Gilles HUNAULT 2000

Probabilités et Statistiques

Université d'Angers

[...] *Et au troisième coup, il n'y a plus que 12,5 % de chances que nous continuions de gagner ou de perdre. Cela est d'ailleurs purement théorique, car à partir de là, je vous prie de remarquer que nous ne sommes plus du tout dans l'ordre du réel, mais dans celui de la signification symbolique que nous avons définie* [...]. *Du point de vue du réel, il y a toujours à chaque coup autant de chances que vous gagniez ou que vous perdiez. La notion même de probabilités et de chances suppose l'introduction d'un symbole dans le réel. C'est à un symbole que vous vous adressez, et vos chances ne portent que sur un symbole. Dans le réel, à chaque coup, vous avez tout autant de chances de gagner ou de perdre qu'au tour précédent. Il n'y a aucune raison que, par pur hasard, vous ne gagniez pas dix fois de suite. Cela ne commence à prendre un sens que quand vous écrivez un signe, et tant que vous n'êtes pas là pour l'écrire, il n'y a aucune espèce de gain.*

Jacques LACAN, le séminaire, livre II:
le moi dans la théorie de Freud et dans
la technique de la psychanalyse

Table des matières

Chapitre 1.

Introduction

1.1 Présentation générale

Les probabilités et les statistiques sont deux domaines complémentaires des mathématiques. Les probabilités fournissent le cadre théorique, indépendamment de toute "réalité". Les statistiques partent de la réalité pour donner une vue globale, tentent de synthétiser ou de modéliser en se servant des modèles théoriques issus des probabilités.

On trouvera ici à la fois une introduction aux calculs mathématiques qui servent de base aux probabilités et aux statistiques mais aussi des exemples concrets, une discussion de l'interaction entre calculs et interprétation. Nous fournirons aussi des exemples de programmation de façon à rendre le lecteur, la lectrice à la fois autonome quant à la lecture d'ouvrages sur les deux disciplines qui fondent ce cours mais aussi capable de s'adapter aux logiciels courants ou apte à (re)programmer ces calculs et formules.

Nous attendons des élèves de ce cours qu'ils sachent calculer (car on ne "fait" pas des mathématiques) à la fois à la main et avec un ordinateur. Nous espèrons qu'ils sauront maitriser l'ensemble des concepts et qu'ils pourront acquérir au long de ces pages suffisamment de recul pour "voir des choses" au-delà des pages de formules et de chiffres, par-delà les hypothèses, les modèles. Enfin, et surtout, nous voudrions qu'ils arrivent à exposer leurs vues, leur résultats chiffrés, leurs conclusions avec précision, concision et culture en un tout lisible, accessible dans cet effort difficile qu'est la rédaction et l'exposé.

1.2 Plan de l'ouvrage

Les pages que vous êtes en train de lire constituent le chapitre 1, c'est l'*Introduction* à l'ouvrage. Le chapitre 2 viendra rappeler les bases du calcul des *Probabilités*, avec de nombreux exemples concrets. On y introduira les *Variables Aléatoires*, leurs grandeurs caractéristiques et quelques *Lois* classiques. On passera en revue dans le chapitre 3 les *Statistiques* classiques et leurs caractéristiques avec un petit détour par la notion de test. Ensuite, des statistiques un peu moins classiques seront traitées dans le chapitre 4 avec des *variables autres* que qualitatives ou quantitatives. Enfin, la théorie reprendra sa place avec les *variables vectorielles* (en particulier *gaussiennes*) et la *convergence* dans le chapitre 5 alors que le chapitre 6 essaiera de conclure via l'*informatisation* des traitements, l'utilisation de *logiciels* et l'importance de la *rédaction*.

Sachant que tous et toutes n'ont pas les mêmes bases théoriques et mathématiques, nous avons souvent regroupé les démonstrations en fin de chapitre de façon à faciliter la lecture des concepts et à ne pas ralentir la progression du cours. Attention toutefois aux démonstrations : de façon à privilégier la réflexion, les démonstrations sont simplifiées. On ne trouve aucune justification des enchainements. C'est tout à fait volontaire, contraignant et sans doute plus pédagogique qu'il n'y parait. Le lecteur peu intéressé par les mathématiques nous en saura gré, l'expert se fera un plaisir de justifier ici l'utilisation de la commutativité de l'addition dans $\mathbb{R}$, là l'existence des intégrales généralisées... Dans le même esprit, des annexes sur les structures algébriques et le dénombrement complètent les 6 chapitres et avec aussi quelques incontournables tables numériques.

De nombreux exercices sont proposés après chaque partie de cours importante. Certains sont des applications pas toujours simples du cours, d'autres viennent fournir des compléments, des éclairages différents sur les notions mises en place. Nous ne saurions trop conseiller de lire soigneusement les énoncés de ces exercices, à défaut de les résoudre...

Chapitre 2.

Espaces Probabilisés et Variables aléatoires

2.1 Probabilités

2.1.1 Définitions et Propriétés élémentaires

Soit E un ensemble dont les éléments sont nommés *évènements élémentaires* ; E lui-même est nommé *espace des épreuves*. On le note classiquement Ω. On appelle *évènement* de E tout élément de $\mathcal{T} = \mathcal{P}(E)$. Le complémentaire de l'évènement F par rapport à E est noté $\overline{F}$ et nommé évènement *contraire* de F.[1] Une probabilité p sur E est une fonction de $\mathcal{T}$ dans $[0,1]$ qui vérifie les axiomes suivants nommés *axiomes de Kolmogorov* :

[P1] $p(E) = 1$
[P2] $p(A \cup B) = p(A) + p(B)$ si $A \cap B = \emptyset$

Tout évènement X tel que $p(X) = 1$ est appelé *évènement certain*. Tout évènement Y tel que $p(Y) = 0$ est appelé *évènement impossible*. En particulier, E est certain et $\emptyset$ est impossible. $(E,\mathcal{T})$ est nommé espace probabilisable et $(E,\mathcal{T},p)$ est nommé espace probabilisé.

1. On note aussi $\mathcal{C}(F)$ ou comp(F) ce complémentaire suivant le contexte (mathématique, informatique...). De même, on précise aussi par rapport à quoi on prend le complémentaire s'il peut y avoir ambiguité, comme avec la notation traditionnelle $\complement_E F$.

De **[P1]** et **[P2]** on déduit que

[P3] $p(\emptyset) = 0$

[P4] $p(\overline{F}) = 1 - p(F)$

[P5] $p(A \cup B) = p(A) + p(B) - p(A \cap B)$ (formule de *Poincaré*)

[P6] p est croissante de $(\mathcal{T}, \subset)$ dans $([0,1], \leq)$

$\mathcal{T}$ est nommé l'univers des possibles (ou plus simplement *univers*). Au lieu d'évènement, on emploie aussi le terme d'éventualité. Deux événements A et B tels que $A \cap B = \emptyset$ sont dits *incompatibles* (ou encore disjoints). Deux événements A et B tels que $p(A \cap B) = p(A).p(B)$ sont dits *indépendants.* Ces deux notions ne sont pas liées. Un *système complet* d'évènements de E est une partition de E.

2.1.2 Probabilités conditionnelles, décompositions

On appelle probabilité conditionnelle de B quand A la quantité

$$p(B|A) = \frac{p(A \cap B)}{p(A)}.$$

Cette notion n'a bien sûr de sens que si $p(A)$ est non nul. On prendra garde au fait que $p(B|A)$ n'est *qu'une notation* et qu'elle n'a aucun sens "classique" puisque $B|A$ ne désigne pas un évènement au même titre que $A \cap B$, $A \cup B$...[2]

Grâce aux propriétés ci-dessus, pour deux évènements A et B, on a :

la *formule des probabilités totales*

$$p(A) = p(B).p(A|B) + p(\overline{B}).p(A|\overline{B})$$

et la *formule de Bayes*

$$p(B|A) = \frac{p(B).p(A|B)}{p(B).p(A|B) + p(\overline{B}).p(A|\overline{B})}$$

2. $A \subset B$ non plus n'est pas un ensemble. Mais il n'y a ici aucune confusion puisqu'on n'écrit jamais $p(A \subset B)$.

Ces formules se généralisent, pour un système complet $\{\omega_i\}$ de E en

[P7] $p(A) = \Sigma_i\ p(\omega_i).p(A|\omega_i)$
[P8] $p(\omega_j|A) = p(\omega_j).p(A|\omega_j)/\Sigma_i\ p(\omega_i).p(A|\omega_i)$

Pour A fixé, notons maintenant $p_{|A}$ l'application $X \mapsto p(X|A)$. Alors :

[P9] $p_{|A}$ est une probabilité.

2.1.3 Généralisation de la notion de probabilité

Une probabilité est en fait un cas particulier de mesure sur une tribu. Une *tribu* (ou σ-algèbre ou algèbre σ-additive) sur un ensemble Ω est une famille de sous-ensembles de Ω qui contient $\emptyset$ et Ω, stable par passage au complémentaire et par union finie ou dénombrable. Si $\mathcal{T}$ est une tribu sur Ω, le couple $(\Omega,\mathcal{T})$ est appelé *espace mesurable.* Une *mesure* μ sur un espace mesurable est une fonction σ*-additive* de Ω dans $\overline{R}_+$ (où $\overline{R}_+$ désigne $R_+ \cup \{\infty\}$) c'est à dire une fonction qui vérifie : $\mu(\bigcup_i\ A_i) = \Sigma_i\ A_i$ où les A_i sont une famille finie ou dénombrable d'évènements deux à deux incompatibles. Le triplet $(\Omega,\mathcal{T},\mu)$ est alors appelé *espace mesuré.*

2.1.4 Calculs effectifs de Probabilités

Les définitions et propriétés précédentes montrent comment calculer la probabilité d'un évènement à partir d'une décomposition ensembliste en évènements élémentaires dont on connait les probabilités. Celles-ci sont soit déterminées par des calculs de dénombrements soit en utilisant une hypothèse raisonnable, pratique parce que simple mais pas forcément fondée : la répartition équitable ou équirépartition. Ainsi obtenir un nombre pair en lançant un dé se ramène à l'union disjointe de l'obtention d'un deux, d'un quatre... Sauf hypothèse contraire, tous les chiffres sont équiprobables et sortir un chiffre pair a donc une probabilité de 0.5.

Les probabilités ainsi construites sont toutes *théoriques* en ce sens qu'elles ne reflètent pas la *réalité*. Un dé n'a jamais ses faces parfaitement équiprobables, le "**hasard**" intervient plus souvent qu'un "ordre" uniformément réparti...

2.2 Probabilités Intuitives et Analyse Combinatoire

2.2.1 Formules des probabilités naives

Pour calculer pratiquement des probabilités sur des situations concrètes finies, on utilise la fonction probabilité suivante, nommé probabilité-cardinal. Soit E un ensemble fini. Si A est un sous-ensemble de E, on nomme probabilité-cardinal ou fonction poids-pondéré de A, noté $p_c(A)$ le rapport du cardinal de A à celui de E, soit en formule $p_c(A) = card(A)/card(E)$. Il est facile de démontrer que p_c est une probabilité sur $\mathcal{P}(E)$ et la formule $p_c(A \cup B)$ se comprend bien car le nombre d'éléments dans $A \cup B$ est la somme du nombre d'éléments dans A et du nombre d'éléments dans B, somme à laquelle on retire les éléments comptés deux fois (c'est à dire ceux dans $A \cap B$).

Le calcul de la probabilité de A résulte du comptage ou du dénombrement des éléments dans A. Une interprétation classique nomme alors "nombre de cas favorables" le nombre d'éléments de A et "nombre de cas total" le nombre d'éléments de E d'où la formule des *probabilités naives* :

$$p(A) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas total}}$$

Le calcul des probabilités finies se ramène alors à des dénombrements et autres calculs combinatoires que nous rappelons ici avec les interprétations en termes de "fonctions à propriétés".

Commençons par un exemple, celui du tirage (ou de l'extraction, de la production...) de 2 éléments parmi 3. Nommons a, b et c les éléments de base. On peut prendre les éléments avec ou sans répétitions, avec ou sans ordre, ce qui fait 4 types de comptages à savoir

$\{a,b\}$ $\{a,c\}$ $\{b,c\}$	(a,b) (a,c) (b,a) (b,c) (c,a) (c,b)
(a,a) (a,b) (a,c) (b,b) (b,c) (c,c)	(a,a) (a,b) (a,c) (b,a) (b,b) (b,c) (c,a) (c,b) (c,c)

Le choix de deux éléments avec ordre, avec répétition éventuelle permet d'utiliser le vocabulaire classique de "couple" des mathématiciens. Pour p éléments plutot que 2, on parlera de p-uplets. Les modèles de répartition peuvent donc être nommées p-uplets ou structures ou dispositions et nous les repérerons comme suit

sans répétition, sans ordre	sans répétition avec ordre
avec répétition, sans ordre	avec répétition avec ordre

mais on préfére un autre vocabulaire, plus concis mais sans doute moins explicite, puisqu'on parle de

combinaisons	arrangements
dérangements	p-uplets (quelconques)

Soit dans le cas général un ensemble E avec p éléments et F un ensemble à n éléments. Nous conviendrons de noter 1,2,3..p les éléments de E et 1,2,3..n ceux de F. Une application (ou fonction) f de E dans F est complétement déterminée par la donnée de $f(1)$, $f(2)$...$f(p)$ et donc un p-uplet d'éléments dans F correspond à la donnée d'une application quelconque de E dans F. Notons temporairement V_n^p le nombre de ces applications quelconques.

Si l'on se restreint aux applications injectives, il ne peut plus y avoir le même élément plusieurs fois de suite (car la définition d'une injection est "à deux éléments distincts correspondent deux images distinctes") et on peut donc interpréter un arrangement comme la donnée d'une injection. Nous noterons A_n^p le nombre d'injections. Une combinaison de p éléments parmi n correspond à l'ensemble des éléments fournis par une injection et avec un peu d'imagination on peut l'interpréter comme la donnée d'une fonction strictement croissante. Enfin, un dérangement correspond à la donnée d'une fonction croissante au sens large. Il est d'usage de noter C_n^p le nombre de combinaisons c'est à dire le nombre de sous-ensembles à p éléments dans un ensemble à n éléments et K_n^p le nombre de dérangements correspondants, soit le tableau des notations et des propriétés des fonctions

C_n^p	A_n^p
K_n^p	V_n^p

croissantes (strictes)	injectives
croissantes (larges)	quelconques

Lorsque $n = p$, toute injection est une surjection et on nomme traditionnellement *permutation* une bijection d'un ensemble fini sur lui-même. Soit $B(n)$ le nombre de bijections d'un ensemble de n éléments. Pour une bijection donnée, le premier élément peut prendre l'une quelconque des n valeurs, le deuxième élément n'a plus le choix qu'entre les $n-1$ valeurs restantes, etc. ce qui fait que $B(n) = n(n-1)(n-2)...1$ c'est à dire $n!$.

Le nombre de p-uplets est facile à calculer puisqu'il y a n choix possibles pour le premier élément, n choix aussi pour le deuxième, donc $V_n^p = n^p$. C'est ce qui justifie la notation F^E pour désigner l'ensemble $\mathcal{F}(E,F)$ des fonctions de E dans F. Le calcul du nombre d'arrangements est assez simple aussi : on a n possibilités pour le premier élément, $n-1$ pour le second etc. mais on s'arrête au bout de p soit la valeur $A_n^p = n(n-1)(n-2)...(n-p+1) = n!/(n-p)!$. Enfin, comme il y a $p!$ façons de permuter les p élements de l'arrangement, $C_n^p = A_n^p/p!$. Nous laissons au lecteur de démontrer la dernière formule pour K_n^p dans le tableau

$\dfrac{n!}{p!(n-p)!}$	$\dfrac{n!}{(n-p)!}$
$\dfrac{(n+p-1)!}{p!(n-1)!}$	n^p

Pour conclure, donnons une interprétation en termes de boules : le tirage avec remise autorise les répétitions, le tirage par paquets (p fois 1 boule plutot que 1 fois p boules) permet de jouer sur l'ordre donc les modèles théoriques de tirages correspondants sont

1 fois p boules sans remise	p fois 1 boule sans remise
	p fois 1 boule avec remise

La modélisation de nombreux phénomènes, de coefficients dans des formules du genre le binome de *Newton* fournit de nombreuses formules mettant en jeu les C_n^k. En effet, à partir de

$$(a+b)^n \;=\; \sum_{i=0}^{n} C_n^i a^i b^{n-i}$$

que l'on démontre par récurrence, on voit que le coefficient de a^i est C_n^i d'où $C_n^0 = 1$, $C_n^1 = n$, $C_n^2 = n(n-1)/2$. Citons au passage quelques formules usuelles plus ou moins simples à démontrer :

[F1] $C_n^p = C_n^{n-p}$ **[F2]** $C_{n-1}^{p-1} + C_{n-1}^p = C_n^p$

[F3] $\sum_{i=0}^{n} C_n^i = 2^n$ **[F4]** $\sum_{i=0}^{n} (C_n^i)^2 = C_{2n}^n$

[F5] $\sum_{i=0}^{n} (-1)^i C_n^i = 0$ **[F6]** $\sum_{i=1}^{n} i C_n^i = n2^{n-1}$

[F7] $\sum_{i=2}^{n} i(i-1) C_n^i = n(n-1)2^{n-2}$ **[F8]** $\sum_{i=2}^{n} i^2 C_n^i = n(n+1)2^{n-2}$

Terminons enfin par quelques (jolies) formules plus compliquées :

[G1] $\sum_{i=p}^{n} C_n^i = C_{n+1}^{p+1}$

[G2] $\sum_{i=0}^{n+1} (-1)^i C_{n+1}^i P(x+i) = 0$ où P est un polynôme de degré $\leq n$

[G3] $1 + \sum_{k=0}^{p-1} C_p^k \, s_k(n) = (n+1)^p$

si $s_k(n) = \sum_{j=1}^{n} j^k$ est la somme des puissances k-ièmes des n premiers entiers, avec donc, comme chacun sait :

$$\begin{array}{llllll} s_1(n) & = 1 & + 2 & + 3... & + n & = n(n+1)/2, \\ s_2(n) & = 1^2 & + 2^2 & + 3^2... & + n^2 & = n(n+1)(2n+1)/6... \end{array}$$

2.2.2 Quelques calculs de Probabilités et Dénombrements

Si n tomes d'une encyclopédie sont rangés sur une étagère, la probabilité que les tomes 1 et 2 soient rangés cote à cote dans cet ordre est $(n-1).(n-2)!/n!$ soit $1/n$. De même, la probabilité que les tomes 1,2...t soient rangés cote à cote dans cet ordre est $(n-t+1)!/n!$.

Dans un polygone convexe déterminé par n points, il y a $n(n-1)/2$ droites dont n cotés. La probabilité qu'une droite soit une diagonale dans un tel polygone est donc $(n-3)/(n-1)$.

Prenons $2r$ chaussures parmi n paires. La probabilité de n'avoir pris aucune paire correcte est $2^{2r}C_n^{2r}/C_{2n}^{2r}$.

2 personnes écrivent au hasard un nombre de deux chiffres exactement (entre 10 et 99). Cette expérience est répétée n fois. La probabilité d'avoir écrit au moins une fois le même nombre est $1-(89/90)^n$.

r personnes sont dans un ascenseur à n étages. On suppose que la probabilité qu'une personne s'arrête à un étage donné est $1/n$ et que les décisions des r personnes sont indépendantes. La probabilité que les r personnes s'arrêtent à des étages différents est A_n^r/n^r et la probabilité que deux personnes exactement s'arrêtent au même étage et que les $r-2$ autres personnes s'arrêtent chacune à un étage différent (qu'on supposera différent de l'étage de ces deux personnes) est $C_r^2 A_n^{r-1}/n^r$.

Admettons qu'un "dé" A comporte 4 faces rouges et 2 faces blanches et qu'un "dé" B comporte 2 faces rouges et 4 faces blanches. On lance une pièce de monnaie et si on obtient pile, on utilise le dé A ; si c'est face, on utilise le dé B. La probabilité d'avoir utilisé A sachant qu'on a obtenu n fois rouge en lançant n fois le dé est $2^n/(2^n+1)$.

Le nombre de surjections $S(n,p)$ de E avec n éléments dans F avec p éléments est

$$S(n,p) = \sum i = 1^{p-1}(-1)^{i-1}C_p^{p-i}(p-i)^n$$

Le nombre de solutions entières de l'équation $x_1 + x_2 ... + x_p = n$ est C_{n+p-1}^n.

2.3 Variables aléatoires

2.3.1 Définitions

Maitenant que nous savons reconnaitre, décomposer et plus généralement manipuler des évènements, nous allons introduire des fonctions sur ces évènements. Ainsi, nous savons déterminer les probabilités associés aux différentes faces d'un dé, que ce soit dans des tirages avec ou sans remise, mais que faire avec ces tirages ? Par exemple, pour deux dés, on peut s'intéresser à la somme des chiffres obtenus en un tirage, à leur produit, ou au nombre de lancers qu'il faut effectuer pour obtenir le 6 sur chaque dé en même temps...

C'est ici qu'interviennent les variables aléatoires. Formellement, une variable aléatoire nommée X sur un ensemble E muni d'une probabilité p est une fonction de $\mathcal{T}_E$ dans une partie de $F = \mathbb{R}$ telle que $\forall u \in F, X^{-1}(B) \in \mathcal{T}_E$ où B est un *borélien*[3] de F. Si $Im(X)$ est fini ou dénombrable, le vocabulaire classique[4] l'appelle variable *discrète* ; sinon, elle est dite *continue*. On peut aussi prendre comme ensemble d'arrivée non pas $\mathbb{R}$ mais $\mathbb{R}^n$; la variable aléatoire est alors dite *vectorielle*. Un premier problème posé est alors celui des notations et un second celui de la probabilité sous-jacente à une "expérience aléatoire empirique". Pour le premier problème, on conviendra de noter $X \in B$ l'ensemble des ω tel que $X(\omega) \in B$ et on s'autorisera même à écrire $X = k$ si B se réduit à une seule valeur, alors que l'écriture correcte serait $X^{-1}(\{k\})$. On trouvera aussi un sens à une expression comme $X \leq k$, $a < X < b$... Pour le second problème, ce sera souvent plus délicat car il y de nombreux sous-entendus : si on lance deux dés et qu'on s'intéresse à la somme des valeurs obtenues, la probabilité sous-jacente n'est pas très difficile à trouver, même si elle requiert un peu de calcul.

Il est facile à partir de la notion de variable aléatoire de comprendre une différence fondamentale entre probabilités et statistiques. En probabilités, on fera l'hypothèse que chaque face du dé est *équiprobable*, on s'intéressera à toutes la valeurs possibles et on déterminera théoriquement ce qu'on doit obtenir.

---

3. Nous utilisons ce terme sans l'avoir défini, ce qui ne gêne pas si l'on prend comme équivalent la périphrase "réunion d'intervalles". L'intérêt de donner le "bon" mot est de faciliter l'approfondissement vers la théorie de la mesure...

4. On peut être étonné de ce vocabulaire puisque'une variable aléatoire est... une fonction ! De même pour la qualification variable aléatoire continue puisque ce n'est pas la fonction qui est continue mais l'ensemble d'arrivée qui est non dénombrable.

En statistiques, on réalise une expérience et on compte. C'est en comparant les valeurs observées et les valeurs théoriques qu'on peut se faire une idée exacte du phénomène, (se) convaincre que les hypothèses sont exactes, ou remettre en question le modèle (ici l'équiprobabilité).

Soit donc X une variable aléatoire discrète sur (E,p) à valeurs dans I et définissons l'application p_X de I dans $[0,1]$ par $p_X(k) = p(X^{-1}(\{k\}))$ soit encore, avec la notation précédente : $p_X(k) = p(X = k)$. Alors :

[V1] p_X est une probabilité appelée probabilité image de p par X.

Passons maintenant à une variable aléatoire continue Y à valeurs dans $\mathbb{R}$, et notons p_Y l'application qui à un intervalle $[a,b]$ de $\mathbb{R}$ associe la quantité $p(a \leq Y \leq b) = p(Y^{-1}([a,b]))$. Alors

[V2] p_Y est une probabilité.

Là encore, p_Y est appelée probabilité image de p par Y.

2.3.2 Valeurs caractéristiques

Essayons maintenant de caractériser nos variables aléatoires. Comme elle peuvent avoir de nombreuses valeurs, il est intéressant de les résumer. Le premier indicateur numérique global est la *moyenne* qui donne la *tendance centrale*. Si X est une variable aléatoire discrète, notons x_i les différentes valeurs possibles de X et p_i les probabilités associées. Alors la moyenne de X est $m = \Sigma\ x_i.p_i$.

Pour une variable aléatoire continue Y, comme il n'est pas possible de d'utiliser $p(x)$, [5], on a recours à la densité de Y qui est une fonction f_Y telle que $\int_{\mathbb{R}} f = 1$ avec $p(Y \leq y) = \int_{-\infty}^{y} f$. La moyenne de Y est alors $m = \int Id.f_y$.

Il existe diverses notations pour désigner la moyenne. Lorsqu'on utilise plusieurs variables aléatoires, on peut noter m_X, $m(X)$, $moy(x)$, $average(X)$, $av(X)$ ou encore $\overline{X}$ pour préciser de quelle variable il s'agit.

Le deuxième indicateur numérique global est la variance. C'est un indicateur de *dispersion absolue* moyenne autour de m. Si on note $mt(X,n) = \Sigma\ x_i^n.p_i$ le moment d'ordre n pour une variable aléatoire discrète X, alors on peut remarquer que $m(X) = mt(X,1)$.

5. Pourquoi ?

La variance peut alors être définie par $V = mt(X{,}2) - m(X)^2$. En d'autre termes, "la variance de X est la différence entre la moyenne du carré des valeurs et la carré de la moyenne des valeurs". On peut aussi introduire les moments centrés : $mt_c(X{,}n) = mt(X - m(X){,}n)$. On démontre alors que

[V3] $V(X) = mt_c(X{,}2)$

soit encore : "la variance de X est son moment centré d'ordre 2". Pour une raison peut-être pas très claire pour l'instant, on définit aussi l'écart-type $\sigma(X)$ de X comme la racine carrée de la variance : $\sigma(X) = \sqrt{V(X)}$. Comme pour la moyenne, de nombreuses notations existent pour $V(X)$ et $\sigma(X)$, notamment $var(X)$, $\overline{\overline{\overline{X}}}$, $etyp(X)$, $ect(X)$, $std(X)$... Le rapport σ/m est nommé *coefficient de variation* et fournit un indicateur de *dispersion relative.*

Pour une variable aléatoire continue Y, le moment d'ordre n est défini par :

$$mt(Y{,}n) = \int_R Id^n.f_y$$

et la variance par

$$V = mt(Y{,}2) - m(Y)^2$$

Il existe d'autres indicateurs globaux (*kurtosis, skewness*) mais ils ne nous intéressent pas pour l'instant. Nous en parlerons au prochain chapitre quand nous traiterons les séries statistiques. De même, des indicateurs locaux (comme minimum, maximum...) ne sont pas a priori intéressants en probabilitiés. On verra par contre qu'ils prennent tout leur sens en statistiques et en informatique, à la fois comme indicateur et comme vérificateur de données.

Passons maintenant aux propriétés des opérateurs m, V et donc σ. m est additive c'est à dire que

[V4] $m(A + B) = m(A) + m(B)$

sous réserve que A et B soient deux variables aléatoires définies pour la même probabilité. m est aussi homogène soit :

[V5] $m(k.A) = k.m(A)$

où k est une constante. Ces deux propriétés font de m une *application linéaire.* Par contre la moyenne n'est pas multiplicative car $m(A.B) \neq m(A)m(B)$ en général (un contre-exemple suffit à le montrer). La variance n'est en général pas additive

[V6] $V(A + B) \neq V(A) + V(B)$

et non homogène puisque

[V7] $V(k.A) = k^2.V(A)$

On en déduit que σ n'est pas additive et que de plus

[V8] $\sigma(k.A) = |k|.\sigma(A)$

Le fait que V ne soit pas additive résulte du fait que m n'est pas multiplicative. En effet :

[V9] $V(A+B) - V(A) - V(B) = 2.(m(A.B) - m(A).m(B))$

On appelle covariance de A et B la quantité $m(A.B) - m(A).m(B))$ notée $cov(A,B)$ qu'on "normalise" pour obtenir le coefficient de corrélation linéaire entre A et B noté souvent ρ plutôt que $corr$:

$$\rho(A,B) = \frac{cov(A,B)}{\sigma(X).\sigma(Y)}$$

Les résultats suivants sont simples à démontrer

[R1] $Y = aX + b \Rightarrow |\rho(X,Y)| = 1$

[R2] $\forall A,B - 1 \leq \rho(A,B) \leq 1$

[R3] A et B indépendantes $\Rightarrow cov(A,B) = 0$

[R4] la relation $\mathcal{R}$ définie par
$x\mathcal{R}y \;\Leftrightarrow\;$ x et y sont en corrélation linéaire
est une relation d'équivalence.

Attention toutefois à ce que la réciproque que **[R3]** est fausse.

2.4 Lois Classiques

Les lois discrètes sont en général assez simples à aborder et le calcul de leurs caractéristiques (moyenne, écart-type) est assez facile à mener. Par contre, pour les lois continues et en particulier pour les dernières (Khi-deux, Student, Fisher-Snédécor) ces calculs demandent des techniques poussées d'intégration. Nous les citons cependant car elles sont importantes pour les test statistiques. De plus, elles font partie de la culture statistique et il est bon de les connaitre, au moins par leur nom et leur définition, si ce n'est par leurs formules.

2.4.1 Lois Discrètes

Loi de Bernoulli

C'est sans doute la loi discrète la plus simple à étudier puisqu'elle n'a que deux valeurs 0 et 1. Il ne faut pas pour autant la négliger puisqu'elle sert de support à tous les calculs sur les phénomènes binaires comme succés/échec, oui/non, présence/absence... On note traditionnellement $b(p)$ cette loi où p désigne la probabilité de $X = 1$. Sa moyenne est p, son écart-type $\sqrt{p.(1-p)}$.

Loi Binomiale

Cette loi requiert deux paramètres : n et p. On la note $\mathcal{B}(n,p)$. La façon la plus naturelle de présenter cette loi discrète est de la considérer comme la somme de n variables de Bernoulli indépendantes de même paramètre p. Sa moyenne est np et son écart-type $\sqrt{n.p.(1-p)}$. On peut aussi la présenter en termes de tirages de boules. Supposons qu'on tire avec remise n boules parmi un ensemble de boules de deux types (par exemple de deux couleurs, blanc et noir). Soit p la proportion de boules blanches. Alors $\mathcal{B}(n,p)$ peut s'interpréter comme la variable "nombre de boules blanches tirées". L'évènement "on a tiré k boules blanches" a comme probabilité $p_k = p(X = k) = C_n^k.p^k.(1-p)^{n-k}$ pour k de 0 à n.

Loi Uniforme Discrète

C'est là encore une loi très simple : p_X est constante. On la définit classiquement sur l'intervalle $[\![1,n]\!] = [1,n] \cap \mathbb{N}$. Chaque probabilité élémentaire vaut donc $1/n$. La moyenne de $\mathcal{U}(n)$ est $(n+1)/2$ et son écart-type $\sqrt{(n^2-1)/12}$.

Loi Géométrique

Reprenons l'interprétation en termes de boules pour la loi Binomiale : on a des boules de deux couleurs, blanc et noir. Les boules blanches sont en proportion p. La loi Géomètrique $\mathcal{G}(p)$ correspond à la variable aléatoire X où $X = k$ est l'évènement "on a obtenu une (première) boule blanche seulement au k-ième tirage". La probabilité associée est $p_k = (1-p)^{k-1}.p$ d'où une moyenne de $1/p$ et un écart-type de $\sqrt{(1-p)/p^2}$.

Loi de Pascal

Nommée aussi loi binomiale négative, cette loi généralise la loi géométrique. On s'intéresse à l'obtention de s boules blanches au bout de k tirages pour des boules de deux couleurs dont la proportion de boules blanches est p. La probabilité associée est $p_k = C_{k-1}^{s-1}.p^s.(1-p)^{k-s}$. La moyenne de $\mathcal{P}(s,p)$ est s/p et son écart-type de $\sqrt{s(1-p)/p^2}$.

Loi Hypergéométrique

On note $H(N,K,n)$ cette loi. C'est la loi de comptage d'un caractère binaire dans un tirage sans remise de n boules. Ce caractère binaire correspond à K boules parmi N boules en tout. Elle prend les valeurs $a,a+1...b$ avec $a = max(0,n-(N-K))$ et $b = min(n,K)$. La probabilité d'obtenir la valeur k est $C_K^k C_{N-K}^{n-k}/C_N^n$.

Sa moyenne est nK/N, sa variance $(N-n)(nK/N)(1-K/N)/(N-1)$.

Loi de Poisson

De paramètre λ, cette loi discrète est un peu surprenante puisqu'elle prend un nombre infini de valeurs à savoir 0,1,2.... Nommée aussi "loi des évènements rares", elle sert à modéliser les phénomènes de files d'attente. La probabilité d'obtenir le nombre entier k est $e^{-\lambda}\lambda^k/k!$. Sa moyenne est λ et son écart-type $\sqrt{\lambda}$. On la note $\mathcal{P}(\lambda)$.

On pourra utiliser le tableau suivant pour reconnaitre une "loi de boules"

	loi de comptage	loi d'apparition (1ère fois)
avec remise	*Binomiale*	*Géométrique*
sans remise	*Hypergéométrique*	(sans nom)

2.4.2 Lois Continues

Loi Uniforme Continue

Elle est constante, comme son nom l'indique. On la définit classiquement sur $[a,b]$. Sa moyenne est donc $(a+b)/2$ et son écart-type $\sqrt{(b-a)^2/12}$. On la note $\mathcal{U}([a,b])$.

Loi Triangulaire

Cette loi $\mathcal{T}([a,b])$ est définie sur $[a,b]$. Elle est composée de deux segments de droite qui se rencontrent au milieu de l'intervalle à une hauteur de $2/(b-a)$ ce qui dessine donc un triangle isocèle. On laisse en exercice la détermination de la probabilité élémentaire, le calcul de la moyenne qui vaut $(a+b)/2$ et de l'écart-type $(b-a)/2\sqrt{6}$.

Loi Exponentielle

Avec un paramètre α strictement positif, la loi $\mathcal{E}(\alpha)$ est définie sur $\mathbb{R}_+$ par la densité $f(x) = \alpha.e^{-\alpha x}$. Sa moyenne est $1/\alpha$ et son écart-type aussi.

Loi de Erlang-Gamma

Cette loi est définie par la densité $f(x) = \alpha^k.x^{k-1}.e^{-\alpha x}/\Gamma(k)$ pour $x \geq 0$, $\alpha > 0$ et $k > 0$ fixés, où Γ est bien sùr l'intégrale eulérienne $\Gamma(k) = \displaystyle\int_0^{+\infty} y^{k-1}.e^{-y}.dy$ qui généralise la fonction factorielle : $\Gamma(n) = (n-1)!$ pour n entier. La moyenne de la loi d'Erlang est k/α, son écart-type $\sqrt{k}/\alpha$.

Loi de Weibull-Rayleigh

La loi de Weibull est notée $\mathcal{W}(\alpha,\beta,\gamma)$. Elle est définie par la densité $f(x) = (\beta/\alpha).y^{\beta-1}e^{-y^\beta}$ pour $\gamma \leq x$, $\alpha > 0$, $\beta > 0$ et γ fixés, où $y = (x-\gamma)/\alpha$. On peut vérifier que α est un paramètre d'échelle, β un paramètre de forme et γ un paramètre de position. Sa moyenne est $\gamma + \alpha.\Gamma(1+1/\beta)$, son écart-type $\alpha.\sqrt{\Gamma(1+2/\beta) - \Gamma(1+1/\beta)}$. La loi de Rayleigh est $\mathcal{W}(\alpha,2,0)$.

Loi Normale

Cette loi à deux paramètres μ et ν a pour densité

$$f(x) = \frac{1}{\sigma.\sqrt{2\pi}}.e^{-t^2/2}$$

où t est $(x-\mu)/\nu$. Sa moyenne vaut μ, son écart-type ν et on la note donc classiquement $\mathcal{N}(m,\sigma)$.

Loi de Cauchy

Si U et V sont deux lois normales centrées réduites, la variable $X = U/V$ est appelée loi de Cauchy. Elle n'admet ni moyenne ni écart-type.

Loi Log-Normale

Si U est une loi normale de moyenne μ et d'écart-type σ, la variable X définie par $X = e^U$ est appelée loi log-normale. Sa moyenne est $x^{\mu+\sigma^2/2}$, son écart-type $\sqrt{e^{2\mu+\sigma^2}e^{\sigma^2} - 1}$.

Loi du Khi-Deux

Si Z_1, Z_2 Z_ν... sont ν variables normales centrées réduites, alors la variable $\Sigma_i Z_i^2$ est une variable de Khi-deux à ν degrés de liberté. Sa moyenne est ν, son écart-type $\sqrt{2\nu}$.

Loi de Student

Si U est une variable normale centrée réduite, si V est une variable de Khi-deux à ν degrés de liberté et si U et V sont indépendantes, alors la variable $T = U/\sqrt{V/\nu}$ est une variable de Student à ν degrés de liberté. Sa moyenne est 0, son écart-type $\sqrt{\nu/(\nu-2)}$ pour $\nu > 2$.

Loi de Fisher-Snédécor

Si U est une variable de Khi-deux à ν_1 degrés de liberté et si V est une variable de Khi-deux à ν_2 degrés de liberté alors la variable $T = (U/\nu_1)/(V/\nu_2)$ est une variable de Fisher-Snédécor à (ν_1,ν_2) degrés de liberté. Sa moyenne est $\nu_2/(\nu_2 - 2)$ pour $\nu_2 > 2$.

2.5 Remarques sur les lois

S'il est difficile *a priori* de s'y retrouver dans toutes ces lois – et nous n'en avons donné que la définition mathématique – c'est avec la pratique de l'approximation statistique, c'est à dire avec l'usage que l'on "reconnait" une distribution normale, fishérienne etc.

Nous conseillons vivement aux lecteurs et lectrices de cours de trouver un logiciel qui permet de tracer ces différentes lois pour "jouer" avec les paramètres et reconnaitre ici un paramètre de forme, là un paramètre d'allongement...

2.6 Démonstrations des propriétés

• *Démonstration de la propriété* **[P3]**

D'après **[P2]** appliqué à $A = E$ et $B = \emptyset$ on a $p(E \cup \emptyset) = p(E) + p(\emptyset)$ car $E \cap \emptyset = \emptyset$. Puisque $E \cup \emptyset = E$, le membre de gauche est aussi égal à $p(E)$. On en déduit donc que $p(E) = p(E) + p(\emptyset)$. Dans R, puisque $p(E) = 1$, on obtient, en simplifiant par $p(E)$, le résultat recherché. Sous forme uniquement calculatoire, on peut réécrire cela en :

$$\begin{array}{llllll} E \cup \emptyset = E & \Rightarrow & p(E \cup \emptyset) & = & p(E) & \\ & \Rightarrow & p(E) + p(\emptyset) & = & p(E) & \text{car } E \cap \emptyset = \emptyset \\ & \Rightarrow & p(\emptyset) & = & 0 \ \diamondsuit & \end{array}$$

• *Démonstration de la propriété* **[P4]**

D'après **[P2]** appliqué à $A = F$ et $B = \overline{F}$ on a $p(F \cup \overline{F}) = p(F) + p(\overline{F})$ car $F \cap \overline{F} = \emptyset$. Puisque $F \cup \overline{F} = E$, le membre de gauche est aussi égal à $p(E)$. On en déduit donc que $1 = p(F) + p(\overline{F})$ grâce à **[P1]**. Passant p(F) dans le membre de gauche on en déduit le résultat recherché. Sous forme d'équations :

$$\begin{array}{llllll} F \cup \overline{F} = E & \Rightarrow & p(F \cup \overline{F}) & = & p(E) & \\ & \Rightarrow & p(F) + p(\overline{F}) & = & 1 & \text{car } F \cap \overline{F} = \emptyset \text{ et } p(E) = 1 \\ & \Rightarrow & p(\overline{F}) & = & 1 - p(F) \ \diamondsuit & \end{array}$$

• *Démonstration de la propriété* **[P5]**

Nous aurons besoin du résultat suivant : $p(B \backslash C) = p(B) - p(C)$ si $C \subset B$ qu'on démontre via **[P2]** pour $B = C \cup (B \backslash C)$ puisque $C \cap (B \backslash C) = \emptyset$. Nous donnons la solution sous forme d'équations et d'implications, la justification des calculs étant laissée au lecteur. On a posé $D = B \backslash (A \cap B)$.

$$\begin{array}{llll} A \cup B = A \cup D & \Rightarrow & p(A \cup B) & = p(A \cup D) \\ & & & = p(A) + p(D) \\ & & & = p(A) + p(B \backslash (A \cap B)) \\ & & & = p(A) + p(B) - p(A \cap B) \ \diamondsuit \end{array}$$

• *Démonstration de la propriété* **[P6]**

Dire que p est croissante revient à dire que $A \subset B \Rightarrow p(A) \leq p(B)$. Or $A \subset B \Rightarrow B = A \cup (B \backslash A)$. En appliquant **[P1]**, on en déduit que $p(B) = p(A) + p(B \backslash A)$. Comme $\forall X p(X) \geq 0$, il vient $p(B) - p(A) \geq 0. \diamondsuit$

• *Démonstration de la propriété* **[P7]**

Commençons par démontrer la *formule des probabilités totales* qui en est un cas particulier. La définition de probabilité conditionnelle permet d'écrire "linéairement" $p(X \cap Y) = p(X).p(Y|X)$. Les calculs sont simples à mener :

$$\begin{array}{rrcl} & A & = & A \cap E \\ & & = & A \cap (B \cup \overline{B}) \\ & & = & (A \cap B) \cup (A \cap \overline{B}) \\ \text{d'où} & p(A) & = & p(A \cap B) + p(A \cap \overline{B}) \end{array}$$

et on conclut en remplaçant les deux termes du membre de droite par leur écriture "linéaire". Pour la formule générale, on applique la même méthode.

$$\begin{array}{rrcl} & A & = & A \cap E \\ & & = & A \cap (\bigcup_i \ \omega_i) \\ & & = & \bigcup_i \ (A \cap \omega_i) \\ \text{d'où} & p(A) & = & \Sigma_i \ p(A \cap \omega_i) \ \diamondsuit \end{array}$$

• *Démonstration de la propriété* **[P8]**

Commençons par démontrer la *formule de Bayes* qui en est un cas particulier. On commence par remarquer que $p(A \cap B) = p(B|A).p(A) = p(A|B).p(B)$. Puis :

$$\begin{array}{rcl} p(B|A) & = & p(A \cap B)/p(A) \text{ par définition} \\ & = & p(A|B).p(B)/p(A) \text{ compte-tenu de la remarque précédente} \\ & = & p(A|B).p(B)/p(A \cap B) + p(A \cap \overline{B}) \\ & & \text{grâce à la formule précédente des probabilités totales. } \diamondsuit \end{array}$$

Avec un système complet d'évènements, c'est à peine plus difficile :

$$\begin{array}{rcl} p(\omega_i|A) & = & p(A \cap \omega_i)/p(A) \\ & = & p(A|\omega_i).p(\omega_i)/p(A) \\ & = & p(A|\omega_i).p(\omega_i)/\Sigma_i \ p(A \cap \omega_i) \ \diamondsuit \end{array}$$

• *Démonstration de la propriété* **[P9]**

Il n'y a que les propriétés **[P1]** et **[P2]** à démontrer pour $p_{|A}$.

$$\begin{array}{rcl} p_{|A}(E) & = & p(E \cap A)/p(A) \\ & = & p(A)/p(A) \\ & = & 1 \ \diamondsuit \end{array}$$

donc la première propriété est démontréee.

De même :

$$\begin{array}{rcl} p_{|A}(X \cup Y) & = & p((X \cup Y) \cap A)/p(A) \\ & = & p((X \cap A) \cup (Y \cap A))/p(A) \\ & = & p(X \cap A)/p(A) + p(Y \cap A)/p(A) \\ & = & p_{|A}(X) + p_{|A}(Y) \ \diamondsuit \end{array}$$

et la deuxième propriété est aussi démontrée.

• *Démonstration de la propriété* **[V1]**

Remarquons tout d'abord que p_X n'est pas bien définie. Nous disposons des fonctions

$$\begin{array}{lcccl} p & : & \mathcal{P}(E) & \to & [0,1] \\ X & : & \mathcal{P}(E) & \to & I \\ p_X & : & I & \to & [0,1] \end{array}$$

Etendons la définition de p_X des singletons $\{k\}$ à un ensemble quelconque par $p_X(J) = \sum_{j \in J} p_X(\{j\})$ où $J \subset I$. Soient A et B sous-ensembles disjoints de I. Alors

$$\begin{array}{rcl} p_X(A \cup B) & = & \displaystyle\sum_{x \in A \cup B} p_X(\{x\}) \\ & = & \displaystyle\sum_{a \in A} p_X(\{a\}) + \sum_{b \in B} p_X(\{b\}) \\ & = & p_X(A) + p_X(B) \end{array}$$

De plus $p_X(I) = p(X^{-1}(I)) = p(E) = 1$ donc p_X est bien une probabilité.$\diamondsuit$

• *Démonstration de la propriété* **[V2]**

Là encore p_Y n'est définie que pour un intervalle. On l'étend naturellement pour tout *borélien*, c'est à dire pour toute réunion disjointe d'intervalles (notée $\dot\cup$ au lieu de $\cup$) par $p_Y(J) = \sum_{\dot\cup[a_i,b_i]=J} p_Y([a_i,b_i])$. Soient A et B sous-ensembles disjoints de $\mathbb{R}$. Alors

$$\begin{array}{rcl} p_Y(A \cup B) & = & \displaystyle\sum_{\dot\cup[x_i,y_i]=A\cup B} p_Y([x_i,y_i]) \\ & = & \displaystyle\sum_{\dot\cup[a_i,b_i]=A} p_Y([a_i,b_i]) + \sum_{\dot\cup[c_i,d_i]=B} p_Y([x_i,c_i]) \\ & = & p_Y(A) + p_Y(B) \ \diamondsuit \end{array}$$

• *Démonstration de la propriété* **[V3]**

$$\begin{aligned} mt_c(X) &= moy(\ (X-m)^2\ ) \\ &= moy(X^2 - 2.m(X).X + m(x)^2) \\ &= moy(X^2) - 2.moy(X)m(X) + m(X)^2 \\ &= moy(X^2) - 2.m(X)^2 + m(X)^2 \\ &= moy(X^2) - m(X)^2 \\ &= V(X)\ \diamondsuit \end{aligned}$$

• *Démonstration de la propriété* **[V4]**

Pour le cas discret :

$$\begin{aligned} m(A+B) &= \Sigma(a_i + b_i).p_i \\ &= (\Sigma a_i.p_i) + (\Sigma b_i.p_i) \\ &= m(A) + m(B)\ \diamondsuit \end{aligned}$$

Pour le cas continu :

$$\begin{aligned} m(A+B) &= \int \Big(a(x) + b(x)\Big).p(x)dx \\ &= \int a(x).p(x)dx + \int b(x).p(x)dx \\ &= m(A) + m(B)\ \diamondsuit \end{aligned}$$

• *Démonstration de la propriété* **[V5]**

Pour le cas discret :

$$\begin{aligned} m(k.A) &= \Sigma k.a_i.p_i \\ &= k.\Sigma a_i.p_i \\ &= k.m(A)\ \diamondsuit \end{aligned}$$

Pour le cas continu :

$$\begin{aligned} m(k.A) &= \int k.a(x).p(x)dx \\ &= k.\int a(x).p(x)dx \\ &= k.m(A)\ \diamondsuit \end{aligned}$$

La Démonstration de la propriété **[V6]** consiste à trouver un contre-exemple. Elle est laissée en exercice.

• *Démonstration de la propriété* **[V7]**

Pour le cas discret :

$$\begin{aligned} mt(k.A,2) &= \Sigma(k.a_i)^2.p_i \\ &= \Sigma k^2.a_i^2.p_i \\ &= k^2.\Sigma a_i^2.p_i \\ &= k.mt(A,2) \ \diamondsuit \end{aligned}$$

De même

$$\begin{aligned} m(k.A)^2 &= (k.m(A))^2 \\ &= k^2.m(A)^2 \end{aligned}$$

Donc finalement

$$\begin{aligned} V(k.A) &= mt(k.A,2) - m(k.A)^2 \\ &= k^2.mt(A,2) - k^2.m(A)^2 \\ &= k^2.(mt(A,2) - .m(A)^2) \\ &= k^2.V(A) \ \diamondsuit \end{aligned}$$

Pour le cas continu, même démarche puisque

$$\begin{aligned} mt(k.A,2) &= \int (\ k.a(x)\ )^2.p(x)dx \\ &= \int k^2.a(x)^2.p(x)dx \\ &= k^2.\int a(x)^2.p(x)dx \\ &= k.mt(A,2) \end{aligned}$$

• *Démonstration de la propriété* **[V8]**

$$\begin{aligned} \sigma(k.A) &= \sqrt{V(k.A)} \\ &= \sqrt{k^2.V(A)} \\ &= |k|.\sqrt{V(A)} \\ &= |k|.\sigma(A) \ \diamondsuit \end{aligned}$$

• *Démonstration de la propriété* **[V9]**

Comme :

$$\begin{aligned} mt((A+B),2) &= moy(\ (A+B)^2\ ) \\ &= moy(\ A^2 + B^2 + 2.A.B\ ) \\ &= moy(A^2) + moy(B^2) + 2.moy(A.B) \end{aligned}$$

et puisque :

$$\begin{aligned} moy((A+B)^2) &= (\ moy(A) + moy(B)\ )^2 \\ &= moy(A)^2 + moy(B)^2 + 2.moy(A).moy(B) \end{aligned}$$

on en déduit

$$\begin{aligned} var(A+B)) &= mt(A+B,2) - m(A+B)^2 \\ &= moy(A^2) + moy(B^2) + 2.moy(A.B) \\ &\quad \text{-}\ (\ moy(A)^2 + moy(B)^2 + 2.moy(A).moy(B)\ ) \\ &= moy(A^2) - moy(A)^2 + moy(B^2) - moy(B)^2 \\ &\quad + 2.(\ moy(A.B) - moy(A).moy(B)\ ) \\ &= var(A) + var(B) + 2.cov(A,b)\ \diamondsuit \end{aligned}$$

• *Démonstration de la propriété* **[R1]**

Si $Y = aX + b$ alors

$$\begin{aligned} cov(X,Y) &= m(\ X.(a.X+b)\ ) - m(X).m(a.X+b) \\ &= m(a.X^2 + b.X) - m(X).(\ a.m(X) + b\ ) \\ &= a.m(X^2) + b.m(X) - a.m(X)^2 - b.m(X) \\ &= a.(\ m(X^2)m(X)^2\ ) \\ &= a.V(X) \end{aligned}$$

Donc $\rho(X,Y) = cov(X,Y)/\sigma(X).\sigma(Y) = a.V(X)/|a|.\sigma(X).\sigma(X)$ et on a bien $|\rho(X,Y)| = |\ a/|a|\ | = 1.\diamondsuit$

La réciproque de la propriété peut s'énoncer ainsi :

$$|\rho(X,Y)| = 1 \Rightarrow \exists a,b\ ;\ m(\ (Y - (a.X+b))^2\ ) = 0.$$

Pour le montrer, considérons la variable $Z = m_X + \rho.(Y - m_Y).\sigma_X/\sigma_Y$ où ρ est une notation simplifiée de $\rho(X,Y)$. Notons aussi $\tau = \rho.\sigma_X/\sigma_Y$ pour simplifier les calculs. Alors

$$\begin{aligned} m(\ (Z-X)^2\ ) &= m(\ (X-m_X)^2\ ) - 2.\tau.m(\ (X-m_X).(Y-m_Y)\ ) \\ &\quad +\tau^2.m(\ (Y-m_Y)^2\ ) \\ &= V(X) - 2.\tau.cov(X,Y) + \tau^2.V(Y) \\ &= V(X) - 2.(\rho.\sigma_X/\sigma_Y).(\rho.\sigma_X.\sigma_Y) + (\rho.\sigma_X/\sigma_Y)^2.V(Y) \\ &= V(X) - 2.\rho.\sigma_X^2 + \rho.\sigma_X^2 \\ &= V(X).(1-\rho^2) \end{aligned}$$

Donc $|\rho| = 1 \ \Rightarrow\ m(\ (Z-X)^2\ ) = 0$. Cela signifie donc qu'en moyenne (quadratique) Z est ègal à X. Or $X = Z$ s'écrit $X = m_X + \rho.(Y - m_Y).\sigma_X/\sigma_Y$ soit encore $\sigma_Y.(X - m_X) = \rho.(Y - m_Y).\sigma_X.(Y - m_Y)$ ce qui donne la relation "linéaire" $Y = a.X + b$ avec $a = \rho.\sigma_Y/\sigma_X$ et $b = m_Y - a.m_x.\diamondsuit$

• *Démonstration de la propriété* **[R2]**

Soit Z la variable aléatoire $Z = X - m_X + \lambda.(Y - m_Y)$. Alors la quantité positive ou nulle $m(Z^2)$ est le trinome du second degré $T(\lambda) = a.\lambda^2 + b.\lambda + c$ avec :

$$\begin{array}{lclcl} a & = & m(\ (Y - m_Y)^2\ ) & = & V(Y) \\ b & = & 2.m(\ (X - m_X).(Y - m_Y)\ ) & = & 2.cov(X,Y) \\ c & = & m(\ (X - m_X)^2\ ) & = & V(X) \end{array}$$

Puique $T(\lambda)$ doit être de signe constant, son discrimant réduit est négatif, soit la relation :

$$cov(X,Y)^2 \leq V(X).V(Y)$$

nommée *inégalité de Schwartz*. En prenant la racine carrée des deux termes, on obtient

$$|cov(X,Y)| \leq \sigma(X).\sigma(Y)$$

et si $\sigma(X).\sigma(Y)$ est non nul, on peut diviser par $\sigma(X).\sigma(Y)$ d'où la relation cherchée $|\rho(X,Y)| \leq 1.\diamondsuit$

• *Démonstration de la propriété*

[R3]

Cet énoncé n'a aucun sens. La notion de variables aléatoires indépendantes n'a jamais été introduite. On consultera le chapitre 6 pour voir cette notion.

2.7 Démonstrations de certains calculs de m et σ pour les lois classiques

• *Calcul de m et σ pour la loi de Bernoulli*

Vérifions que les p_i définissent une probabilité. Puisque $0 \leq p \leq 1$, $p_1 = p$ et $p_0 = 1 - p$ sont des nombres positifs compris entre 0 et 1. De plus $p_0 + p_1 = (1-p) + p = 1.\Diamond$

$$\begin{aligned}
m_{b(p)} &= \Sigma\ x_i.p_i \\
&= 0.(1-p) + 1.p \\
&= p \\
mt^2_{b(p)} &= \Sigma\ x_i^2.p_i \\
&= 0^2.(1-p) + 1^2.p \\
&= p \\
V_{b(p)} &= p - p^2 \\
&= p(1-p)\Diamond
\end{aligned}$$

• *Calcul de m et σ pour la loi Binomiale*

La méthode la plus simple consiste à utiliser la définition de $\mathcal{B}(n,p)$ comme somme de n variables de bernoulli indépendantes car alors

$$m_{\mathcal{B}(n,p)} = m(\ \Sigma\ b(p)\ ) = \Sigma\ m(\ b(p)\ ) = n.p$$

$$m_{\mathcal{B}(n,p)} = V(\ \Sigma\ b(p)\ ) = \Sigma\ V(\ b(p)\ ) = n.p.(1-p)$$

Mais il est également possible de calculer explicitement ces valeurs. Vérifions que les p_i définissent une probabilité. Puisque $0 \leq p \leq 1$, p^k et $(1-p)^{n-k}$ sont des nombres positifs compris entre 0 et 1. C_n^k est un entier positif et $\Sigma p_k = C_n^k.p^k.(1-p)^{n-k} = (p + (1-p))^n = 1$ d'après la formule du binome de Newton.$\Diamond$

$$\begin{aligned}
m_{\mathcal{B}(n,p)} &= \sum_{k=0}^{n} x_k.p_k \\
&= \sum_{k=0}^{n} k.C_n^k.p^k.(1-p)^{n-k} \\
&= \sum_{k=1}^{n} (k.n!/k!(n-k)!).p^k.(1-p)^{n-k} \\
&= \sum_{k=1}^{n} n.((n-1)!/(\ (k-1)!((n-1)-(k-1))!\ ).p.p^{k-1}.(1-p)^{n-k}
\end{aligned}$$

Après simplification,

$$\begin{aligned}
m_{\mathcal{B}(n,p)} &= n.p.\sum_{k=1}^{n} C_{n-1}^{k-1}.p^{k-1}.(1-p)^{n-k} \\
&= n.p.\sum_{k=0}^{n-1} C_{n-1}^{k}.p^{k}.(1-p)^{n-k} \\
&= n.p.[\ (p+(1-p)\ ]^{n-1} = n.p.1^{n-1} = n.p.
\end{aligned}$$

Pour le moment d'ordre 2, on utilise la relation $k^2 = k(k-1)+k$:

$$\begin{aligned}
mt^2_{\mathcal{B}(n,p)} &= \sum_{k=0}^{n} x_k^2.p_k \\
&= \sum_{k=0}^{n} k^2.C_n^k.p^k.(1-p)^{n-k} \\
&= \sum_{k=0}^{n} (\ k.(k-1)+k)\ ).C_n^k.p^k.(1-p)^{n-k} \\
&= n.(n-1).p^2.\sum_{k=2}^{n} .C_{n-2}^{k}.p^{k-2}.(1-p)^{n-k} + n.p \\
&= n.(n-1).p^2.1^{n-2} + n.p = n.(n-1).p^2 + n.p \\
V_{\mathcal{B}(n,p)} &= n.(n-1).p^2 + n.p - (n.p)^2 \\
&= n.p.(1-p)\diamondsuit
\end{aligned}$$

• *Calcul de m et σ pour la loi Uniforme Discrète*

Vérifions que les p_k définissent une probabilité. Puisque $a < b$, les p_k qui valent tous $1/n$ sont des nombres positifs compris entre 0 et 1. De plus $\sum_{i=1}^{n} p_k = \sum_{i=1}^{n} 1/n = n.1/n = 1.\diamondsuit$

$$\begin{aligned}
m_{\ \mathcal{U}(n)} &= \sum_{i=1}^{n} x_i.p_i \\
&= \sum_{i=1}^{n} i.\frac{1}{n} \\
&= \frac{n.(n+1)}{2}.\frac{1}{n} \quad = \quad (n+1)/2
\end{aligned}$$

$$\begin{aligned}
mt^2{}_{\mathcal{U}(n)} &= \sum_{i=1}^{n} x_i^2.p_i \quad = \sum_{i=1}^{n} i^2.\frac{1}{n} \\
&= \frac{n.(n+1).(2n+1)}{6}.\frac{1}{n} \quad = (n+1).(2n+1)/6
\end{aligned}$$

$$\begin{aligned}
V_{\mathcal{U}(n)} &= \frac{(n+1).(2n+1)}{6} - \left[\frac{n.(n+1)}{2}\right]^2 \\
&= \frac{(n+1)}{2}.\left[\frac{2n+1}{3} - \frac{n+1}{2}\right] \\
&= \frac{(n+1)}{2}.\left[\frac{4n+2-3n-3}{2}\right] \\
&= \frac{(n+1)}{2}.\left[\frac{n-1}{2}\right] \quad = (n^2-1)/12.
\end{aligned}$$

- *Calcul de m et σ pour la loi Géométrique*

On utilise la fonction f définie en x par la somme de la série $\sum x^{k-1}$ pour $x < 1$. On démontre que $f(x) = 1/(1-x)$. On a également besoin de $f'(x) = 1/(1-x)^2$ et de $f''(x) = 2/(1-x)^3$.

Vérifions que les p_i définissent une probabilité. Puisque $0 \leq p \leq 1$, le nombre $p.(1-p)^{n-k}$ est positif et compris entre 0 et 1. De plus

$$\sum_{k \in \mathbb{N}} p.(1-p)^{k-1} = p.f(1-p) = p/(1-(1-p)) = p/p = 1/\diamondsuit$$

ce qui permet de calculer

$$\begin{aligned}
m_{\mathcal{G}(n,p)} &= \sum_{k \in \mathbb{N}} x_k.p_k \\
&= \sum_{k \in \mathbb{N}} k.p.(1-p)^{k-1} \\
&= p.f'(1-p) \\
&= p.1/p^2 \\
&= 1/p
\end{aligned}$$

$$\begin{aligned} mt^2{}_{\mathcal{G}(n,p)} &= \sum_{k\in\mathbb{N}} x_k^2.p_k \\ &= \sum_{k\in\mathbb{N}} k^2.p.(1-p)^{k-1} \\ &= \sum_{k\in\mathbb{N}} (\ k(k-1)+k\ ).p.(1-p)^{k-1} \\ &= p.(1-p).f''(1-p)+1/p \\ &= 2(1-p)/p^2+1/p\ =\ (2-p)/p^2 \\ V_{\mathcal{G}(n,p)} &= (2-p)/p^2-(1/p)^2\ =\ (1-p)/p^2. \end{aligned}$$

• *Calcul de m et σ pour la loi de Pascal*

Plutôt que d'effectuer des calculs, remarquons que la loi de Pascal peut être considérée comme la somme de s lois géométriques indépendantes. Car obtenir s boules peut être vu comme obtenir s fois 1 boule (*sic !*). D'où

$m\ _{\mathcal{P}}(n,p) = s/p$ et $V_{\mathcal{P}}(n,p) = s/(1-p)/p^2$.

• *Calcul de m et σ pour la loi de Poisson*

Vérifions que les p_i définissent une probabilité. Pour tout λ $e^{-\lambda}.\dfrac{\lambda^k}{k!}$ est un nombre positif. De plus, $e^x = \sum_{k\in\mathbb{N}} \dfrac{x^k}{k!}$

$$\text{Donc}\quad \begin{aligned} \sum_{k\in\mathbb{N}} e^{-\lambda}.\frac{\lambda^k}{k!} &= e^{-\lambda}.\sum_{k\in\mathbb{N}} \frac{\lambda^k}{k!} \\ &= e^{-\lambda}.e^{\lambda} \\ &= 1.\diamondsuit \end{aligned}$$

$$\begin{aligned}
m_{\ \mathcal{P}(\lambda)} &= \sum_{k\in\mathbb{N}} x_k.p_k = \sum_{k\in\mathbb{N}} k.e^{-\lambda}.\frac{\lambda^k}{k!} \\
&= e^{-\lambda}.\lambda \sum_{k\in\mathbb{N}} \frac{\lambda^{k-1}}{(k-1)!} = e^{-\lambda}.\lambda.e^{\lambda} = \lambda \\
{mt^2}_{\mathcal{P}(\lambda)} &= \sum_{k\in\mathbb{N}} x_k^2.p_k = \sum_{k\in\mathbb{N}} k^2.e^{-\lambda}.\frac{\lambda^k}{k!} \\
&= \sum_{k\in\mathbb{N}} (\ k(k-1)+k\ ).e^{-\lambda}.\frac{\lambda^k}{k!} = e^{-\lambda}.\lambda^2 \sum_{k\in\mathbb{N}} \frac{\lambda^{k-1}}{(k-1)!} + \lambda \\
&= \lambda^2+\lambda \\
V_{\ \mathcal{P}(\lambda)} &= \lambda^2+\lambda-(\lambda)^2 = \lambda
\end{aligned}$$

• *Calcul de m et σ pour la loi Uniforme Continue*

Vérifions que f est une densité de probabilité. f est constante sur $[a,b]$ et nulle ailleurs, soit : $f(x) = k$ si $x \in [a,b]$.

$$\begin{aligned}
\int_{\mathbb{R}} f &= \int_{-\infty}^{a} f + \int_a^b f + \int_b^{+\infty} f \\
&= \int_a^b k.dx = k.[x]_a^b = k(b-a)
\end{aligned}$$

Donc $\int f = 1$ pour $k(b-a) = 1$ soit $k = 1/(b-a)$.$\diamondsuit$

$$
\begin{aligned}
m_{\mathcal{U}([a,b])} &= \int_{\mathbb{R}} Id.f = \int_a^b x.k.dx \\
&= k.\left[\frac{x^2}{2}\right]_a^b = (1/(b-a)).(b^2-a^2)/2 = (a+b)/2 \\
{mt^2}_{\mathcal{U}([a,b])} &= \int_{\mathbb{R}} Id^2.f = \int_a^b x^2.k.dx \\
&= k.\left[\frac{x^3}{3}\right]_a^b = (1/(b-a)).(b^3-a^3)/2 = (a^2+ab+b^2)/3 \\
V_{\mathcal{U}([a,b])} &= (a^2+ab+b^2)/3 - ((a+b)/2))^2 \\
&= (4a^2+4ab+4b^2-3a^2-3b^2-6ab)/12 \\
&= (a^2-2ab+b^2)/12 = (a-b)^2/12.
\end{aligned}
$$

- *Calcul de m et σ pour la loi Loi Triangulaire*

f est un segment de droite sur $[a,m]$ pour $m = (a+b)/2$ qui vaut 0 en a. Donc $f(x) = \alpha.(x-a)$ et comme on impose $f(m) = 2/(b-a)$ on trouve $\alpha = 4/(b-a)^2$. Par le même raisonnement, $f(x) = -\alpha.(x-b)$ sur $[m,b]$. Montrons que f est une densité de probabilité.

$$
\begin{aligned}
\int_{\mathbb{R}} f &= \int_a^b f = \int_a^m f + \int_m^b f \\
&= \int_a^m \alpha.(x-a)dx + \int_m^b -\alpha.(x-b)dx \\
&= \alpha.[(x-a)^2/2]_a^m - \alpha.[(x-b)^2/2]_m^b \\
&= \frac{\alpha\, a^2}{4} - \frac{\alpha\, ab}{2} + \frac{\alpha\, b^2}{4} \\
&= \alpha.(a^2-2ab+b^2)/4 = 1
\end{aligned}
$$

$$
\begin{aligned}
m_{\mathcal{T}([a,b])} &= \int_{\mathbb{R}} Id.f = \int_a^m Id.f + \int_m^b Id.f \\
&= \alpha.\left[\frac{x(x-a)}{2}\right]_a^{(a+b)/2} - \alpha.\left[\frac{x(x-b)}{2}\right]_{(a+b)/2}^b \\
&= \frac{\alpha\ (2\,a+b)\,(a-b)^2}{24} + \frac{\alpha\ (a+2\,b)\,(a-b)^2}{24} \\
&= \frac{\alpha\ (a+b)\,(a-b)^2}{8} = (a+b)/2 \\
mt^2{}_{\mathcal{T}([a,b])} &= \int_{\mathbb{R}} Id^2.f = \int_a^m Id^2.f + \int_m^b Id^2.f \\
&= \alpha.\left[\frac{x^2(x-a)}{2}\right]_a^{(a+b)/2} - \alpha.\left[\frac{x^2(x-b)}{2}\right]_{(a+b)/2}^b \\
&= \frac{\alpha\ (11\,a^2+10\,ab+3\,b^2)\,(a-b)^2}{192} \\
&\quad + \frac{\alpha\ (3\,a^2+10\,ab+11\,b^2)\,(a-b)^2}{192} \\
&= \frac{\alpha\ (7\,a^2+10\,ab+7\,b^2)\,(a-b)^2}{96} \\
&= (7a^2+10ab+7b^2)/24 \\
V_{\mathcal{T}([a,b])} &= (7a^2+10ab+7b^2)/24 - ((a+b)/2))^2 \\
&= (a^2-2ab+b^2)/24 = (a-b)^2/24.
\end{aligned}
$$

• *Calcul de m et σ pour la loi Exponentielle*

Vérifions que f est une densité.

$$
\begin{aligned}
\int_{\mathbb{R}_+} f &= \int_0^{+\infty} \alpha.e^{-\alpha x}.dx \\
&= \alpha.\left[\frac{e^{-\alpha x}}{-\alpha}\right]_0^{+\infty} = -1.(0-1) = 1
\end{aligned}
$$

$$
\begin{aligned}
m_{\ \mathcal{E}(\alpha)} &= \int_{\mathbb{R}_+} Id.f = \int_0^{+\infty} x.\alpha.e^{-\alpha x}.dx \\
&\text{en intégrant par parties avec } u = x,\ v = e^{-\alpha x}/-\alpha \\
&= \alpha.\left[\frac{x.e^{-\alpha x}}{-\alpha}\right]_0^{+\infty} - \int_0^{+\infty} -1.e^{-\alpha x}.dx \\
&= 0 - \left[\frac{e^{-\alpha x}}{-\alpha}\right]_0^{+\infty} = 1/\alpha \\
mt^2{}_{\mathcal{E}(\alpha)} &= \int_0^{+\infty} x^2.\alpha.e^{-\alpha x}.dx \\
&\text{en intégrant par parties avec } u = x^2,\ v = e^{-\alpha x}/-\alpha \\
&= \alpha.\left[\frac{x^2.e^{-\alpha x}}{-\alpha}\right]_0^{+\infty} - \int_0^{+\infty} -2.x.e^{-\alpha x}.dx \\
&= 0 - 2.m/\alpha = 2/\alpha^2 \\
V_{\ \mathcal{E}(\alpha)} &= 2/\alpha^2 - (1/\alpha)^2 = 1/\alpha^2 \ \diamondsuit
\end{aligned}
$$

• *Calcul de m et σ pour la loi Gamma*

Vérifions que f est une densité.

$$
\begin{aligned}
\int_{\mathbb{R}_+} f &= \int_0^{+\infty} \alpha^k.x^{k-1}.e^{-\alpha x}/\Gamma(k).dx \\
&\text{avec le changement de variable } y = \alpha x \\
&= \int_0^{+\infty} \alpha^k.(y/\alpha)^{k-1}.e^{-y}/\Gamma(k).d(y/\alpha) \\
&= \frac{\alpha^k}{\alpha^{k-1}.\Gamma(k)}.\int_0^{+\infty} y^{k-1}.e^{-y}.dy/\alpha \\
&= = \Gamma(k)/\Gamma(k) = 1
\end{aligned}
$$

$$\begin{aligned}
m_{\ \gamma(\alpha,k)} &= \int_{\mathbb{R}_+} Id.f = x.\alpha^k.x^{k-1}.e^{-\alpha x}/\Gamma(k).dx \\
&\quad \text{en intégrant par parties avec } u = x^k,\ dv = e^{-\alpha x} \\
&= \frac{\alpha^k}{\Gamma(k)}\left[\frac{e^{-\alpha x}}{-\alpha}x^k\right]_0^{+\infty} + \frac{k}{\alpha}.\int_{\mathbb{R}_+} f \\
&= 0 + (k/\alpha).1 = k\alpha
\end{aligned}$$

$$\begin{aligned}
mt^2{}_{\gamma(\alpha,k)} &= \int_{\mathbb{R}_+} Id^2.f = x^2.\alpha^k.x^{k-1}.e^{-\alpha x}/\Gamma(k).dx \\
&\quad \text{en intégrant par parties avec } u = x^{k+1},\ dv = e^{-\alpha x} \\
&= \frac{\alpha^k}{\Gamma(k)}\left[\frac{e^{-\alpha x}}{-\alpha}x^{k+1}\right]_0^{+\infty} + \frac{k+1}{\alpha}.\int_{\mathbb{R}_+} Id.f \\
&= 0 + ((k+1)/\alpha).m = k.(k+1)/\alpha^2 \\
V_{\ \gamma(\alpha,k)} &= k.(k+1)/\alpha^2 - (k/\alpha)^2 = k/\alpha^2
\end{aligned}$$

- *Calcul de m et σ pour la loi de Weibull*

Vérifions que f est une densité.

$$\begin{aligned}
\int_{\gamma}^{+\infty} f &= \int_{\gamma}^{+\infty} \frac{\beta}{\alpha}.\frac{x-\gamma}{\alpha}^{\beta-1} e^{-\frac{x-\gamma}{\alpha}^{\beta}}.dx \\
&\quad \text{avec le changement de variable } u = \tfrac{x-\gamma}{\alpha} \\
&= \int_0^{+\infty} \beta.u^{\beta-1}e^{-u^\beta} \\
&\quad \text{effectuons un second changement de variable } v = e^{-u^\beta} \\
&= \int_0^{+\infty} -d(\ e^{-v}\ ) \\
&= \left[-e^{-u^\beta}\right]_0^{+\infty} = -(1-0) = 1
\end{aligned}$$

Nous aurons besoin du résultat suivant :

$$
\begin{aligned}
I_n &= \int_0^{+\infty} u^n.\beta.e^{-u^\beta}.du \text{ avec le changement de variable } v = u^\beta \\
&= \int_0^{+\infty} v^{n/beta}.e^{-v}.dv = \Gamma(\frac{n}{\beta}+1) \\
m_{\ \mathcal{W}(\alpha,\beta,\gamma)} &= \int_0^{+\infty} Id.f = \int_0^{+\infty} x.\frac{\beta}{\alpha}.\frac{x-\gamma}{\alpha}^{\beta-1}.e^{-\frac{x-\gamma}{\alpha}^\beta}.dx \\
&\quad \text{avec le changement de variable } u = \tfrac{x-\gamma}{\alpha} \\
&= \int_0^{+\infty} (\alpha.u+\gamma).\beta.u^{\beta-1}.e^{-u^\beta}.du \\
&= \alpha.I_1 + \gamma.1 = \alpha.\Gamma(1+1/\beta)+\gamma \\
mt^2{}_{\mathcal{W}(\alpha,\beta,\gamma)} &= \int_0^{+\infty} Id^2.f = x^2.\frac{\beta}{\alpha}.\frac{x-\gamma}{\alpha}^{\beta-1} e^{-\frac{x-\gamma}{\alpha}^\beta}.dx \\
&= \int_0^{+\infty} (\alpha.u+\gamma)^2.\beta.u^{\beta-1}.e^{-u^\beta}.du \\
&= \alpha^2.I_2 + 2.\alpha.\gamma.I_1 + \gamma^2.1 \\
&= \alpha^2.\Gamma(2+1/\beta) + 2.\alpha.\gamma.\Gamma(1+1/\beta) + \gamma^2 \\
V_{\ \mathcal{W}(\alpha,\beta,\gamma)} &= \alpha^2.\Gamma(1+1/\beta) + 2.\alpha.\gamma.\Gamma(1+1/\beta) + \gamma^2 - (\alpha.\Gamma(1+1/\beta)+\gamma)^2 \\
&= \alpha.^2(\ \Gamma(1+2/\beta) - \Gamma(1+1/\beta)^2\ ).
\end{aligned}
$$

• *Calcul de m et σ pour la loi Normale*

Commençons par établir quelques résultats annexes.

$$
\begin{aligned}
I &= \int_0^{+\infty}\int_0^{+\infty} e^{-(x^2+y^2)}.dx.dy \\
&\quad \text{avec le changement de variable } x=\rho.cos(\theta),\ y=\rho.sin(\theta) \\
&= \int_0^{+\pi/2}\int_0^{+\infty} e^{-\rho^2}\rho.d\rho.d\theta \\
&= \int_0^{+\pi/2} d\theta.\int_0^{+\infty} \rho.e^{-\rho^2}.d\rho \\
&= [\theta]_0^{\theta/2}.\left[\frac{-1}{2}e^{-r^2}\right]_0^{+\infty} \\
&= (\pi/2).(\ 0-(-1/2)\ )\ =\ \pi/4 \\
J &= \int_0^{+\infty}\int_0^{+\infty} x^2y^2e^{-(x^2+y^2)}.dx.dy \\
&\quad \text{avec le changement de variable } x=\rho.cos(\theta),\ y=\rho.sin(\theta) \\
&= \int_0^{+\pi/2} sin^2(\theta).cos^2(\theta).d\theta.\int_0^{+\infty} \rho^5.e^{-\rho^2}.d\rho \\
&\quad \text{effectuons un second changement de variable } u=\rho^2 \\
&= \int_0^{+\pi/2} \frac{1-cos(4\theta)}{8}.d\theta.\frac{1}{2}\int_0^{+\infty} u^2.e^{-u^2}.du \\
&= \frac{1}{8}\left[\theta-\frac{sin(4\theta)}{4}\right].\frac{1}{2}\Gamma(3) \\
&= (\pi/16).2/2\ =\ \pi/16
\end{aligned}
$$

Montrons que f est bien une densité.

$$I = \int_{-\infty}^{+\infty} \frac{1}{\sigma.\sqrt{2\pi}}.e^{-(\frac{x-\mu}{\sigma})^2/2}.dx$$

avec le changement de variable $u = (x-\mu)/\sigma$

$$= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}}.e^{-u^2/2}.du$$

effectuant un second changement de variable $v = u/\sqrt{2}$

$$= \frac{1}{\sqrt{\pi}}.2\int_{0}^{+\infty} e^{-v^2}.dv$$

$$= \frac{2}{\sqrt{\pi}}.\sqrt{I} = 1$$

$$m_{\mathcal{N}(m,\sigma)} = \int_{-\infty}^{+\infty} Id.f = \int_{-\infty}^{+\infty} \frac{1}{\sigma.\sqrt{2\pi}}.x.e^{-(\frac{x-\mu}{\sigma})^2/2}.dx$$

avec le changement de variable $u = (x-\mu)/\sigma$

$$= \frac{\sigma}{\sqrt{\pi}}.\int_{-\infty}^{+\infty} u.e^{-u^2/2}.du + \mu.\int_{-\infty}^{+\infty} f$$

$$= 0 + \mu.1 = \mu$$

$$\begin{aligned}
mt^2{}_{\mathcal{N}(m,\sigma)} &= \int_{-\infty}^{+\infty} Id^2.f = \int_{-\infty}^{+\infty} \frac{1}{\sigma.\sqrt{2\pi}}.x^2.e^{-(\frac{x-\mu}{\sigma})^2/2}.dx \\
&\quad \text{avec les mêmes changements de variable qu'au-dessus} \\
&= \frac{\sigma^2}{\sqrt{2\pi}}.\int_{-\infty}^{+\infty} u^2.e^{-u^2/2}.du \ + \frac{2\sigma\mu}{\sqrt{2\pi}}.\int_{-\infty}^{+\infty} u.e^{-u^2/2}.du \ + \mu^2.\int_{-\infty}^{+\infty} f \\
&= \frac{\sigma^2}{\sqrt{2\pi}}.2\int_{-\infty}^{+\infty} u^2.e^{-u^2/2}.du \ + 0 + \mu^2.1 \\
&= \frac{\sigma^2}{\sqrt{\pi}}.4.\int_{0}^{+\infty} v^2.e^{-v^2}.dv \ + \mu^2 \\
&= \frac{\sigma^2}{\sqrt{\pi}}.4.\sqrt{J} + \mu^2 \ = \ \sigma^2 + \mu^2 \\
V_{\mathcal{N}(m,\sigma)} &= \sigma^2 + \mu^2 - (\mu)^2 \ = \ \sigma^2
\end{aligned}$$