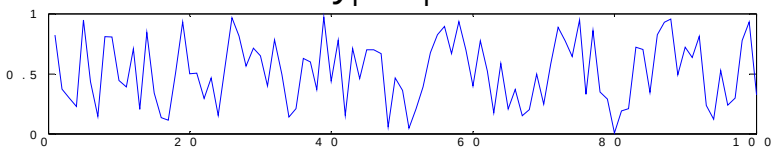
Intérêt

- Il est toujours possible de générer une séquence aléatoire à partir de réalisations d'une expérience aléatoire mais c'est laborieux (lancer un dé 10000 fois, compter le nombre de voitures passant à un péage par minute pendant 1h déterminée pendant 10 jours…).
- On recherche une méthode arithmétique qui puisse faire cette génération automatiquement. La séquence sera alors pseudo- aléatoire (elle aura les mêmes caractéristiques qu'une séquence vraiment aléatoire tout en ayant été générée par des opérations déterministes)

Générateur aléatoire uniforme

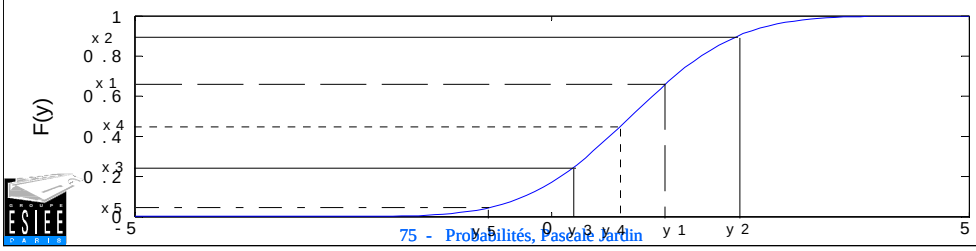
- Algorithme classique : formule de Lehmer
  - $x_{i+1} = a\, x_i \mod b$ pour i=0,1,…,n-1

  avec $b \in \mathbb{N}^+$,$a \in \{1,2,\ldots,b-1\}$ et $x_0 \in \{1,2,\ldots,b-1\}$
  - Sous certaines conditions sur a et b les valeurs x_i sont des réalisations d'une v.a. uniforme sur l'ensemble {0,1,2,…,b-1}
  - De bons résultats sont obtenus par exemple avec $a=2^7 - 1$ et $b=2^{31}-1$
  - En général on préfère générer des nombres uniformément répartis sur [0,1[en effectuant la transformation: $y_i = x_i/b$



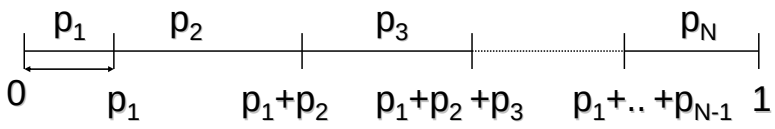
Génération suivant loi non uniforme par inversion de la fonction de répartition

- On génère des réalisations x_i d'une v.a. X de loi uniforme: X ~ Un(0,1)
- On considère une v.a. Y de fonction de répartition $F_Y(y)$.
- On applique l'inversion : $y_i = F_Y^{-1}(x_i)$.
- Les y_i sont des réalisations de la v.a Y



Génération suivant une loi discrète de distribution $P(X= x_i)=p_i$ pour i=1,…,N

- Division de l'intervalle [0,1[en sous intervalles de longueur p_i.



- Génération de réalisations u_n d'une v.a. U de loi uniforme: U ~ Un(0,1).
- lorsque $u_n \in [\Sigma_{k=0:i-1} p_k, \Sigma_{k=0:i} p_k[$ on génère x_i

Statistique descriptive – Estimation
Introduction

Objectifs

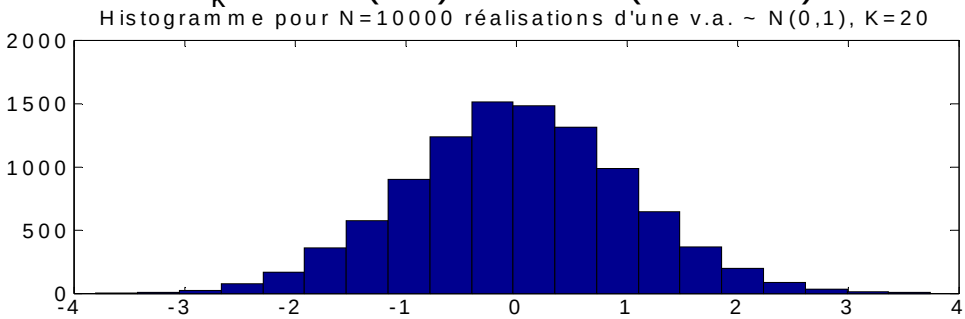
- A partir d'un ensemble de N réalisations X_i (issues de mesures ou de simulation) d'une ou plusieurs v.a. (vecteur aléatoire) on cherche à décrire cette v.a. (ou ce vecteur)
  - En Statistique cet ensemble est appelé "échantillon représentatif" et la variable ou le vecteur aléatoire caractérise une "population"
  - Exemples:
    - taille, poids, âge des étudiants de l'ESIEE
    - PIB, taux de natalité, taux de mortalité, espérance de vie, revenu moyen des pays d'un continent
    - Activité des noeuds d'un réseau de communication

Statistiques courantes pour X=(X_1,…,X_N)

- Min(X)=$X_{(1)}$, Max(X)=$X_{(N)}$, étendue= $X_{(N)}$ - $X_{(1)}$
- Moyenne: $$\hat{m} = \frac{1}{N}\sum_{i=1}^{N} X_i$$
- Variance: $$\hat{\sigma}^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(X_i - \hat{m}\right)^2$$

Histogramme

- L'histogramme est la représentation graphique des valeurs H_k en ordonnée (et x_k en abscisse) telles que :
  - H_k = card{$x_k < X_i < x_{k+1}$} pour k=1 à K
  - avec x_k = Min+(k-1)Δx et Δx=(Max-Min)/K



  - l'histogramme permet l'estimation des probabilités

Estimation des probabilités

- $\hat{P}(a < X \leq b) = \dfrac{card\{a < X_i \leq b\}}{N}$
- $\hat{F}(x) = \dfrac{card\{X_i \leq x\}}{N}$
- $\hat{p}(x) = \dfrac{card\{X_i = x\}}{N}$
- $\hat{f}(x) = \dfrac{card\left\{x - \dfrac{\Delta x}{2} < X_i \leq x + \dfrac{\Delta x}{2}\right\}}{N\Delta x}$

Estimation des paramètres de différentes lois à partir de l'estimée de la moyenne

- X~Bi(n,p) : $\hat{p} = \frac{\hat{m}}{n}$
- X~Ge(p) : $\hat{p} = \frac{1}{\hat{m}}$
- X~Po(μ) : $\hat{\mu} = \hat{m}$
- X~Exp(λ) : $\hat{\lambda} = \frac{1}{\hat{m}}$

Qualités d'un estimateur

- Biais : $b\left(\hat{\theta}\right) = E\left(\hat{\theta}\right) - \theta$
  - On cherchera à avoir des estimateurs non biaisés
- EQM: $EQM\left(\hat{\theta}\right) = E\left(\left(\hat{\theta} - \theta\right)^2\right)$
  - L'estimateur est efficient si son EQM est finie et minimale
- Consistance: un estimateur est consistant si: $\lim_{N\to\infty} P\left(\left|\hat{\theta} - \theta\right| < \varepsilon\right) = 1$

Exercice: Biais pour estimateur de la moyenne et de la variance