

MTH TC3 -PROBABILITÉS

ECOLE CENTRALE DE LYON - PREMIÈRE ANNÉE

2017-2018

Table des matières

1	Variables aléatoires réelles	1
1.1	Modèle probabiliste	1
1.2	Lois à densité	2
1.3	Lois usuelles	2
1.3.1	Loi uniforme.	2
1.3.2	Loi exponentielle, loi d'Erlang, loi Gamma	3
1.3.3	Loi normale ou gaussienne.	4
1.4	Propriétés qualitatives d'une variable aléatoire	5
1.5	Fonction de répartition d'une variable aléatoire réelle	6
2	Moments de variables aléatoires réelles	7
2.1	Définition de l'espérance	8
2.2	Propriétés de l'espérance	9
2.3	Formule de transfert-Variance.	9
2.3.1	Formule de transfert	9
2.3.2	Variance et écart-type d'une variable aléatoire réelle.	10
2.4	Changement de variables.	11
2.4.1	Changement de variable par la fonction de répartition.	11
2.4.2	Changement de variable pour les v.a. à densité.	12
3	Vecteurs aléatoires	13
3.1	Définition	13
3.2	Lois à densité.	13
3.2.1	Définition.	13
3.2.2	Loi uniforme sur un ensemble de mesure bornée.	14
3.3	Lois marginales.	14
3.4	Moments de vecteurs aléatoires.	15
3.4.1	Formule de transfert	15
3.4.2	Covariance de deux variables réelles.	16
3.4.3	Espérance d'un vecteur aléatoire.	17
3.4.4	Covariance de deux vecteurs aléatoires.	18
3.4.5	Variance (ou covariance) d'un vecteur aléatoire.	18
3.4.6	Espérance et variance sous transformation affine d'un vecteur aléatoire	19
3.5	Variables aléatoires indépendantes.	19
3.6	Somme de v.a. indépendantes.	21
3.7	Vecteurs gaussiens	23
3.7.1	Propriétés remarquables	23
3.7.2	Lois du chi-deux, de Student, de Fisher-Snedecor	25

4	Suites de variables aléatoires - Résultats asymptotiques	26
4.1	Un exemple important.	26
4.2	Différents types de convergence.	26
4.3	Théorèmes de convergence.	27
4.3.1	Les lois des grands nombres	27
4.3.2	Théorème limite central	28
4.4	Simulation de variables aléatoires	30
4.4.1	Simulation d'une v.a. admettant une densité continue par inversion de la fonction de répartition	31
4.4.2	Simulation d'une v.a. bornée à densité bornée par la méthode de rejet	31
4.4.3	Simulation d'une loi normale	32
4.5	Méthode de Monte Carlo pour le calcul des intégrales et des moments.	32
4.5.1	Méthode	32
4.5.2	Précision (estimation d'erreur, intervalle de confiance)	33
4.5.3	Intérêts de la méthode	34
4.5.4	Calcul approché des moments	34
4.6	Méthode de "simulation Monte Carlo".	34
4.7	Rééchantillonnage (bootstrap)	35
5	Enoncés de TD	36
A	Annexe : Eléments de théorie des ensembles et de théorie de la mesure	41
A.1	Ensembles dénombrables	41
A.2	Ensembles négligeables. Intégrale des fonctions nulles presque partout.	41
A.3	Mesure de Lebesgue des ensembles de \mathbb{R}^n	42
B	Formulaire : Lois à densité usuelles	43
B.1	Lois, espérances, variances.	43
B.2	Propriétés.	46
C	Rappel : Variables Aléatoires Discrètes	47
C.1	Définition et Propriétés	47
C.1.1	Espérance	47
C.1.2	Formule de transfert	47
C.1.3	Variance	48
C.1.4	Somme de v.a discrètes indépendantes	48
C.2	Lois usuelles discrètes.	48
C.2.1	Lois, espérances, variances.	48
C.2.2	Propriétés.	52
C.3	Simulation de variables aléatoires discrètes	52
C.3.1	Simulation d'une v.a. de Bernoulli	52
C.3.2	Simulation d'une v.a. binomiale	53
C.3.3	Simulation d'une v.a. discrète	53

Théorie des probabilités - Le modèle continu

Les notions vues en classes préparatoires (à peu près tout ce qui concerne le modèle discret) ne seront pas refaites en cours et sont absentes de ce poly. Elles sont considérées comme acquises. Si nécessaire, on pourra consulter par exemple les références suivantes :

Paul Pichereau : PROBABILITÉS POUR LA PRÉPA. HK Editions 2013.

Probabilités et Statistique 2015. Poly de Centrale. Disponible sur Pédagogie.

Pour compléter le présent poly, sur le programme de la première année de Centrale Lyon, on pourra consulter au choix par exemple

- . un ouvrage destiné aux ingénieurs mais avec une base mathématique solide, avec une partie supplémentaire sur l'analyse des données :
Gilbert Saporta : PROBABILITÉS, ANALYSE DES DONNÉES ET STATISTIQUE. Editions Technip 2011.
- . un ouvrage américain très peu mathématique, avec beaucoup d'exemples :
Douglas C. Montgomery et George C. Runger : APPLIED STATISTICS AND PROBABILITY FOR ENGINEERS. Wiley Edition 2013.
- . un ouvrage plus mathématique :
Benjamin Jourdain : PROBABILITÉS ET STATISTIQUE. Ellipses 2009.

L'objectif de secours est de compléter le programme de classe préparatoire en étudiant les phénomènes aléatoires qui sont plus naturellement représentés par des variables ou des vecteurs aléatoires prenant un continuum de valeurs plutôt qu'un ensemble discret (dénombrable) de valeurs. En outre, on verra que ces lois apparaissent naturellement de façon asymptotique. Enfin, on pourra appliquer les théorèmes limites à des fins de simulation aléatoire et de calculs approchés.

1 Variables aléatoires réelles

1.1 Modèle probabiliste

Dans l'ensemble du cours, le modèle probabiliste sera décrit par l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, qui est un triplet composé de :

- . Ω l'univers des états du monde, un ensemble non vide.
- . \mathcal{F} une tribu sur Ω (un ensemble de parties de Ω contenant Ω et stable par intersection dénombrable et passage au complémentaire).
- . \mathbb{P} une probabilité sur Ω .

Les ensembles de \mathcal{F} sont des sous-ensembles de Ω appelés *événements* et leur probabilité d'occurrence est donnée à travers la probabilité (la fonction)

$$\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$$

qui satisfait $\mathbb{P}[\Omega] = 1$ et la propriété de sigma-additivité

$$\mathbb{P} \left[\bigcup_{n \in \mathbb{N}} A_n \right] = \sum_{n \in \mathbb{N}} \mathbb{P}[A_n]$$

pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'événements disjoints.

1.2 Lois à densité

Une *variable aléatoire (réelle)* (v.a.r. en abrégé) est une application

$$X : \Omega \rightarrow \mathbb{R}$$

qui associe à un état du monde $\omega \in \Omega$ le résultat d'une expérience aléatoire $X(\omega)$. Sa loi est la donnée des probabilités $\mathbb{P}[X \in B] = \mathbb{P}[X^{-1}(B)]$ pour $B \subset \mathbb{R}^*$, où $X^{-1}(B) = \{\omega \in \Omega / X(\omega) \in B\}$.

Les *variables aléatoires à densité* modélisent des expériences dont le résultat est quantitatif et dont les valeurs prises sont réparties « suffisamment régulièrement » sur \mathbb{R} . Par exemple : taille d'un composant produit par une machine, durée de vie d'une particule, tirage d'un point au hasard dans $[0, 1]$.

Définition 1.1

Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction positive d'intégrale 1. On dit qu'une variable aléatoire X a pour densité f si sa loi est définie pour tout $B \subset \mathbb{R}$ par

$$\mathbb{P}[X \in B] = \int_B f(x)dx.$$

Remarque : Si X est une variable réelle de densité f alors, pour $a < b$ réels,

$$\mathbb{P}[a < X < b] = \mathbb{P}[X \in]a, b[] = \int_{]a, b[} f(x)dx = \int_a^b f(x)dx.$$

1.3 Lois usuelles

Les lois à densité sont en général moins faciles à appréhender que les lois discrètes. Elles apparaissent naturellement par passage à la limite dans des lois discrètes, par changement de variable à partir d'une loi uniforme, ou « par expérience » dans les séries statistiques. Dans la suite, nous proposons quelques exemples usuels. L'ensemble des lois classiques à connaître se trouve en annexe.

1.3.1 Loi uniforme.

Remarque : Dans le cas discret, la loi uniforme donne une probabilité de réalisation d'un événement égale à la proportion de résultats de l'événement donné (rapport des cardinaux). Quand l'ensemble des résultats est infini dénombrable, il n'est pas possible d'avoir de l'uniformité des événements élémentaires. En revanche, dans le cas continu, c'est possible lorsque l'on peut faire un rapport de mesure (longueurs, aires, volumes, etc.)

Une variable aléatoire X est dite uniforme si la probabilité qu'elle prenne ses valeurs dans un ensemble donné est proportionnelle à la mesure de cet ensemble, c'est donc l'intégrale d'une constante sur l'ensemble. L'ensemble des valeurs prises par X est bien infini mais, de même qu'en discret, la mesure de l'ensemble des valeurs prises doit rester bornée (c'est le cardinal, la mesure de comptage, dans le cas discret).

Pour un intervalle $[a, b]$ la longueur $b - a$ est aussi l'intégrale sur l'intervalle. Cela implique que la densité d'une variable aléatoire uniforme sur $[a, b]$ est la constante $(b - a)^{-1}$.

Plus précisément : on dit que X suit la loi uniforme sur l'ensemble borné et de mesure finie non nulle B_0 si elle admet pour densité la fonction

$$f(x) = \begin{cases} 1/\lambda(B_0) & \text{si } x \in B_0 \\ 0 & \text{sinon} \end{cases}$$

*. Pour être tout à fait rigoureux, cette loi est définie pour les ensembles B dits boréliens qui constituent la plus petite tribu contenant les intervalles de \mathbb{R} , qui est également la plus petite tribu contenant les ouverts de \mathbb{R} . Par définition, si X est une variable aléatoire définie sur $(\Omega, \mathcal{F}, \mathbb{P})$ alors $X^{-1}(B) \in \mathcal{F}$ pour tout borélien de \mathbb{R} et $\mathbb{P}[X^{-1}(B)]$ est ainsi bien défini.

où $\lambda(B_0)$ désigne la mesure de B_0 c'est-à-dire la quantité

$$\lambda(B_0) = \int_{B_0} dx.$$

La densité de X est donc l'application f définie sur \mathbb{R} par

$$f = \frac{1}{\lambda(B_0)} \mathbf{1}_{B_0}$$

avec

$$\mathbf{1}_{B_0}(x) = \begin{cases} 1 & \text{si } x \in B_0 \\ 0 & \text{sinon} \end{cases}$$

Si X suit une loi uniforme sur B_0 , on note $X \sim \mathcal{U}(B_0)$ et, d'après la définition de la loi de X , on a, pour tout $B \subset \mathbb{R}$,

$$\mathbb{P}[X \in B] =$$

$$=$$

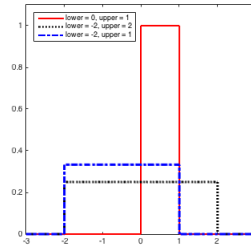


FIGURE 1 — Densité de lois uniformes dans \mathbb{R}

1.3.2 Loi exponentielle, loi d'Erlang, loi Gamma

La loi exponentielle est le pendant continu de la loi géométrique : c'est la seule loi à densité continue sans mémoire, c'est-à-dire satisfaisant

$$\mathbb{P}[X > s + t | X > t] = \mathbb{P}[X > s] \text{ pour tous } t, s > 0. \quad (1)$$

On peut montrer que la densité d'une loi sans mémoire à densité continue est nécessairement définie par

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$$

où $\lambda > 0$ est un paramètre. On dit qu'une v.a. admettant cette fonction pour densité est une loi exponentielle de paramètre λ que l'on note $\mathcal{E}(\lambda)$. Une telle variable aléatoire satisfait la propriété (1).

On peut également voir la loi exponentielle comme celle de l'espacement (temporel ou spatial) entre deux "événements poissonniens". En effet, la loi de Poisson de paramètre λx modélise le nombre d'occurrences N_x d'événements identiques, indépendants et non simultanés dans un intervalle (de temps) de longueur x , le paramètre λ étant le nombre

moyen d'occurrences dans l'intervalle de longueur 1. Si on note X l'écart entre deux occurrences que l'on suppose indépendant du nombre d'occurrences, alors la probabilité que ce temps dépasse x est la probabilité qu'il n'y ait aucune occurrence sur un intervalle de temps x , soit

$$\mathbb{P}[X > x] =$$

De l'égalité

on tire en dérivant

qui est le résultat attendu.

La loi d'Erlang généralise la loi précédente en ce sens qu'elle est celle du temps (ou de l'espacement) jusqu'à l'occurrence du second, troisième, ou plus généralement k -ième événement poissonnien.

Pour $k \in \mathbb{N}^*$ et $\lambda > 0$, la densité d'une v.a. d'Erlang de paramètres λ , k est la fonction définie sur \mathbb{R} par

$$f(x) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!} \mathbf{1}_{\mathbb{R}_+}(x), \text{ pour } x \in \mathbb{R}.$$

La loi d'Erlang est un cas particulier de la loi Gamma qui prend comme paramètres le paramètre $\lambda > 0$ et un paramètre $\alpha > 0$ réel en lieu et place de k et dont la densité est

$$f(x) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} \mathbf{1}_{\mathbb{R}_+}(x), \text{ pour } x \in \mathbb{R},$$

où

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx.$$

On vérifie aisément que

$$\Gamma(\alpha + 1) = \alpha \Gamma(\alpha) \text{ pour tout } \alpha > 0;$$

et donc que

$$\Gamma(k) = (k-1)! \text{ pour tout } k \in \mathbb{N}^*.$$

1.3.3 Loi normale ou gaussienne.

Cette loi apparaît comme loi limite de la moyenne d'une infinité de résultats d'expériences identiques et indépendantes. Elle est également utilisée très fréquemment en statistique. Sa densité est strictement positive sur \mathbb{R} , symétrique, et présente une forme caractéristique dite « en cloche ».

On dit que X suit une loi normale de paramètres $m \in \mathbb{R}$ et $\sigma^2 > 0$ et on note $X \sim \mathcal{N}(m, \sigma^2)$ si X admet pour densité la fonction définie par

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(x-m)^2}{\sigma^2}\right) \text{ pour tout } x \in \mathbb{R}.$$

On note en particulier que

$$\int_{\mathbb{R}} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(x-m)^2}{\sigma^2}\right) dx = 1.$$

Les variables jouent un rôle tout particulier en théorie des probabilités et en statistique. Elles vérifient des propriétés exceptionnelles que nous explorerons ultérieurement.

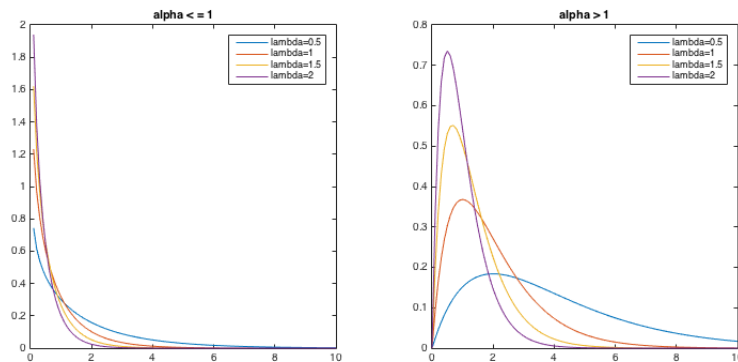


FIGURE 2 — Densités de lois Gamma pour différentes valeurs de α et λ

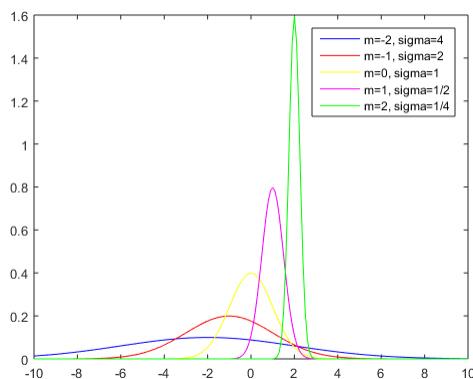


FIGURE 3 — Densités de lois gaussiennes 1D

1.4 Propriétés qualitatives d'une variable aléatoire

La représentation de l'ensemble des valeurs prises par une variable aléatoire, que l'on appelle le support[†] de sa loi, est une information pertinente lorsque l'on souhaite modéliser un problème.

Une variable aléatoire réelle positive a son support inclus dans \mathbb{R}^+ . Une variable aléatoire bornée a son support inclus dans une « boule ». Les variables dont les lois présentent des symétries ont des supports qui présentent les mêmes symétries. Par exemple, les variables X telles que $\mathbb{P}[-X \in B] = \mathbb{P}[X \in B]$ pour tout $B \subset \mathbb{R}$, ont un support symétrique par rapport à l'origine (on parle alors de lois symétriques).

Remarque : Support et densité

1. Toutes les variables n'admettent pas de densité.

Par exemple une loi discrète ne peut en admettre. En effet, si D est l'ensemble dénombrable (= "dont on peut compter les éléments", voir l'annexe A pour une définition précise et des propriétés utiles) des valeurs qu'elle prend alors

[†]. Plus précisément, on définit le support d'une loi d'une variable aléatoire qui prend ses valeurs dans \mathbb{R} comme l'intersection D_0 de tous les fermés D de \mathbb{R} tels que $\mathbb{P}[X \in D] = 1$. Dans les cas non pathologiques qui sont les seuls qui nous concernent dans ce cours, on a également $\mathbb{P}[X \in D_0] = 1$. Par exemple, le support d'une loi de v.a. discrète est l'ensemble des valeurs prises par la variables aléatoire. Pour une variable aléatoire admettant une densité continue, c'est l'adhérence de $\{x \in \mathbb{R} / f(x) > 0\}$.

$\mathbb{P}[X \in D] = 1$ mais pourtant, pour toute fonction positive

$$\int_D f(x)dx = 0.$$

Parfois, des modélisations différentes d'un même problème peuvent conduire à des variables avec ou sans densité. Par exemple, une variable prenant uniformément ses valeurs sur le cercle a un support de mesure (d'aire) nulle dans \mathbb{R}^2 et ne peut donc pas admettre de densité sur \mathbb{R}^2 car l'intégrale, sur le cercle, d'une fonction positive sur \mathbb{R}^2 est nulle. En revanche, on peut choisir un modèle repérant un point sur le cercle par son abscisse curviligne, représentation qui a une densité en dimension 1.

2. La valeur d'une densité en un point n'importe pas : une densité peut-être modifiée sur un ensemble dénombrable de points sans que son intégrale change. En fait elle pourra même être modifiée sur un ensemble sans changer les calculs de probabilité du moment qu'il reste « négligeable » (voir l'annexe A).
3. Ces propriétés viennent du fait que toute loi de probabilité peut se décomposer comme somme de deux mesures dont l'une admet une densité et l'autre prend ses valeurs dans un ensemble négligeable (et donc n'admet pas de densité). On pourra voir ce point plus en détail en consultant l'annexe A du poly.

1.5 Fonction de répartition d'une variable aléatoire réelle

La fonction de répartition d'une variable aléatoire est une caractérisation de tous les types de variables aléatoires réelles.

Définition 1.2

Soit X une v.a. réelle sur Ω . On appelle fonction de répartition de X et on note F_X la fonction, définie pour tout $t \in \mathbb{R}$ par

$$F_X(t) = \mathbb{P}[X \leq t] = \mathbb{P}[\{\omega \in \Omega / X(\omega) \leq t\}].$$

Proposition 1.1

Soit F_X la fonction de répartition d'une v.a.r. X définie sur Ω . Alors, pour tous $a, b \in \mathbb{R}$ tels que $a < b$, on a :

$$\mathbb{P}[a < X \leq b] = F_X(b) - F_X(a).$$

Démonstration :

Proposition 1.2

Soit F_X la fonction de répartition d'une v.a.r. X définie sur Ω . Alors F_X est croissante, continue à droite et satisfait

$$\lim_{t \rightarrow -\infty} F_X(t) = 0, \quad \lim_{t \rightarrow +\infty} F_X(t) = 1.$$

Remarque : La loi d'une v.a.r. X est entièrement déterminée par la fonction de répartition F_X de X . Plus précisément, pour tout $t \in \mathbb{R}$,

$$F_X(t) = \mathbb{P}[X \in]-\infty, t]]$$

et la donnée de ces probabilités suffit pour caractériser la loi de X .

Exemple : Soit X une v.a. réelle. Si X est une v.a.r. admettant une densité f alors

$$F_X(t) = \int_{-\infty}^t f(x)dx \text{ pour tout } t \in \mathbb{R}.$$

En effet,

Par exemple, si X suit une loi uniforme sur $[0, a]$, où $a > 0$:

En particulier, dans le cas à densité, on a le résultat suivant :

Proposition 1.3

Soit X une v.a.r. de densité f et de fonction de répartition F . Alors F est continue sur \mathbb{R} et si f est continue en $x \in \mathbb{R}$, alors F est dérivable en x et

$$F'(x) = f(x).$$

Réciproquement, si X admet une fonction F continue et dérivable par morceaux, alors X admet pour densité $f = F'$.

Définition 1.3

Soit X une v.a. réelle de fonction de répartition F_X . On appelle quantile d'ordre α le réel q_α défini par

$$q_\alpha = F_X^{-1}(\alpha)$$

Représentation des quantiles

2 Moments de variables aléatoires réelles

Les moments sont, comme la moyenne et la variance, des données de synthèse qui donnent des indication de tendance sur une variable aléatoire réelle. En outre, on peut caractériser la loi d'une variable aléatoire à travers l'espérance

de fonctions test de cette variable.

2.1 Définition de l'espérance

Rappel : espérance dans le cas discret

Pour les variables à densité, l'idée pour déterminer l'espérance est la suivante : la valeur moyenne de la variable aléatoire, doit être la somme des valeurs prises par la variable pondérées par la densité de probabilité de la variable. Cela conduit à la définition suivante :

Définition 2.1

Soit X est une variable aléatoire admettant pour densité f . Alors si

$$\int_{\mathbb{R}} |x|f(x)dx < +\infty$$

on dit que X est intégrable et on définit l'espérance de X par

$$\mathbb{E}[X] = \int_{\mathbb{R}} xf(x)dx.$$

L'espérance (ou la *moyenne*) d'une variable aléatoire est une caractéristique de tendance centrale de la variable. On l'appelle également le *moment d'ordre 1* de la variable aléatoire.

Exemple : [Calcul d'espérance dans le cas à densité.]

Soit X une v.a. suivant la loi exponentielle de paramètre $\lambda > 0$. On rappelle qu'une telle variable admet pour densité la fonction

$$x \mapsto f(x) = \lambda e^{-\lambda x} \mathbf{1}_{\mathbb{R}_+}(x).$$

Comme X est une v.a. positive, elle est intégrable si son intégrale est finie. L'intégrale à calculer est

Cette intégrale est finie et se calcule par intégration par parties. Donc X est intégrable et

$$\mathbb{E}[X] =$$

Finalement, X est intégrable et

$$\mathbb{E}[X] =$$

L'espérance des lois usuelles se trouve dans le formulaire en fin de polycopié. Le calcul de ces quantités constitue autant d'exercices d'entraînement. On notera en particulier

Proposition 2.1

Soit X une v.a. de loi $\mathcal{N}(m, \sigma^2)$. Alors X est intégrable et son espérance vaut m .

Démonstration :

X a pour densité la fonction f définie sur \mathbb{R} par

On a alors :

2.2 Propriétés de l'espérance

L'espérance est d'une variable aléatoire X est en réalité une intégrale en un sens que l'on ne veut pas préciser ici. A ce titre, elle vérifie un certain nombre de propriétés que nous énonçons sans démonstration.

Proposition 2.2

Soient X et Y des v.a.r. intégrables définies sur Ω . Alors,

1. pour tous $a, b \in \mathbb{R}$, $aX + b$ est intégrable et $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$ et $\mathbb{E}[b] = b$;
2. $X + Y$ est intégrable et $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$;
3. si X est positive sur Ω alors $\mathbb{E}[X] \geq 0$.

Proposition 2.3

Si X est une variable aléatoire bornée alors X est intégrable et

$$\mathbb{E}[|X|] \leq \sup_{\omega \in \Omega} |X(\omega)|.$$

Proposition 2.4 (Inégalité de Markov (rappel).)

Soit X une v.a.r. intégrable. On a,

$$\mathbb{P}[|X| > \varepsilon] \leq \frac{\mathbb{E}[|X|]}{\varepsilon} \text{ pour tout } \varepsilon > 0.$$

2.3 Formule de transfert-Variance.

Il arrive que l'on veuille calculer non pas l'espérance d'une variable aléatoire X mais l'espérance de $\varphi(X)$. On verra dans cette section que l'on peut calculer l'espérance de $\varphi(X)$ en utilisant la loi de X . Ce résultat nous permettra de calculer la variance de X .

2.3.1 Formule de transfert

Les formules suivantes permettent de calculer l'espérance de fonctions de variables aléatoires.

Proposition 2.5 (Formule de transfert.)

Soit X une v.a. sur Ω et à valeurs dans \mathbb{R} . Soit $\varphi : \mathbb{R} \rightarrow \mathbb{R}$. Si X est une v.a. admettant pour densité f et si $\varphi(X)$

est intégrable, c'est-à-dire si

$$\int_{\mathbb{R}} |\varphi(x)| f(x) dx < +\infty \text{ alors } \mathbb{E}[\varphi(X)] = \int_{\mathbb{R}} \varphi(x) f(x) dx.$$

2.3.2 Variance et écart-type d'une variable aléatoire réelle.

Définition 2.2

Soit X une v.a.r. définie sur Ω telle que X^2 soit intégrable. On appelle alors la variance de X le réel positif

$$\mathbf{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2].$$

La variance est une caractéristique de dispersion (on utilise aussi la racine de la variance, appelée l'écart-type de la variable). On appelle également la variance le *moment centré d'ordre 2* de la variable aléatoire réelle.

Remarque : Si X une v.a.r. définie sur Ω telle que X^2 soit intégrable, alors X est intégrable.

Grâce à la formule de transfert, on peut montrer le résultat suivant :

Proposition 2.6

Si X est une v.a.r. admettant une densité f et si

$$\int_{\mathbb{R}} x^2 f(x) dx < +\infty,$$

alors la variance de X est la quantité

$$\mathbf{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \int_{\mathbb{R}} (x - \mathbb{E}[X])^2 f(x) dx.$$

Proposition 2.7

Soit X une v.a.r. de carré intégrable.

1. $\mathbf{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.
2. Pour tous $a, b \in \mathbb{R}$, $aX + b$ est de carré intégrable et

$$\mathbf{Var}[aX + b] = a^2 \mathbf{Var}[X].$$

Exemple : [Calculs de variances.]

- loi exponentielle. On reprend l'exemple sur le calcul d'espérance ci-dessus : $X \rightsquigarrow \mathcal{E}(\lambda)$.

- loi normale. $X \rightsquigarrow \mathcal{N}(m, \sigma^2)$. Alors X a pour variance σ^2 .

Remarque : On a donc montré que les paramètres m et σ^2 de la loi normale sont son espérance et sa variance. En outre, si X suit une loi normale de paramètres m et σ^2 alors, comme on l'a vu dans la section sur les changements de variable, la variable aléatoire

$$X^* = \frac{X - m}{\sigma},$$

dite « centrée-réduite », suit une loi normale de paramètres

$$\frac{1}{\sigma}m - \frac{m}{\sigma} = 0 \quad \text{et} \quad \frac{1}{\sigma^2}\sigma^2 = 1.$$

Proposition 2.8 (Inégalité de Bienaymé-Tchebychev (rappel).)

Soit X une v.a.r. de carré intégrable. Alors,

$$\mathbb{P}[|X - \mathbb{E}[X]| > \varepsilon] \leq \frac{\text{Var}[X]}{\varepsilon^2} \text{ pour tout } \varepsilon > 0.$$

2.4 Changement de variables.

Il arrive qu'un modèle aléatoire utilise une variable aléatoire Y dont on ne connaît pas la loi mais qui s'exprime fonctionnellement à travers une autre variable X mieux connue. Nous venons de voir que la formule de transfert permet de calculer l'espérance de Y en utilisant la loi de X . Si l'on veut aller plus loin, à savoir connaître la loi de Y , il faut une méthode qui permet de calculer l'une des caractérisations de cette loi. Il n'existe pas de méthode générale mais des cas pratiques.

2.4.1 Changement de variable par la fonction de répartition.

Exemple : Soit X une variable aléatoire de fonction de répartition F_X . Soit ψ une application bijective croissante de \mathbb{R} dans \mathbb{R} et soit $Y = \psi(X)$. Calculer la fonction de répartition F_Y de Y .

Ce résultat n'est pas aisément généralisable alors que les calculs peuvent manifestement être poussés plus loin. Cela est fait ci-dessous à travers les densités.

2.4.2 Changement de variable pour les v.a. à densité.

Il existe une sorte de réciproque à la formule de transfert qui sera utile pour déterminer des lois inconnues dans le cas des variables à densité.

Proposition 2.9

Soit X une v.a. définie sur Ω et à valeurs dans \mathbb{R} . La loi de X est entièrement caractérisée par la donnée des

$$\mathbb{E}[\varphi(X)], \text{ pour tout } \varphi : \mathbb{R} \rightarrow \mathbb{R} \text{ bornée.}$$

Plus précisément, s'il existe une fonction f telle que pour toute fonction $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ bornée

$$\mathbb{E}[\varphi(X)] = \int_{\mathbb{R}} \varphi(x) f(x) dx,$$

alors f est une fonction positive d'intégrale 1 et X a pour densité f .

Remarque : Dans le résultat précédent, les applications φ sont bornées, ce qui garantit l'existence de l'espérance et rend inutile sa vérification.

Application : Utilisation de la formule de transfert pour un changement de variable. On connaît la loi d'une v.a. à densité et une seconde v.a. s'exprime, par l'intermédiaire d'une application, en fonction de la première. Comment calculer la loi de la seconde v.a. ? On donne ci-dessous un exemple pour présenter la méthode préconisée.

Exemple : On suppose que Θ suit une loi uniforme sur $]-\pi/2, \pi/2[$. On pose $X = \tan \Theta$. Quelle est la loi de X ?

Soit $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ une application bornée. Alors $\varphi(X)$ est intégrable et

$$\mathbb{E}[\varphi(X)] =$$

On applique la formule de transfert à $\varphi \circ \tan$. Il vient :

Puisque cette égalité est satisfaite pour toute application bornée, on en déduit que la v.a. X admet pour densité

Cette méthode nous permet de montrer le résultat suivant, qui reste assez limité. Dans tous les autres cas à densité, il conviendra de reprendre la méthode pas à pas pour procéder au changement de variable dans l'intégrale.

Proposition 2.10

Soit X une variable aléatoire de densité f . On pose $Y = \psi(X)$ où ψ est une fonction de \mathbb{R} sur \mathbb{R} (respectivement d'un intervalle ouvert I tel que $X(\Omega) \subset I$ sur $\psi(I)$), de classe C^1 , bijective et d'inverse de classe C^1 telle $\psi'(x) \neq 0, \forall x \in \mathbb{R}$ (resp. $\forall x \in I$). Alors Y admet pour densité la fonction g définie, pour tout $y \in \mathbb{R}$, par

$$g(y) = \frac{f(\psi^{-1}(y))}{|\psi'(\psi^{-1}(y))|}.$$

$$\left(\text{respectivement } g(y) = \frac{f(\psi^{-1}(y))}{|\psi'(\psi^{-1}(y))|} \mathbf{1}_{\psi(I)}(y). \right)$$

Proposition 2.11

Application Soit X une v.a. suivant une loi normale de paramètres m et σ^2 . Soient $a \neq 0$ et b des réels. Alors $Y = aX + b$ suit une loi normale de paramètres $am + b$ et $a^2\sigma^2$.

Démonstration : Comme $a \neq 0$, Y est obtenue par changement de variable bijectif sur \mathbb{R} et en appliquant la formule ci-dessus, puisque sur \mathbb{R} ,

$$\psi(x) = \quad \text{et} \quad \psi^{-1}(y) = \quad ,$$

sa densité g est définie, pour tout $y \in \mathbb{R}$, par

$$g(y) =$$

Avec

$$f(x) = \quad \text{pour tout } x \in \mathbb{R},$$

on obtient bien

$$g(y) = \quad \text{pour tout } y \in \mathbb{R}.$$

3 Vecteurs aléatoires

3.1 Définition

Dans cette partie, il s'agit de modéliser des phénomènes aléatoires prenant leurs valeurs dans \mathbb{R}^n , $n \geq 2$.

Exemple : Prise en compte d'une dimension physique, ici en continu : On lance une flèche sur une cible circulaire. On peut repérer l'impact de la flèche par un couple de réels, l'abscisse et l'ordonnée de l'impact dans un repère orthonormé par exemple.

Dans l'exemple ci-dessus, le résultat peut être considéré comme un couple de variables aléatoires réelles ou encore un vecteur aléatoire de dimension deux (X, Y) . Le but de ce chapitre sera de décrire la loi des vecteurs aléatoires et de comprendre les liens entre loi du vecteur et loi des composantes du vecteur.

Définition 3.1

Un vecteur aléatoire $X = (X_1, \dots, X_n)$ est une application

$$\begin{aligned} X : \Omega &\longrightarrow \mathbb{R}^n \\ \omega &\longmapsto X(\omega) = (X_1(\omega), \dots, X_n(\omega)). \end{aligned}$$

3.2 Lois à densité.

3.2.1 Définition.

Exemple : Si l'on considère que l'impact de la flèche est réparti uniformément sur la cible et que la cible peut être assimilée à un disque D de rayon 1 du plan, alors on dira que le couple (X, Y) admet comme densité la fonction

$$f = \quad ,$$

où $\lambda(D)$ est l'aire du disque D ($\lambda(D) = \pi$).

Définition 3.2

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction positive d'intégrale 1. On dit qu'un vecteur aléatoire $X = (X_1, \dots, X_n)$ admet pour densité f si sa loi est définie, pour tout $B \subset \mathbb{R}^n$ par

$$\mathbb{P}[X \in B] = \int_B f(x) dx = \int \dots \int_B f(x_1, \dots, x_n) dx_1 \dots dx_n.$$

La densité d'un vecteur permet de faire le calcul des probabilités d'événements mettant en jeu ses composantes.

Exemple : Soient X et Y deux variables aléatoires réelles. On suppose que le couple (X, Y) admet pour densité la fonction f et on cherche la probabilité que X dépasse Y . On cherche donc $\mathbb{P}[X > Y]$. On a

3.2.2 Loi uniforme sur un ensemble de mesure bornée.

Soit $B \subset \mathbb{R}^n$ un ensemble de mesure non nulle et bornée. On dit que X suit la loi uniforme sur B si elle admet pour densité la fonction f définie, pour tout $x \in \mathbb{R}^n$, par

$$f(x) = \frac{1}{\lambda_n(B)} \mathbf{1}_B(x),$$

où $\lambda_n(B)$ est la mesure de B définie par

$$\lambda_n(B) = \int \dots \int_B dx_1 \dots dx_n.$$

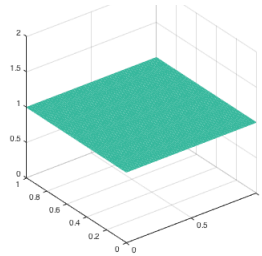


FIGURE 4 – Densité de loi uniforme dans \mathbb{R}^2

3.3 Lois marginales.

Les lois marginales sont les lois des composantes d'un vecteur aléatoire. Connaissant la loi d'un vecteur aléatoire, on peut déduire les lois marginales mais l'inverse est faux en général.

Exemple : [Cas à densité] On reprend l'exemple de la flèche lancée sur le disque D de centre 0 et de rayon 1. On suppose toujours que le point d'impact de la flèche est uniformément réparti sur la cible et donc que la densité de ce couple est $f = \mathbf{1}_D/\pi$. On cherche les lois de X et Y . Par symétrie, ces v.a. ont même loi. Pour avoir la loi de X , l'idée est de « sommer » sur les valeurs de Y . Plus précisément, la densité de X (et donc de Y) sera définie par

Graphiquement :

Généralisons ces résultats.

Proposition 3.1

Soit (X, Y) un couple de variables aléatoires réelles définies sur Ω .

Si (X, Y) admet pour densité $f_{(X,Y)} : \mathbb{R}^2 \rightarrow \mathbb{R}$, la densité de X est la fonction f_X définie par

$$f_X(x) = \int_{\mathbb{R}} f_{(X,Y)}(x, y) dy \text{ pour tout } x \in \mathbb{R}.$$

Remarque : Par symétrie, on obtient des résultats analogues pour la v.a.r. Y . Les lois de X et de Y , quand elles sont calculées à partir de la loi du couple (X, Y) , sont appelées des *lois marginales*.

Nous présentons ici le cas de la dimension n .

Proposition 3.2

Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire défini sur Ω .

Si X admet pour densité $f : \mathbb{R}^n \rightarrow \mathbb{R}$, la densité de X est la fonction f_X définie par

$$f_X(x) = \int_{\mathbb{R}^{n-1}} f(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n \text{ pour tout } x \in \mathbb{R}.$$

3.4 Moments de vecteurs aléatoires.

3.4.1 Formule de transfert

Pour pouvoir faire le calcul des moment de vecteurs aléatoires, il faudra savoir calculer l'espérance de fonctions réelles de vecteurs aléatoires. Ce résultat est donné par les formules de transfert.

Définition 3.3

Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire de dimension n et de densité f , et soit φ une application de \mathbb{R}^n dans

\mathbb{R} . On dit que la v.a.r. $\varphi(X)$ est intégrable si

$$\int_{\mathbb{R}^n} |\varphi(x)| f(x) dx < +\infty.$$

Proposition 3.3 (Formule de transfert)

Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire de dimension n et soit φ une application de \mathbb{R}^n dans \mathbb{R} . Si X admet une densité f et si $\varphi(X)$ est intégrable, alors

$$\mathbb{E}[\varphi(X)] = \int_{\mathbb{R}^n} \varphi(x) f(x) dx.$$

Cette formule sert comme en dimension 1 dans le calcul des probabilités ou pour les changements de variables.

On a également un théorème qui donne la densité sous changement de variable injectif comme en dimension 1.

Théorème 3.1

Soit X un vecteur aléatoire de densité f et à valeurs dans un ouvert $D \subset \mathbb{R}^d$. On pose $Y = \psi(X)$ avec ψ est une fonction bijective de classe C^1 de D sur $\psi(D) \subset \mathbb{R}^d$, d'inverse de classe C^1 telle que $\det J_\psi(x) \neq 0$, $x \in D$. Alors $\psi(X)$ est un vecteur aléatoire de \mathbb{R}^d de densité

$$g(y) = \frac{f(\psi^{-1}(y))}{|\det J_\psi(\psi^{-1}(y))|} \mathbf{1}_{\psi(D)}(y), \quad y \in \mathbb{R}^d,$$

J_ψ étant la matrice jacobienne de ψ .

(Par changement de variable dans la formule de transfert).

Exemple : [Test 2015-2016] Soient X et Y deux variables aléatoires uniformes sur $[0, 1] \times [0, 1]$. On pose U et V par

$$U = XY, \quad V = \frac{X}{Y}.$$

Déterminer la loi du couple (U, V) .

3.4.2 Covariance de deux variables réelles.

Définition 3.4

Soient X et Y deux v.a.r. de carré intégrable.

– la covariance de X et Y est définie par

$$\text{cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])].$$

On dit que X et Y sont décorrélées si $\text{cov}[X, Y] = 0$.

– la corrélation de X et Y est définie par

$$\text{corr}[X, Y] = \frac{\text{cov}[X, Y]}{\sqrt{\text{Var}[X]\text{Var}[Y]}}.$$

si $\text{Var}[X] \neq 0$ et $\text{Var}[Y] \neq 0$

Remarque : Pour calculer la covariance de deux variables aléatoires, on a donc besoin de l'espérance du produit des variables $X - \mathbb{E}[X]$ et $Y - \mathbb{E}[Y]$. Ce calcul, comme on le voit dans les formules de transfert ci-dessus, nécessite la loi jointe des variables aléatoires. La corrélation de deux variables aléatoires est adimensionnée et comprise entre -1 et $+1$.

Proposition 3.4

1. La covariance est bilinéaire, i.e. les applications $X \rightarrow \text{cov}[X, Y]$ et $Y \rightarrow \text{cov}[X, Y]$ sont linéaires. La covariance est symétrique.
2. Pour tout X de carré intégrable et pour réel a , $\text{cov}[a, X] = 0$.

Proposition 3.5

Soient X et Y deux v.a.r. de carré intégrable. Alors

1. $\text{cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$.
2. $X + Y$ est de carré intégrable et $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2\text{cov}[X, Y]$. En particulier, si X et Y sont décorrélées, alors $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$.

Proposition 3.6

Si X_1, \dots, X_n sont n v.a.r. alors

$$\text{Var}\left[\sum_{k=1}^n X_k\right] = \sum_{k=1}^n \text{Var}[X_k] + \sum_{i \neq j} \text{cov}[X_i, X_j]$$

et si les v.a.r. X_1, \dots, X_n sont deux à deux décorrélées alors

$$\text{Var}\left[\sum_{k=1}^n X_k\right] = \sum_{k=1}^n \text{Var}[X_k].$$

3.4.3 Espérance d'un vecteur aléatoire.

Définition 3.5

Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire défini sur Ω et à valeurs dans \mathbb{R}^n . On suppose que toutes les composantes du vecteur sont intégrables. Alors X est dit intégrable et son espérance est le vecteur de \mathbb{R}^n

$$\mathbb{E}[X] = (\mathbb{E}[X_1], \dots, \mathbb{E}[X_n]).$$

On déduit aisément des propriétés de l'espérance des variables réelles, celles de l'espérance d'un vecteur aléatoire :

Proposition 3.7

Soient X et Y deux vecteurs aléatoires intégrables à valeur dans \mathbb{R}^n et soient $a \in \mathbb{R}$ et $b \in \mathbb{R}^n$ des constantes. Alors

1. $aX + b$ est intégrable et $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$.
2. $X + Y$ est intégrable et $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$.

3.4.4 Covariance de deux vecteurs aléatoires.

On peut prendre la covariance de vecteurs aléatoires de tailles différentes du moment qu'ils sont définis sur le même univers.

Définition 3.6

Soient $X = (X_1, \dots, X_n)$ et $Y = (Y_1, \dots, Y_m)$ deux vecteurs aléatoires définis sur Ω et à valeurs dans \mathbb{R}^n et \mathbb{R}^m respectivement. On suppose que les composantes des deux vecteurs sont de carré intégrable. Alors, la covariance des vecteurs X et Y est la matrice réelle $n \times m$

$$\text{cov}[X, Y] = (\text{cov}[X_i, Y_j])_{i=1..n, j=1..m}.$$

Si $\text{cov}[X, Y] = 0$, X et Y sont dits décorrélés.

Remarque : Remarquer que X et Y sont décorrélés si et seulement si les composantes de X et les composantes de Y sont deux à deux décorrélés.

Remarque : On peut également définir la covariance de X et Y par

$$\text{cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])^t],$$

où l'exposant t désigne l'opération de transposition.

Proposition 3.8

La covariance est bilinéaire.

3.4.5 Variance (ou covariance) d'un vecteur aléatoire.**Définition 3.7**

Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire défini sur Ω et à valeurs dans \mathbb{R}^n . On suppose que toutes les composantes du vecteur sont de carré intégrable. On définit alors la variance ou la covariance de X par la matrice de dimension $n \times n$

$$\text{Var}[X] = \text{cov}[X, X] = (\text{cov}[X_i, X_j])_{i=1..n, j=1..n}.$$

De la bilinéarité de la covariance de vecteurs, on déduit la

Proposition 3.9

1. Soient X un vecteur aléatoire de \mathbb{R}^n de composantes de carré intégrable, a un réel et b un vecteur de \mathbb{R}^n . Alors

$$\text{Var}[aX + b] = a^2 \text{Var}[X].$$

2. Soient X et Y deux vecteurs aléatoires de \mathbb{R}^n de composantes de carré intégrable. Alors

$$\text{Var}[X + Y] = \text{Var}[X] + \text{cov}[X, Y] + \text{cov}[Y, X] + \text{Var}[Y].$$

3.4.6 Espérance et variance sous transformation affine d'un vecteur aléatoire

Proposition 3.10

Soit X un vecteur aléatoire de \mathbb{R}^d dont les composantes sont de carré intégrable. On définit $Y = AX + B$ un vecteur de $\mathbb{R}^{d'}$ par transformation affine de X (A est une matrice de dimension $d' \times d$ et B un vecteur de dimension d'). Alors

1. $\mathbb{E}[Y] = A\mathbb{E}[X] + B$;
2. $\text{Var}[Y] = A\text{Var}[X]A^t$.

Démonstration : On suppose $A = (a_{ij})_{1 \leq i \leq d', 1 \leq j \leq d}$, $x = (x_i)_{1 \leq i \leq d}$, $y = (y_i)_{1 \leq i \leq d'}$.

3.5 Variables aléatoires indépendantes.

Définition 3.8

Soient X et Y deux v.a.r. sur Ω . On dit que X et Y sont indépendantes si, pour tous $A, B \subset \mathbb{R}$,

$$\mathbb{P}[X \in A, Y \in B] = \mathbb{P}[X \in A]\mathbb{P}[Y \in B].$$

Proposition 3.11

Si (X, Y) est un couple de variables aléatoires qui admet pour densité $f_{(X,Y)} : \mathbb{R}^2 \rightarrow \mathbb{R}$, et si f_X, f_Y sont les densités marginales de X et Y respectivement, X et Y sont indépendantes si et seulement si

$$f_{(X,Y)}(x, y) = f_X(x)f_Y(y) \text{ pour presque tout } (x, y) \in \mathbb{R}^2.$$

Définition 3.9

Plus généralement, si (X_1, \dots, X_n) est un vecteur aléatoire, on dit que les v.a.r. X_1, \dots, X_n sont indépendantes si, pour tous $B_1, \dots, B_n \subset \mathbb{R}$,

$$\mathbb{P}[X_1 \in B_1, \dots, X_n \in B_n] = \mathbb{P}[X_1 \in B_1] \dots \mathbb{P}[X_n \in B_n].$$

L'expression de cette indépendance dans les cas discrets et à densité est une extension évidente du cas $n = 2$ ci-dessus et est laissée au lecteur.

Exemple : Dans le cas de la flèche lancée sur la cible, les variables X et Y sont-elles indépendantes ?

Graphiquement :

Plus généralement, on peut montrer qu'une condition nécessaire pour que les v.a. X et Y d'un couple de v.a. (X, Y) admettant une densité f soient indépendantes est que le support de f soit du type $A \times B$ où $A, B \subset \mathbb{R}$.

Proposition 3.12

Si X et Y sont deux v.a.r. indépendantes et de carré intégrable alors elles sont décorrélées, c'est-à-dire que $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

Remarque : La réciproque est fautive en général. En effet, toujours avec notre cible, il est clair que

$$\mathbb{E}[X] = \mathbb{E}[Y] =$$

D'autre part, XY prend ses valeurs dans $[-1, 1]$ et, grâce à la formule de transfert, on a :

$$\mathbb{E}[XY] = \int_{\mathbb{R}^2} xy \frac{1}{\pi} \mathbf{1}_D(x, y) dx dy =$$

Par conséquent, $\text{cov}[(X, Y)] = 0$, les variables sont décorrélées mais non indépendantes (vu plus haut).

Démonstration : On démontre la proposition dans le cas d'un couple de variables à densité (pour des variables discrètes, la preuve est similaire). On suppose X et Y indépendantes, de carré intégrable, et que le couple (X, Y) admet une densité f . On note f_X, f_Y les densités marginales de X et Y respectivement. Par la formule de transfert,

$$\mathbb{E}[XY] =$$

Comme X et Y sont indépendantes, on a :
théorème de Fubini,

. Et donc, par le

Proposition 3.13

Soient X_1, \dots, X_n n v.a.r. indépendantes.

1. Elles sont indépendantes deux à deux mais la réciproque est fausse.
2. On connaît la loi de (X_1, \dots, X_n) dès que l'on connaît les lois de chacune des variables $X_i, i \in \{1, \dots, n\}$.
3. Soit $p \in \{1, \dots, n-1\}$. Alors, pour toutes fonctions $\phi : \mathbb{R}^p \rightarrow \mathbb{R}$ et $\psi : \mathbb{R}^{n-p} \rightarrow \mathbb{R}$, les variables $\phi(X_1, \dots, X_p)$ et $\psi(X_{p+1}, \dots, X_n)$ sont indépendantes.

3.6 Somme de v.a. indépendantes.

Un cas où l'indépendance de variables aléatoires est particulièrement utile est celui du calcul de la loi de la somme de v.a. En effet, en général, on ne sait rien de la somme de v.a. si l'on ne fait pas d'hypothèse supplémentaire.

En toute généralité, la loi de $X + Y$ est connue dès que l'on connaît la loi du couple. En effet, si par exemple (X, Y) admet une densité f , alors pour toute fonction bornée $\varphi : \mathbb{R} \rightarrow \mathbb{R}$,

$$\mathbb{E}[\varphi(X + Y)] = \int_{\mathbb{R}^2} \varphi(x + y) f(x, y) dx dy.$$

Cela dit, lorsque les v.a. sont indépendantes, on peut pousser les calculs plus loin.

Proposition 3.14

Soient X, Y deux v.a. indépendantes et soit $Z = X + Y$. On suppose que X et Y admettent des densités f et g respectivement. Alors Z admet pour densité $h = f * g$ (la convolution de f et g) définie, pour presque tout $z \in \mathbb{R}$, par

$$h(z) = \int_{\mathbb{R}} f(x) g(z - x) dx = \int_{\mathbb{R}} f(z - y) g(y) dy$$

Démonstration : Les v.a. étant indépendantes, le couple (X, Y) admet pour densité $(x, y) \mapsto f(x)g(y)$.

Pour trouver la loi de $Z = X + Y$, on utilise la méthode vue dans le chapitre sur les moments : soit φ une application bornée de \mathbb{R} dans \mathbb{R} . On a, en utilisant la formule de transfert,

$$\mathbb{E}[\varphi(Z)] =$$

$$\Rightarrow \mathbb{E}[\varphi(Z)] =$$

Puis, en faisant le changement de variable $z = x + y$, on obtient

$$\mathbb{E}[\varphi(Z)] =$$

$$\mathbb{E}[\varphi(Z)] =$$

Donc Z a pour densité la fonction $h = f * g$ (la convolution de f et g) définie, pour presque tout $z \in \mathbb{R}$, par

$$h(z) = \int_{\mathbb{R}} f(x)g(z-x)dx = \int_{\mathbb{R}} f(z-y)g(y)dy$$

L'exemple qui suit et des résultats du même type sont dans l'annexe « lois usuelles ». Il faut connaître ces résultats.

Exemple : On suppose que X et Y sont indépendantes et suivent des lois normales $\mathcal{N}(0, 1)$. Alors $X + Y \rightsquigarrow \mathcal{N}(0, 2)$.

Alors, d'après la proposition précédente, si on note respectivement f et g les densités de X et Y (ici, $f = g$), Z a pour densité la fonction h définie sur \mathbb{R} par

$$h(z) = \int_{\mathbb{R}} f(x)g(z-x)dx.$$

Rappelons que la densité d'une v.a. $\mathcal{N}(m, \sigma^2)$ est

$$x \mapsto \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-m)^2/(2\sigma^2)} \text{ pour tout } x \in \mathbb{R}.$$

Donc on a ($\sigma^2 = 1$ et $m = 0$)

$$f(x) = g(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \text{ pour tout } x \in \mathbb{R}$$

et on veut ($\sigma^2 = 2$, $m = 0$)

$$h(z) = \frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-z^2/4} \text{ pour tout } z \in \mathbb{R}.$$

Or,

$$h(z) =$$

On remarque que

$$(*) \quad x^2 + (x-z)^2 =$$

Remarquer que dans le formulaire « lois usuelles » le résultat est plus général :

Proposition 3.15

Si X_1 et X_2 sont deux v.a. indépendantes suivant respectivement des lois $\mathcal{N}(m_1, \sigma_1^2)$ et $\mathcal{N}(m_2, \sigma_2^2)$, alors $X_1 + X_2 \rightsquigarrow \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$.

Démonstration : Pour montrer cela, on commence par montrer que la propriété est satisfaite lorsque $m_1 = m_2 = 0$ (comme ci-dessus mais la factorisation (*) est plus difficile) puis on remarque que, si $X_i \rightsquigarrow \mathcal{N}(m_i, \sigma_i^2)$ alors $X_i - m_i \rightsquigarrow \mathcal{N}(0, \sigma_i^2)$.

La proposition établit en outre que la somme de deux v.a. indépendantes à densité est une v.a. à densité. On a aussi le résultat que la somme de deux v.a. indépendantes discrètes est une v.a. discrète. Cette remarque suggère qu'il existe des variables aléatoires ni discrètes ni à densité (en sommant deux v.a. indépendantes de nature différente).

3.7 Vecteurs gaussiens

Définition 3.10

On dit que X vecteur aléatoire de \mathbb{R}^d est un vecteur gaussien de paramètres $M \in \mathbb{R}^d$ et $\Gamma \in S^d$ et on note $X \rightsquigarrow \mathcal{N}_d(M, \Gamma)$ si X a pour densité sur \mathbb{R}^d

$$f(x) = \frac{1}{\sqrt{|\det \Gamma|} \sqrt{2\pi}^d} \exp \left(-\frac{1}{2} \langle \Gamma^{-1}(x - M), x - M \rangle \right), x \in \mathbb{R}^d$$

où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire dans \mathbb{R}^d .

Les vecteurs gaussiens ont des propriétés bien particulières. Il faudra retenir en particulier :

- Les vecteurs gaussiens sont bien des collections de v.a. gaussiennes mais toute collection de v.a. gaussiennes ne constitue pas un vecteur gaussien (Cf contre-exemple).
- Un vecteur gaussien dont les composantes sont deux à deux décorrélées forme un vecteur de v.a. mutuellement indépendantes.

Ces particularités, ainsi que l'importance que les v.a. gaussiennes ont dans la modélisation (voir chapitre suivant) font des vecteurs gaussiens un outil très important de la Statistique. Dans cette section, nous détaillons un peu les deux points ci-dessus et nous combinons des v.a. gaussiennes indépendantes pour construire de nouvelles lois qui seront utilisées en statistique.

3.7.1 Propriétés remarquables

Dans la définition de la densité des vecteurs gaussiens, les paramètres, comme en dimension 1, sont respectivement M l'espérance et Γ la variance du vecteur.

Proposition 3.16

Un vecteur aléatoire X est gaussien au sens de la définition du chapitre 1 si et seulement si toute combinaison linéaire de ses composantes est une variable aléatoire réelle gaussienne et si Γ est inversible.

Remarque : La définition suppose l'inversibilité de Γ , ce qui justifie cette formulation. En réalité, on classe parmi les vecteurs gaussiens les vecteurs dont toutes les combinaisons linéaires sont gaussiennes. Ces vecteurs n'admettent pas nécessairement une densité, c'est pourquoi nous ne considérons pas dans ce cours les vecteurs gaussiens dans toute leur généralité.

Proposition 3.17

Soit (X_1, \dots, X_n) un vecteur gaussien. Les v.a. X_1, \dots, X_n sont (mutuellement) indépendantes si et seulement si elles sont deux à deux décorrélées.

Les vecteurs gaussiens à composantes indépendantes se reconnaissent donc à leur matrice de covariance Γ diagonales.

Ces propriétés, ainsi que la définition des vecteurs gaussiens, sont à manier avec précaution. Nous allons ici construire un exemple de vecteur $X = (X_1, X_2)$ de \mathbb{R}^2 tel que

- X_1 et X_2 sont des v.a. de loi $\mathcal{N}(0, 1)$ décorrélées (donc $\Gamma = \text{Id}$).
- X_1 et X_2 ne sont pas indépendantes
- X n'est pas un vecteur gaussien

Exemple : Soient X_1 une variable gaussienne $\mathcal{N}(0, 1)$ et ϵ une variable prenant les valeurs -1 et 1 avec probabilité 1/2. On suppose X_1 et ϵ indépendantes et on définit X_2 par $X_2 = \epsilon X_1$ et on montre les trois propriétés ci-dessus.

On a enfin le résultat évident suivant :

Proposition 3.18

Soient X_1, X_2, \dots, X_n n v.a. gaussiennes indépendantes. Alors (X_1, \dots, X_n) est un vecteur gaussien de paramètre Γ diagonal.

Démonstration : On peut remarquer par exemple que les combinaisons linéaires des composantes sont des v.a. gaussiennes, par la proposition 3.15, ou bien que la densité du vecteur a bien la forme voulue par la propriété 3.11.

3.7.2 Lois du chi-deux, de Student, de Fisher-Snedecor

Lemme 3.1

Soit X une v.a. suivant une loi $\mathcal{N}(0, 1)$. Alors

$$X^2 \rightsquigarrow \text{Gamma}\left(\frac{1}{2}, \frac{1}{2}\right).$$

Définition 3.11

Soient X_1, \dots, X_n n v.a. indépendantes de loi $\mathcal{N}(0, 1)$. La loi de $X_1^2 + \dots + X_n^2$ est appelée une loi du chi-deux à n degrés de liberté, notée χ_n^2 .

En consultant dans « lois usuelle » la partie concernant la somme de v.a. indépendantes suivant une loi Gamma, on se convainc facilement des résultats suivants :

Proposition 3.19

$\chi_n^2 = \text{Gamma}(1/2, n/2)$. Par ailleurs, si X et Y sont indépendantes et suivent respectivement des lois du chi-deux à n et p degrés de liberté alors $X + Y$ suit une loi du chi-deux à $n + p$ degrés de liberté.

Nous définissons maintenant la loi du rapport d'une loi normale centrée réduite et de la racine d'un chi-deux indépendant, la loi de student.

Définition 3.12

Soient X et Y des v.a. indépendantes suivant respectivement des lois $\mathcal{N}(0, 1)$ et χ_n^2 . La loi de

$$T = \frac{X}{\sqrt{Y/n}}$$

s'appelle la loi de Student à n degrés de liberté, notée T_n .

Sa densité est la fonction

$$f(x) = \frac{(1 + x^2/n)^{-(n+1)/2}}{\sqrt{n}B(1/2, n/2)} \text{ pour } x \in \mathbb{R}$$

où B est une fonction définie à partir de la fonction Γ dans la définition de la loi Bêta (voir annexes).

Proposition 3.20

La loi de Student à 1 degré de liberté est la loi de Cauchy de paramètre 1 (qui n'admet aucun moment).

Si $n > 1$, $\mathbb{E}[T_n] = 0$ et, si $n > 2$,

$$\text{Var}[T_n] = \frac{n}{n-2}.$$

Enfin, nous aurons besoin de la loi du rapport de deux chi-deux indépendants.

Théorème 3.2

Soient X et Y deux v.a. indépendantes suivant respectivement des loi du Chi-deux à n et p degrés de liberté. Alors

$$F = \frac{X/n}{Y/p}$$

suit une loi de Fisher-Snedecor à (n, p) degrés de liberté notée $F_{n,p}$.

4 Suites de variables aléatoires - Résultats asymptotiques

4.1 Un exemple important.

Une des grandes questions de la théorie des probabilités a été la suivante : si je lance un très grand nombre de fois une pièce de monnaie et que je calcule la *fréquence observée* de « pile », est-ce que j'obtiens, à la limite quand le nombre de tirages tend vers l'infini, la *fréquence théorique* de « pile », à savoir, si la pièce est équilibrée, $1/2$? De plus, puis-je connaître la distribution de cette fréquence observée quand n devient grand ?

Modélisons le problème. On modélise chaque tirage par une v.a. de Bernoulli de paramètre $1/2$.

Comme le nombre de tirages est destiné à tendre vers l'infini, on se donne donc une suite $(X_n)_{n \geq 1}$ de variables suivant la loi $\mathcal{B}(1/2)$. Il est assez naturel de supposer que les tirages sont indépendants, ce qui se traduit par le fait que les variables sont indépendantes, c'est-à-dire que tout vecteur aléatoire formé par un nombre fini quelconque de v.a. de cette suite forme une famille de v.a. indépendantes.

Si on compte 1 lorsque la pièce tombe sur « pile » et 0 sinon, la fréquence de « pile » observée après n tirages sera

$$\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k.$$

Le première question à laquelle il faut répondre est : a-t-on

$$\lim_{n \rightarrow +\infty} \bar{X}_n = \frac{1}{2} ?$$

Et si oui, en que sens ?

Les théorèmes qui répondent à ce type de questions sont connus sous le nom de « lois des grands nombres ». On en verra un en particulier.

La seconde question, à savoir : « quelle est la distribution, quand n devient grand, de la fréquence observée de « pile » \bar{X}_n » est plus délicate. La réponse est contenue dans un résultat célèbre appelé le « théorème limite central » qui dit que pour n assez grand, \bar{X}_n suit approximativement une loi normale.

Ces résultats dépassent largement le cadre de l'exemple proposé puisqu'ils ne dépendent pas de la loi de la variable dont on évalue la moyenne (ici la loi de Bernoulli).

Avant de voir ces résultats, définissons deux notions de convergence.

4.2 Différents types de convergence.

Définition 4.1

Soit $(X_n)_{n \geq 1}$ une suite de v.a.r. définies sur Ω et soit X une v.a.r. définie sur le même univers.

1. $(X_n)_{n \geq 1}$ converge presque sûrement vers X s'il existe un négligeable $N \subset \Omega$ (i.e. un élément de $\mathcal{P}(\Omega)$ satisfaisant $\mathbb{P}[N] = 0$) tel que, $X_n(\omega) \rightarrow X(\omega)$ pour tout $\omega \in N^c$.
2. On dit que $(X_n)_{n \geq 1}$ converge en loi vers X si, pour tout $t \in \mathbb{R}$ auquel F_X est continue, $\lim_{n \rightarrow +\infty} F_{X_n}(t) = F_X(t)$.

Remarque : La notion de convergence presque sûre correspond à celle de convergence presque partout pour les suites de fonctions réelles. En revanche, la notion de convergence en loi est totalement neuve. On utilise couramment d'autres types de convergence "en moyenne", "en moyenne quadratique", "en probabilité" que nous ne définissons pas ici.

Proposition 4.1 (admise)

Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires réelles définies sur Ω et soit X une v.a.r. définie sur le même univers. Si $(X_n)_{n \geq 1}$ converge vers X presque sûrement alors $(X_n)_{n \geq 1}$ converge en loi vers X .

Remarque : La réciproque de l'implication de la proposition précédente est fautive en général. Deux variables de même loi peuvent être différentes en chaque point $\omega \in \Omega$ et $(X_n)_{n \geq 1}$ converge en loi vers X signifie seulement que la loi de $(X_n)_{n \geq 1}$ converge vers la loi de X . On notera cette convergence : $\mathcal{L}(X_n) \rightarrow \mathcal{L}(X)$ quand $n \rightarrow +\infty$.

Exemple : Soit Y définie sur Ω une variable aléatoire réelle de loi $\mathcal{B}(1/2)$. Soit $X = 1 - Y$. Alors X a la même loi que Y . Pour tout n , on pose $X_n = Y$ et $X = 1 - Y$.

Alors $(X_n)_{n \geq 1}$ converge trivialement vers X en loi. Et pourtant $|X_n - X| = 1$ sur Ω et donc on n'a pas convergence presque sûre.

Proposition 4.2

On définit $(U_N)_{N \geq 1}$ une suite de v.a. telle que, pour N fixé, U_N suit une loi uniforme sur

$$\left\{0, \frac{1}{N}, \frac{2}{N}, \dots, \frac{N-1}{N}\right\}.$$

Alors $(U_N)_{N \geq 1}$ converge en loi vers la loi uniforme sur $[0, 1[$.

Démonstration : Pour tout $N \geq 1$, on définit F_N la fonction de répartition de U_N , soit

La fonction de répartition d'une loi uniforme est

Cette fonction est continue et par conséquent, la convergence en loi est la convergence ponctuelle de la suite $(F_N)_{N \geq 1}$. Pour $t < 0$ et $t \geq 1$, $F_N(t) = 0$ et $F_N(t) = 1$ respectivement. Pour $0 \leq t < 1$, $F_N(t) = \frac{[Nt]}{N}$ où $[x]$ désigne la partie entière d'un réel x . Or, pour tout $t \in \mathbb{R}$, $\lim_{N \rightarrow +\infty} [Nt]/N = t$, d'où le résultat.

4.3 Théorèmes de convergence.

4.3.1 Les lois des grands nombres

Les lois des grands nombres sont des théorèmes selon lesquels une moyenne des n premiers termes d'une suite de variables aléatoires converge vers une constante. Lorsque les variables sont intégrables et de même espérance, cette constante est la moyenne. Plusieurs lois ont été démontrées sous des hypothèses différentes. Lorsque la convergence vers la constante est presque sûre on parle de loi forte et lorsque la convergence est plus faible, de loi faible des grands nombres.

Théorème 4.1 (Une loi forte des grands nombres pour des v.a.r. décorrélées.)

Soit $(X_n)_{n \geq 1}$ une suite de v.a.r. de carré intégrable, décorrélées, d'espérance m et dont la suite des variances est bornée par une constante $C > 0$. Pour tout n , on note

$$\bar{X}_n = \frac{X_1 + \dots + X_n}{n}.$$

Alors, pour tout $\varepsilon > 0$, on a les estimations

$$\mathbb{P} [|\bar{X}_n - m| > \varepsilon] \leq \frac{C}{n\varepsilon^2} \quad \text{et} \quad \mathbb{E} [(\bar{X}_n - m)^2] \leq \frac{C}{n}.$$

De plus, la suite $(\bar{X}_n)_{n \geq 1}$ converge presque sûrement vers m .

Démonstration : On a

$$\mathbb{E} [\bar{X}_n] =$$

et, puisque les v.a. sont décorréelées,

. On en déduit que

On remarque que cette quantité tend vers 0 quand $n \rightarrow +\infty$: cette convergence s'appelle la convergence en moyenne quadratique.

En appliquant l'inégalité de

On remarque que cette quantité tend vers 0 à ε fixé quand n tend vers l'infini : ce type de convergence s'appelle la convergence en probabilité.

On admet la convergence presque sûre.

4.3.2 Théorème limite central

À présent que l'on sait que la moyenne empirique d'une suite de variables aléatoires décorréelées se concentre autour de son espérance, la question suivante est naturelle : que peut-on dire des fluctuations de la moyenne empirique autour de l'espérance, c'est-à-dire de la distribution de $\bar{X}_n - m$? La réponse à cette question, le Théorème Central Limite, est un des résultats majeurs de la théorie des probabilités, et est assez extraordinaire : il affirme que

1. $\bar{X} - m$ est de l'ordre de $\frac{1}{\sqrt{n}}$.
2. La distribution de $(\bar{X} - m)/\frac{\sigma}{\sqrt{n}}$ approche la même distribution, lorsque n devient grand, quelle que soit la distribution des X_i , tant que ceux-ci ont une variance σ^2 finie !

Ce théorème repose non plus sur la décorrélation des variables de la suite mais sur leur indépendance. Précisons la définition.

Définition 4.2

Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires. On dit que les v.a. de cette suite sont indépendantes si, pour tout $N \geq 2$, les v.a. X_1, \dots, X_N sont indépendantes, au sens vu au chapitre précédent.

Théorème 4.2 (Théorème limite central.)

Soit $(X_n)_{n \geq 1}$ une suite de v.a.r. de carré intégrable, indépendantes et identiquement distribuées (c'est-à-dire de

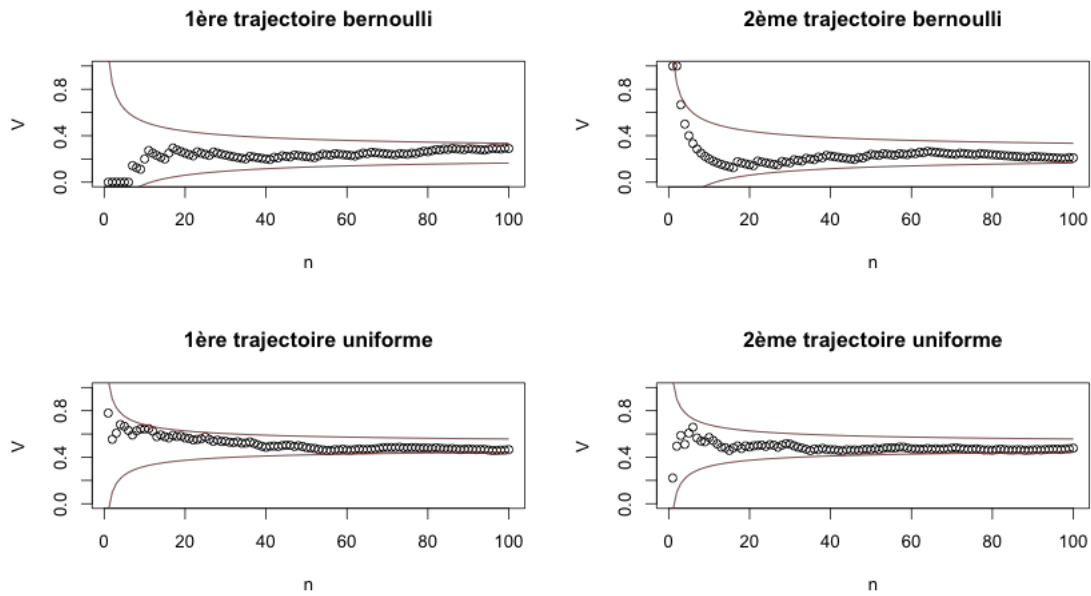


FIGURE 5 – Loi des grands Nombres (Bernoulli(0.25) et Uniforme sur $[0; 1]$)

même loi — i.i.d. en abrégé). On pose $m = \mathbb{E}[X_1]$, $\sigma^2 = \text{Var}[X_1]$. Alors, quand $n \rightarrow +\infty$,

$$\mathcal{L}\left(\frac{X_1 + \dots + X_n - nm}{\sqrt{n}\sigma}\right) \rightarrow \mathcal{N}(0, 1).$$

Ce théorème, qui sera très largement utilisé en Statistique, est admis.

Une application usuelle du théorème limite central est le calcul approché de probabilité.

Exemple : Une chaîne de montage produit des pièces défectueuses avec un taux de 10%. Quelle est la probabilité d'obtenir au moins 50 pièces défectueuses parmi 400 ?

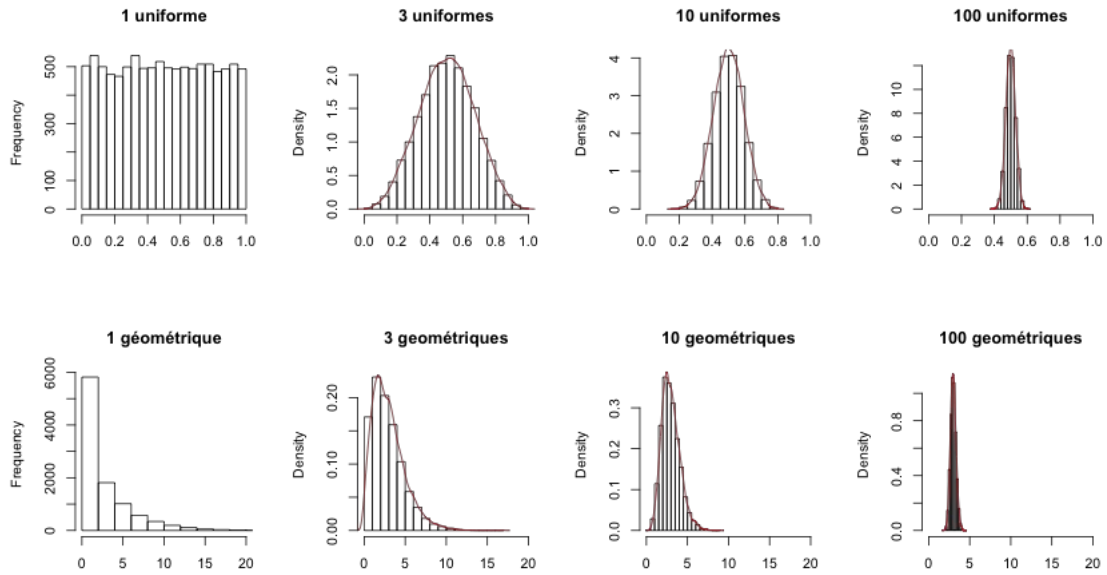


FIGURE 6 – TCL-Illustration

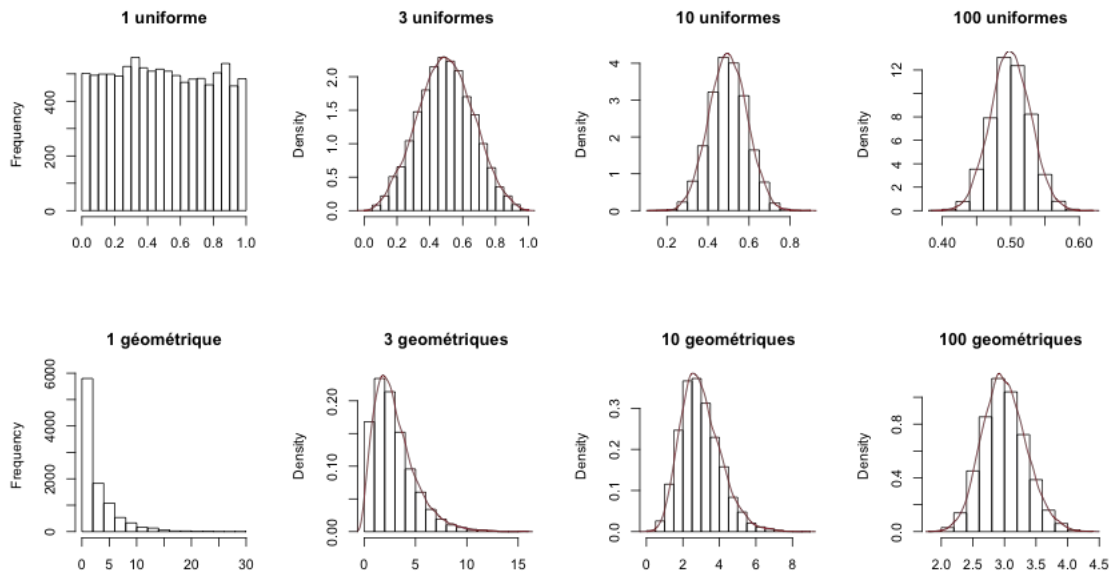


FIGURE 7 – TCL-Illustration - zoom

4.4 Simulation de variables aléatoires

La plupart des langages informatiques mettent à disposition de l'utilisateur un générateur de nombre pseudo-aléatoires. Il s'agit en fait de suites récurrentes entières avec une très grande périodicité qui sont initialisées via des

paramètres liés à l'ordinateur (par exemple l'horloge interne) et qui sont renormalisées pour être à valeurs dans $[0, 1]$. On considère qu'approximativement, via sa fonction "random", un langage fournit une suite de v.a. indépendantes de loi uniforme sur $[0, 1]$.

A partir de cette suite de v.a. indépendantes, on doit être capable de construire une suite de v.a. indépendantes suivant n'importe quelle loi.

4.4.1 Simulation d'une v.a. admettant une densité continue par inversion de la fonction de répartition

Proposition 4.3

Soit X une v.a. de densité f continue et de fonction de répartition F inversible. On pose $Y = F(X)$. Alors Y suit une loi uniforme sur $[0, 1]$.

Démonstration : D'après la proposition 2.10 sur les changements de variables bijectifs pour les v.a. à densité, Y admet pour densité la fonction g définie, pour $y \in]0, 1[$ par

Par ailleurs, pour $y < 0$, $\mathbb{P}[F(X) \leq y] = 0$ et pour $y \geq 1$, $\mathbb{P}[F(X) \leq y] = 1$ et le théorème est démontré.

Pour simuler la v.a. X , si F^{-1} est connue analytiquement, il suffit donc de simuler une loi uniforme sur $]0, 1[$ et de calculer F^{-1} du résultat obtenu.

4.4.2 Simulation d'une v.a. bornée à densité bornée par la méthode de rejet

Proposition 4.4

Soit X une v.a. presque sûrement bornée de densité f presque partout bornée. On suppose pour simplifier que $X \in [0, 1]$ et on choisit m tel que $f \leq m$ p.p. Soient U et V deux v.a. indépendantes de lois uniformes sur $[0, 1]$ et $[0, m]$ respectivement. Alors, conditionnellement à $[V < f(U)]$, U a même loi que X .

Démonstration : Soit $x \in \mathbb{R}$. On calcule $\mathbb{P}[U \leq x | V < f(U)]$.

Par la formule de transfert appliquée à une fonction du couple (U, V) de densité égale au produit de ses marginales, on a, si $x \in [0, 1]$

$$\begin{aligned} \mathbb{P}[U \leq x, V < f(U)] &= \int_0^x \int_0^{f(u)} \frac{dv}{m} du \\ &= \int_0^x \frac{f(u)}{m} du = \int_{-\infty}^x \frac{f(u)}{m} du \end{aligned}$$

puisque X prend ses valeurs dans $[0, 1]$. De plus,

$$\mathbb{P}[V < f(U)] = \int_0^1 \int_0^{f(u)} \frac{1}{m} dv du = \int_0^1 \frac{f(u)}{m} du = \frac{1}{m}.$$

Finalement, $\mathbb{P}[U \leq x | V < f(U)] = \int_{-\infty}^x f(u) du$ pour $x \in [0, 1[$ et on peut vérifier facilement que cela reste vrai pour $x < 0$ (probabilité nulle) et $x \geq 1$ (probabilité égale à 1).

A partir de ce résultat, on construit la procédure de simulation suivante : étant donnée une v.a. X comme dans la proposition, on procède à la simulation de U et V indépendantes. Si les réalisations u et v de ces v.a. satisfont $v < f(u)$ la réalisation de la simulation de X est u , sinon on procède à un autre tirage des v.a. U et V .

Cette méthode est utilisée pour simuler des lois bêta.

Définition 4.3 (Lois Beta)

1. On dit que la v.a. X suit une loi bêta de type I de paramètres réels n et $p > 0$ et on note $X \rightsquigarrow \text{bêta I}(n, p)$ si elle est comprise entre 0 et 1 et si sa densité est

$$f(x) = \frac{1}{B(n, p)} x^{n-1} (1-x)^{p-1} \mathbf{1}_{]0,1[}(x)$$

où la fonction B est donnée par

$$B(n, p) = \int_0^1 x^{n-1} (1-x)^{p-1} dx.$$

2. On dit que la v.a. $Y \geq 0$ suit une loi bêta de type II de paramètres n, p et on note $Y \rightsquigarrow \text{bêta II}(n, p)$ si elle admet pour densité la fonction

$$f(y) = \frac{1}{B(n, p)} \frac{y^{n-1}}{(1+y)^{n+p}} \mathbf{1}_{]0,+\infty[}(y).$$

4.4.3 Simulation d'une loi normale

Parfois, on ne connaît les v.a. qu'à travers un changement de variable et si une v.a. s'exprime comme une fonction de v.a. indépendantes il peut être plus simple de simuler ces v.a. et de reconstituer la v.a. de départ à travers une fonction que de chercher à la simuler directement. La loi normale est un exemple de ce principe.

Proposition 4.5

Soient U et V deux v.a. indépendantes de loi uniforme sur $]0, 1[$. On pose

$$X = \sqrt{-2 \ln U} \cos 2\pi V, \quad Y = \sqrt{-2 \ln U} \sin 2\pi V.$$

Alors X et Y sont des v.a. indépendantes de loi $\mathcal{N}(0, 1)$.

Générer une v.a. $\mathcal{N}(0, 1)$ nécessite donc de générer deux v.a. indépendantes de loi uniforme sur $]0, 1[$ puis d'utiliser l'une des deux formules proposées.

4.5 Méthode de Monte Carlo pour le calcul des intégrales et des moments.

La simulation aléatoire permet de faire des calculs d'intégrales. Cette méthode, appelée la méthode de Monte Carlo, repose sur la loi des grands nombres : sous de bonnes hypothèses, on peut approcher presque sûrement l'espérance d'une variable aléatoire (une intégrale dans le cas à densité) par la moyenne arithmétique de variables aléatoires de même espérance. L'idée de la méthode est alors de simuler ces v.a. et de considérer la réalisation de leur moyenne arithmétique comme une approximation numérique de leur espérance commune.

4.5.1 Méthode

On suppose que l'on cherche à calculer numériquement

$$\int_{\mathbb{R}} f(x) dx \text{ pour une fonction } f \text{ intégrable.}$$

On remarque que, par des changements de variable successifs, on peut se ramener au calcul de

$$\int_0^1 g(u) du, \text{ pour une certaine fonction } g.$$

Mais, étant donné une v.a. U de loi uniforme sur $]0, 1[$, par la formule de transfert,

$$\mathbb{E}[g(U)] = \int_0^1 g(u) du.$$

Plus précisément, si f est intégrable sur \mathbb{R} , on remarque par exemple que

$$\int_{\mathbb{R}} f(x) dx = \int_0^1 f\left(\frac{u}{1-u}\right) + f\left(-\frac{u}{1-u}\right) du.$$

Puis, en posant $g(u) = f\left(\frac{u}{1-u}\right) + f\left(-\frac{u}{1-u}\right)$, on obtient

$$\begin{aligned} \int_{\mathbb{R}} f(x) dx &= \int_0^1 g(u) du \\ &= \int_0^1 g(u) du \end{aligned}$$

avec

$$g(u) = f\left(\frac{u}{1-u}\right) + f\left(-\frac{u}{1-u}\right) \text{ pour } u \in [0, 1].$$

Soit alors $(U_n)_{n \geq 1}$ une suite de v.a. indépendantes suivant une loi uniforme sur $]0, 1[$. D'après la loi des grands nombres,

$$\frac{1}{n} \sum_{i=1}^n g(U_i) \rightarrow \int_0^1 g(u) du \text{ p.s. quand } n \rightarrow +\infty.$$

Une réalisation de cette somme donnera donc une approximation numérique de l'intégrale.

4.5.2 Précision (estimation d'erreur, intervalle de confiance)

On approche donc une constante (l'intégrale ou l'espérance) par la réalisation d'une variable aléatoire. Il n'est donc pas possible de caractériser avec certitude l'erreur d'approximation que l'on commet.

Néanmoins, on a les informations suivantes sur la variable aléatoire :

$$\mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n g(U_i)\right] = \mathbb{E}[g(U)] = \int_0^1 g(u) du,$$

c'est-à-dire que l'on approche la valeur sans erreur en moyenne.

De plus,

$$\text{Var}\left[\frac{1}{n} \sum_{i=1}^n g(U_i)\right] = \frac{1}{n} \text{Var}[g(U)],$$

autrement dit l'erreur quadratique moyenne tend vers 0 en $O(1/n)$ quand n tend vers $+\infty$.

En outre, on verra dans le cours de statistique que l'on connaît la loi approximative de l'erreur (sans supposer connue la variance de l'erreur) et donc il est possible de montrer non pas qu'elle se situe dans un intervalle petit autour de 0 lorsque n est grand, mais qu'elle se situe dans un intervalle petit autour de 0 **avec forte probabilité**, la dépendance entre la taille de l'intervalle et la probabilité faisant intervenir le nombre de simulations n .

De cet intervalle, on déduit un intervalle pour la valeur cherchée que l'on appelle intervalle de confiance qui est donné une fois fixés le nombre de simulations n et la probabilité $1 - \alpha$ de se trouver dans l'intervalle. $1 - \alpha$ est appelé le seuil de confiance (α sera défini sous le nom de seuil de risque).

Plus précisément, on définit :

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n g(U_i), \quad S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (g(U_i) - \bar{X}_n)^2$$

et $u_{1-\alpha/2}$ le quantile d'ordre $1 - \alpha/2$ de la loi normale centrée réduite. Alors l'intervalle

$$\left[\bar{X}_n - u_{1-\alpha/2} \frac{S_n}{\sqrt{n}}, \bar{X}_n + u_{1-\alpha/2} \frac{S_n}{\sqrt{n}} \right]$$

est un intervalle de confiance au seuil $1 - \alpha$ (autrement dit, la valeur de l'intégrale que l'on cherche à approcher est dans cet intervalle avec probabilité $1 - \alpha$).

On constate bien que plus n est grand, plus l'intervalle est petit pour α fixé.

4.5.3 Intérêts de la méthode

Cette méthode converge assez lentement et repose sur la qualité du générateur de nombres aléatoires dont on dispose mais elle est très robuste. En particulier, en grandes dimensions (dès la dimension 3), elle est plus rapide et plus simple d'utilisation que les intégrateurs numériques.

4.5.4 Calcul approché des moments

Une autre utilisation évidente de cette méthode est l'approximation des moments des v.a. En effet, toujours par la loi des grands nombres, l'espérance $\mathbb{E}[\varphi(X)]$ d'une v.a. X pour une fonction à valeurs réelles φ telle que $\varphi(X)$ est intégrable, peut être approchée p.s. quand n grand par

$$\mathbb{E}[\varphi(X)] \approx \frac{1}{n} \sum_{i=1}^n \varphi(X_i)$$

pour toute suite $(X_i)_{i \geq 1}$ iid à X .

Même lorsque X admet une densité f , cette approximation peut-être beaucoup plus précise à n petit que le calcul direct de la valeur

$$\int_{\mathbb{R}} \varphi(x) f(x) dx$$

car les points alors choisis par la méthode de Monte Carlo se font selon la densité f et non uniformément.

4.6 Méthode de “simulation Monte Carlo”.

La simulation permet également de calculer des probabilités faisant intervenir des phénomènes dont on ne connaît pas la distribution explicitement (phénomènes complexes dépendant de variables multiples).

On suppose que l'on sait simuler une v.a. X (par exemple X est fonction d'un vecteur aléatoire dont on connaît la loi). Pour calculer la probabilité $\mathbb{P}[X \in B]$, on remarque que, si X a pour densité f ,

$$\mathbb{P}[X \in B] = \int_B f(x) dx = \int_{\mathbb{R}} \mathbf{1}_B(x) f(x) dx = \mathbb{E}[\mathbf{1}_B(X)].$$

Cette valeur, d'après la section précédente, peut s'approcher en simulant un grand nombre de fois la variable X et en faisant la moyenne arithmétique des valeurs des réalisations obtenues de $\mathbf{1}_B(X)$, c'est-à-dire

$$\mathbb{P}[X \in B] \approx \frac{1}{n} \sum_{i=1}^n \mathbf{1}_B(X_i)$$

pour des v.a. X_i indépendantes de même loi que X . Cela qui peut aussi s'exprimer par

$$\mathbb{P}[X \in B] \approx \frac{\text{card}\{i / X_i \in B\}}{n},$$

autrement dit la probabilité que X soit dans B est approchée par la proportion de simulations observées dans B .

A ce calcul doit être joint un intervalle de confiance que l'on peut légèrement améliorer par rapport au précédent (nous sommes ici dans un cas particulier du calcul précédent) : on note

$$\hat{p} = \frac{\text{card}\{i / X_i \in B\}}{n}$$

l'approximation de la probabilité p cherchée ; alors

$$\left[\hat{p} - \frac{u_{1-\alpha/2}}{\sqrt{n}} \sqrt{\hat{p}(1-\hat{p})}, \hat{p} + \frac{u_{1-\alpha/2}}{\sqrt{n}} \sqrt{\hat{p}(1-\hat{p})} \right]$$

est un intervalle de confiance de p au seuil de risque $1 - \alpha$.

Enfin, grâce à cette méthode, on peut également calculer une distribution approchée de phénomènes aléatoires complexes. Dans le cas d'une variable aléatoire X à densité, par exemple, il s'agit d'obtenir une densité approchée, par exemple par une fonction constante par morceaux. Pour cela, on discrétise l'ensemble des valeurs prises par X en intervalles I « assez petits », puis on approche $\mathbb{P}[X \in I]$ par la méthode de Monte Carlo pour chaque intervalle I . On obtient ainsi une approximation discrète de la densité de X (cf Fig. 8).

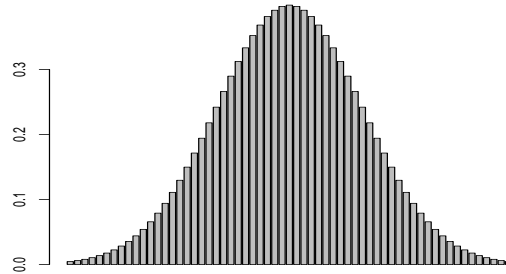


FIGURE 8 – Distribution empirique d'une loi normale

En pratique, cette méthode est utilisée quand la loi de X est connue comme une fonction de variables aléatoires indépendantes, i.e. quand $X = \psi(Y_1, \dots, Y_N)$ où Y_1, \dots, Y_N sont des v.a. indépendantes de lois connues et ψ une fonction connue. Pour chaque simulation souhaitée de X , on simule chacune des variables Y_1, \dots, Y_N puis on applique la fonction ψ au vecteur trouvé.

4.7 Rééchantillonnage (bootstrap)

Le bootstrap est une méthode qui permet de faire des calculs (de moments, de fonction de répartition d'une v.a. X) à partir d'une v.a. dont on ne connaît que des réalisations. Elle consiste à utiliser, comme distribution (inconnue) de la v.a. X sa distribution empirique (l'histogramme des données). Pour bien comprendre cette méthode, il faut déjà avoir vu la notion d'échantillon qui est traitée dans la partie suivante.

5 Enoncés de TD

PROBABILITÉS - SÉANCE DE TD 2

Dans cette séance, on traite des notions de densité d'une variable aléatoire réelle, de fonction de répartition et de moments de variables aléatoires réelles.

L'un des objectifs est de savoir réaliser tout type de changement de variables en une dimension.

Exercices à préparer : 1 et 2. Reconnaître les lois usuelles, savoir calculer leurs moments. Savoir faire des changements de variables simples.

Exercice 1 (Changement de variable, fonction de répartition, calcul des moments par formule de transfert)

Soit X une variable aléatoire de densité f . On pose $Y = X^2$.

1. Calculer la fonction de répartition de Y en fonction de celle de X .
2. On suppose que X suit une loi uniforme sur $[-1, 1]$.
 - (a) Calculer l'espérance et la variance de Y si elles existent.
 - (b) Calculer la densité de Y .

Exercice 2

Un point est choisi au hasard sur une ligne droite de longueur L . On obtient alors 2 segments. On souhaite calculer la probabilité p que le rapport entre le plus petit segment et le plus long soit strictement inférieur à $\frac{1}{4}$.

1. Modélisation : Soit X l'abscisse du point. Quelle est la loi de X ?
2. Exprimer la probabilité à calculer.
3. Déterminer la valeur numérique.

Exercice 3 (Loi log-normale)

On dit qu'une v.a.r. X suit une loi Log-normale de paramètres m et σ^2 , notée $\mathcal{LN}(m, \sigma^2)$ si $X = e^T$, T suivant une loi $\mathcal{N}(m, \sigma^2)$. Quelle est la loi de $Y = \alpha X^\beta$ pour $\alpha > 0$ et $\beta \in \mathbb{R}$.

Exercice 4 (Loi normale)

Lors d'un tir, on admet que les longueurs de tir suivent une loi normale. On constate que :

- (i) 10% des obus tombent à une distance supérieure à 1600km.
- (ii) 25% des obus tombent à une distance inférieure à 1400km.

Déterminer la longueur moyenne et l'écart-type des tirs. On pourra utiliser que si $Z \sim \mathcal{N}(0, 1)$, on a $\mathbb{P}[Z \leq 1, 28] = 0.9$ et $\mathbb{P}[Z \leq -0, 67] = 0.25$.

Exercice 5 (Loi log-normale)

Un tas de sable est composé de n grains homogènes sphériques. Le diamètre X d'un grain pris au hasard suit une loi $\mathcal{LN}(-0.5; 0.09)$ (m en mm). On passe le tas de sable au crible d'un tamis dont les trous sont de diamètres $0.5mm$.

1. Quelle est la probabilité p pour un grain donné de passer à travers le tamis.
2. Soit U la v.a. égale au nombre de grains passant au travers du tamis, quelle est la loi de U .
3. Quelle est la proportion moyenne de grains passant au travers du tamis.

Application numérique : si ϕ est la fonction de répartition d'une $\mathcal{N}(0, 1)$ alors $\phi(-0, 644) = 0, 26$.

Exercice 6

Soit X une v.a. de densité $f(x) = (1 - \frac{x}{2})\mathbf{1}_{[0,2]}$. On pose $Y = \sup(X, X^2)$

1. Calculer $E(X)$ et $E(X^2)$. En déduire que $E(Y) \geq \frac{2}{3}$.
2. Montrer que pour tout $x \in [0, 4]$,

$$F_Y(x) = \mathbb{P}[X \leq \inf(x, 1)] + \mathbb{P}[\{X \leq \sqrt{x}\} \cap \{X > 1\}].$$

En déduire la fonction de répartition de Y , puis sa densité.

3. Calculer $E(Y)$.

PROBABILITÉS - SÉANCE DE TD 3

Dans cette séance, on s'approprie les notions de vecteurs aléatoires et de lois jointes de variables aléatoires. A l'issue de cette séance, on doit maîtriser les changements de variables multidimensionnels et savoir calculer la loi de la somme de variables aléatoires indépendantes. On fera des représentations graphiques de distributions bidimensionnelles avec MATLAB.

Exercices à préparer : 1, 2, 3, 4, 6.

Exercice 1 (Loi de couple de v.a. à densité)

Le vecteur aléatoire (X, Y) possède la densité jointe suivante :

$$f_{X,Y}(x,y) = \begin{cases} 90x^2y(1-y) & \text{si } 0 \leq y \leq 1, 0 \leq x \leq y \\ 0 & \text{sinon.} \end{cases}$$

1. Les variables X et Y sont elles indépendantes ?
2. Calculer la densité marginale de X .
3. Calculer la densité marginale de Y .
4. Calculer la covariance de X et Y

Exercice 2 (Changement de variables - couple)

Soit U une variable aléatoire uniforme sur l'intervalle $[0, 2\pi]$ et Z , indépendante de U , de loi exponentielle de paramètre 1. On pose X et Y par

$$X = \sqrt{2Z} \cos U, Y = \sqrt{2Z} \sin(U).$$

1. Déterminer la loi du couple (X, Y) .
2. Les variables X et Y sont elles indépendantes.

Exercice 3

Loi du minimum et du maximum Soient X_1, \dots, X_N , N variables aléatoires réelles indépendantes de fonction de répartition F . Donner la fonction de répartition des variables aléatoires $U = \max(X_1, \dots, X_N)$ et $V = \min(X_1, \dots, X_N)$.

Exercice 4 (Somme de v.a. indépendantes - cas à densité)

Soient X et Y deux v.a. indépendantes suivant des lois uniformes sur $[0, 1]$. Quelle est la loi de $X + Y$?

Exercice 5 (Loi de couple de v.a. à densité)

On cherche à modéliser le tirage d'un point "au hasard" dans le disque D_0 de centre 0 et de rayon 1.

1. Expériences numériques.
 - (a) Dans un premier temps, on effectue des tirages indépendants d'un angle Θ pris uniformément dans $[0, 2\pi[$ et d'un rayon R uniforme dans $[0, 1]$. Sous MATLAB, effectuer un grand nombre (noté N dans la suite) de réalisations du couple (R, Θ) .
 - (b) Dans un second temps, on simule un couple (\bar{X}, \bar{Y}) de loi uniforme sur $[-1, 1] \times [-1, 1]$. Si le point simulé est dans le disque, on considère que c'est une réalisation du tirage voulu, sinon on recommence jusqu'à obtenir une réalisation dans D_0 . Effectuer ce tirage afin d'obtenir N réalisations dans le disque.
 - (c) Représenter graphiquement les tirages dans deux graphiques de la même figure (avec la commande 'axis square' après l'appel de chaque fonction 'plot'). Lequel des deux représente bien N tirages de la loi uniforme sur D_0 ?

2. Calcul des lois.

(a) Donner la densité du couple (R, Θ) .

(b) On note (X, Y) les points que l'on a effectivement représentés sur le second graphique. Ils valent (\bar{X}, \bar{Y}) conditionnellement à l'événement $[(\bar{X}, \bar{Y}) \in D_0]$. Calculer, pour tout $B \subset \mathbb{R}^2$,

$$\mathbb{P}[(X, Y) \in B] = \mathbb{P}[(\bar{X}, \bar{Y}) \in B / (\bar{X}, \bar{Y}) \in D_0]$$

pour vérifier que (X, Y) suit bien une loi uniforme sur D_0 .

(c) Donner la densité du couple $(\bar{R}, \bar{\Theta})$ obtenu en posant

$$X = \bar{R} \cos \bar{\Theta}, Y = \bar{R} \sin \bar{\Theta}$$

puis calculer les densités marginales de \bar{R} et $\bar{\Theta}$. Les v.a. $(\bar{R}, \bar{\Theta})$ sont-elles indépendantes ?

Exercice 6 (Calculs sur les espérances et les variances)

Soient X_1, \dots, X_n n variables aléatoires ayant la même loi qu'une variable aléatoire X d'espérance m et de variance σ^2 . On suppose de plus que ces variables aléatoires sont deux à deux décorrélées. On note

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \text{ et } S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Calculer l'espérance et la variance de \bar{X} et l'espérance de S^2 . A quelle condition la variable S^2 admet-elle une variance ?

Indication : pour le dernier calcul, on pourra remarquer que $X_i - \bar{X} = (X_i - m) + (m - \bar{X})$ et développer le carré selon ces deux termes.

PROBABILITÉS - SÉANCE DE TD 4

Dans cette séance, on étudie un exemple de convergence en loi, on applique le théorème limite central et la loi des grands nombres. On illustre ces applications avec des calculs numériques reposant sur la méthode de Monte Carlo. Il est très important de comprendre ces deux exemples et leur implémentation. A l'issue de cette séance, on proposera des exercices de simulation variés qu'il faudra réaliser en autonomie. En effet, la séance suivante sera consacrée à des travaux pratiques notés sous MATLAB.

Exercices à préparer : 1, 2, 3.

Exercice 1 (Convergence en loi)

Soit $(U_n)_n$ une suite de variables aléatoires indépendantes suivant toutes la loi uniforme sur $[0, 1]$. On note $M_n = \max(U_1, \dots, U_n)$ et $X_n = n(1 - M_n)$.

1. Quelle est la fonction de répartition de X_n ?
2. Etudier la convergence en loi de la suite $(X_n)_n$.

Exercice 2 (TLC)

Un assemblage comprend 100 sections mises bout à bout. La longueur de chaque section (en centimètres) est une variable aléatoire uniforme sur $[8, 12]$. De plus les sections sont indépendantes. La norme pour la longueur totale de l'assemblage est de $1000\text{cm} \pm 30\text{cm}$.

1. Soit L la longueur totale de l'assemblage. Donner une loi approchant la loi de L .
2. Quelle est approximativement la probabilité que l'assemblage ne respecte pas la norme en question ?

Exercice 3 (Calcul de π par Monte Carlo)

En remarquant que $\pi = \int_{-1}^1 \int_{-1}^1 \mathbf{1}_{\mathcal{D}}(x, y) dx dy$, proposer une estimation de π par la méthode de Monte Carlo.

Exercice 4

On reprend l'exercice 3 sur l'évaluation de π par une méthode de Monte Carlo. Ecrire un programme Matlab permettant d'évaluer π . Pour une valeur fixée du nombre d'itérations n , donner la variance de la variable aléatoire qui évalue π . Vérifier que cette variance évolue en $\frac{1}{n}$.

Exercice 5

On reprend l'exercice 2. Un assemblage comprend n sections mises bout à bout. La longueur de chaque section (en centimètres) est une variable aléatoire uniforme sur $[8, 12]$. De plus les sections sont indépendantes.

1. Obtenir par simulation la loi de L pour $n \in \{2, 5, 100\}$. Que visualisez-vous ?
2. Pour $n = 100$, vérifier que $\mathbb{P}[|L - 1000| \geq 30] = 0.0094$.

Appendice

A Annexe : Eléments de théorie des ensembles et de théorie de la mesure

A.1 Ensembles dénombrables

Définition A.1

Un ensemble D est dit dénombrable si l'on peut compter ses éléments, c'est-à-dire s'il existe une application $\varphi : D \rightarrow \mathbb{N}$ injective (un même numéro ne peut correspondre à deux éléments distincts).

Proposition A.1

- Les ensembles \mathbb{N} , \mathbb{Z} sont dénombrables.
- Tout sous-ensemble d'un ensemble dénombrable est dénombrable.
- Le produit cartésien fini d'ensembles dénombrables est dénombrable.
- \mathbb{Q} est dénombrable.
- L'union dénombrable d'ensembles dénombrables est dénombrable.

A.2 Ensembles négligeables. Intégrale des fonctions nulles presque partout.

Définition A.2

Un ensemble $N \subset \mathbb{R}$ est dit négligeable de \mathbb{R} si, pour tout $\varepsilon > 0$, il existe une famille dénombrable d'intervalles de \mathbb{R} recouvrant N et de somme des longueurs inférieure à ε , c'est-à-dire s'il existe une suite $(I_n^\varepsilon)_{n \geq 1}$ d'intervalles de \mathbb{R} satisfaisant :

$$N \subset \bigcup_{n \geq 1} I_n^\varepsilon \quad \text{et} \quad \sum_{n \geq 1} |I_n^\varepsilon| < \varepsilon,$$

où $|I_n^\varepsilon|$ désigne la longueur de l'intervalle I_n^ε .

Remarque :

- Un ensemble ouvert ne peut donc pas être négligeable puisqu'il contient un intervalle de longueur strictement positive.
- Toute union dénombrable de négligeables est négligeable.

Définition A.3

Un ensemble $N \subset \mathbb{R}^n$ ($n \geq 1$) est dit négligeable de \mathbb{R}^n si, pour tout $\varepsilon > 0$, il existe une famille dénombrable de pavés de \mathbb{R}^n recouvrant N et de somme des mesures inférieure à ε , avec la mesure d'un pavé prise comme le produit des longueurs de ses côtés.

Remarque :

- Les remarques ci-dessus s'appliquent à la dimension n .
- Les droites dans \mathbb{R}^2 , les plans dans \mathbb{R}^3 etc. sont négligeables.

Proposition A.2

Soit f une fonction définie sur \mathbb{R}^n , nulle en dehors d'un négligeable (on dit aussi nulle presque partout). Alors l'intégrale de Lebesgue de f est nulle. Réciproquement, si f , positive (respectivement négative) presque partout, est d'intégrale nulle, alors f est nulle presque partout.

A.3 Mesure de Lebesgue des ensembles de \mathbb{R}^n .

On peut montrer que les notions de longueur (dans \mathbb{R}) d'aire (dans \mathbb{R}^2) et de volume (dans \mathbb{R}^3) peuvent se généraliser à toutes dimensions. La mesure ainsi obtenue s'appelle la mesure de Lebesgue. Elle vérifie certaines propriétés que nous rappelons ici.

La mesure d'un pavé de \mathbb{R}^n (produit de n intervalles de \mathbb{R}) est le produit des longueurs des intervalles.

Un ensemble est négligeable si et seulement si sa mesure de Lebesgue est nulle.

B Formulaire : Loïs à densité usuelles

B.1 Loïs, espérances, variances.

Loi uniforme sur $[a, b]$. On dit que X suit une loi uniforme sur $[a, b]$, où $a < b$, et on note

$$X \rightsquigarrow \mathcal{U}([a, b])$$

si X admet pour densité la fonction

$$f(x) = \frac{1}{b-a} \mathbf{1}_{[a,b]}(x) \text{ pour tout } x \in \mathbb{R}.$$

On a

$$\mathbb{E}[X] = \frac{a+b}{2} \text{ et } \mathbf{Var}[X] = \frac{(b-a)^2}{12}.$$

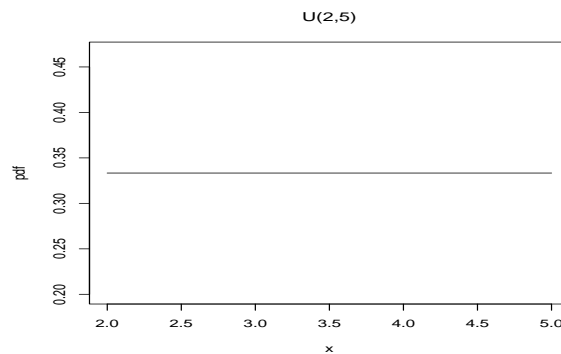


FIGURE 9 – Loi Uniforme

Loi exponentielle. On dit que X suit une loi exponentielle de paramètre $\lambda > 0$ et on note

$$X \rightsquigarrow \mathcal{E}(\lambda)$$

si X admet pour densité la fonction

$$f(x) = \lambda e^{-\lambda x} \mathbf{1}_{]0,+\infty[}(x) \text{ pour tout } x \in \mathbb{R}.$$

On a

$$\mathbb{E}[X] = \frac{1}{\lambda} \text{ et } \mathbf{Var}[X] = \frac{1}{\lambda^2}.$$

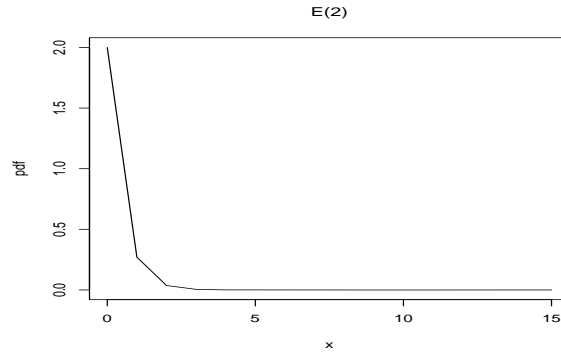


FIGURE 10 – Loi Exponentielle

Loi Gamma. On dit que X suit une loi Gamma de paramètres $\lambda > 0$ et $\alpha > 0$ et on note

$$X \rightsquigarrow \text{Gamma}(\lambda, \alpha)$$

si X admet pour densité la fonction

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \mathbf{1}_{]0, +\infty[}(x) \text{ pour tout } x \in \mathbb{R}$$

avec

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx.$$

On a :

$$\Gamma(\alpha + 1) = \alpha \Gamma(\alpha) \text{ pour tout } \alpha > 0 ;$$

$$\Gamma(n) = (n-1)! \text{ pour tout } n \in \mathbb{N}^*.$$

On en déduit

$$\mathbb{E}[X] = \frac{\alpha}{\lambda} \text{ et } \text{Var}[X] = \frac{\alpha}{\lambda^2}.$$

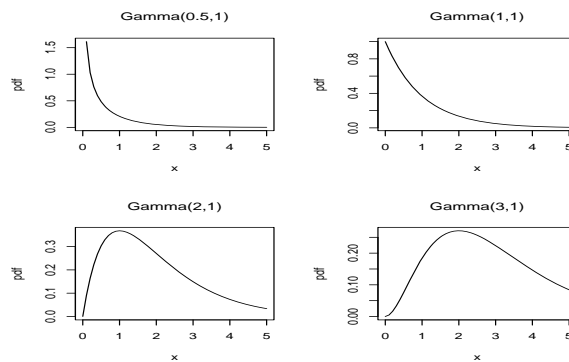


FIGURE 11 – Loi Gamma

Lois Bêta. On dit que X suit une loi Bêta de type I de paramètres $\alpha > 0$ et $\beta > 0$ et on note

$$X \rightsquigarrow \text{bêta } I(\alpha, \beta)$$

si X admet pour densité la fonction

$$f(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} \mathbf{1}_{[0,1]}(x) \text{ pour tout } x \in \mathbb{R}$$

avec

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx.$$

On a :

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \text{ pour tous } \alpha, \beta > 0.$$

On en déduit

$$\mathbb{E}[X] = \frac{\alpha}{\alpha+\beta} \text{ et } \text{Var}[X] = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}.$$

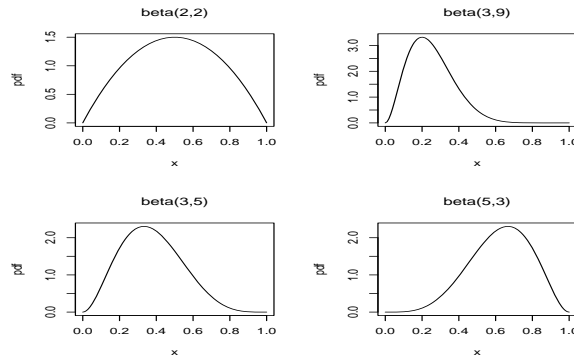


FIGURE 12 – Loi Bêta I

Si X suit une loi Bêta de type I de paramètres $\alpha > 0$ et $\beta > 0$, alors la v.a. $Y = X/(1-X)$ suit une loi Bêta de type II notée $\text{bêta II}(\alpha, \beta)$ et dont la densité s'obtient par changement de variable :

$$g(y) = \frac{1}{B(\alpha, \beta)} \frac{y^{\alpha-1}}{(1+y)^{\alpha+\beta}} \mathbf{1}_{]0,+\infty[}(y).$$

On a :

$$\mathbb{E}[Y] = \frac{\alpha}{\beta-1} \text{ et } \text{Var}[Y] = \frac{\alpha(\alpha+\beta-1)}{(\beta-1)^2(\beta-2)}.$$

Loi de Cauchy. On dit que X suit une loi de Cauchy de paramètre $c > 0$ et on note

$$X \rightsquigarrow \mathcal{C}(c)$$

si X admet pour densité la fonction

$$f(x) = \frac{c}{\pi(c^2 + x^2)} \text{ pour tout } x \in \mathbb{R}.$$

Une telle v.a. n'est pas intégrable.

Loi normale (ou de Gauss ou de Laplace-Gauss). On dit que X suit une loi normale de paramètres $m \in \mathbb{R}$ et $\sigma^2 > 0$ et on note

$$X \rightsquigarrow \mathcal{N}(m, \sigma^2)$$

si X admet pour densité la fonction

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \text{ pour tout } x \in \mathbb{R}.$$

On a :

$$\mathbb{E}[X] = m \text{ et } \text{Var}[X] = \sigma^2.$$

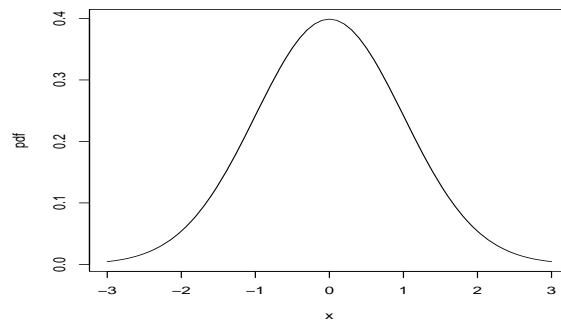


FIGURE 13 – Loi Normale

B.2 Propriétés.

Une loi exponentielle de paramètre λ est une loi Gamma de paramètres λ et 1.

La somme de deux v.a. indépendantes suivant des lois Gamma de paramètres (λ, α) et (λ, β) suit une loi Gamma de paramètres $(\lambda, \alpha + \beta)$.

Si X suit une loi normale de paramètres m et σ^2 alors la v.a.

$$X^* = \frac{X - m}{\sigma}$$

suit une loi normale de paramètres 0 et 1.

La somme de deux v.a. indépendantes suivant des lois normales de paramètres (m_1, σ_1^2) et (m_2, σ_2^2) suit une loi normale de paramètres $m_1 + m_2$ et $\sigma_1^2 + \sigma_2^2$.

C Rappel : Variables Aléatoires Discrètes

C.1 Définition et Propriétés

On redonne dans cette section les expressions pour les cas des variables aléatoires discrètes

Définition C.1

Une variable aléatoire réelle X est dite *discrète* s'il existe une suite $\{x_1, \dots, x_k, \dots\}$ finie ou non de réels et une suite $\{p_1, \dots, p_k, \dots\}$ de $[0, 1]$ telles que, pour tout $k \geq 1$

$$\mathbb{P}[X = x_k] = p_k.$$

La donnée de ces deux suites détermine la loi de la variable aléatoire discrète X .

Une expérience à issues dénombrables modélisée par une partition d'événements $(A_k)_{k \in \mathbb{N}}$ peut être également modélisée par une variable aléatoire à valeurs entières satisfaisant

$$\mathbb{P}[X = k] = \mathbb{P}[A_k], \text{ pour tout } k \in \mathbb{N}.$$

Réciproquement, si X est une variable aléatoire à valeurs dans $\{x_k\}_{k \geq 1}$, alors les événements $([X = x_k])_{k \geq 1}$ forment une partition de Ω . Les propriétés des probabilités entraînent alors que la suite $(p_k)_{k \geq 1}$ de la définition ci-dessus doit satisfaire

$$p_k \geq 0 \text{ pour tout } k \text{ et } \sum_{k \geq 1} p_k = 1.$$

Proposition C.1

Soit X une v.a. discrète prenant ses valeurs dans $\{x_k, k \geq 1\}$ avec les probabilités $(p_k)_{k \geq 1}$. Alors, pour tout $B \subset \mathbb{R}$,

$$\mathbb{P}[X \in B] = \sum_{x_k \in B} p_k.$$

Soit X une v.a. réelle. Si X est une v.a. discrète prenant ses valeurs dans $\{x_1, \dots, x_k, \dots\}$, les valeurs étant distinctes, alors

$$F_X(t) = \sum_{k | x_k \leq t} \mathbb{P}[X = x_k] \text{ pour tout } t \in \mathbb{R}.$$

C.1.1 Espérance

Définition C.2

Soit X est une v.a. prenant ses valeurs dans $\{x_k, k \in \mathbb{N}\}$, alors si

$$\sum_{k \in \mathbb{N}} |x_k| \mathbb{P}[X = x_k] < +\infty$$

on dit que X est intégrable et on définit l'espérance de X par

$$\mathbb{E}[X] = \sum_{k \in \mathbb{N}} x_k \mathbb{P}[X = x_k].$$

C.1.2 Formule de transfert

Proposition C.2 (Formule de transfert.)

Soit X une v.a. sur Ω et à valeurs dans \mathbb{R} . Soit $\varphi : \mathbb{R} \rightarrow \mathbb{R}$. Si X est une v.a. discrète prenant ses valeurs dans

$\{x_k, k \geq 1\}$ et si $\varphi(X)$ est intégrable, c'est-à-dire si

$$\sum_{k \geq 1} |\varphi(x_k)| \mathbb{P}[X = x_k] < +\infty \text{ alors } \mathbb{E}[\varphi(X)] = \sum_{k \geq 1} \varphi(x_k) \mathbb{P}[X = x_k].$$

C.1.3 Variance

Proposition C.3

Si X est une v.a.r. discrète prenant ses valeurs dans $\{x_k, k \geq 1\}$ et si

$$\sum_{k \geq 1} x_k^2 \mathbb{P}[X = x_k] < +\infty$$

alors la variance de X est la quantité

$$\text{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \sum_{k \geq 1} (x_k - \mathbb{E}[X])^2 \mathbb{P}[X = x_k].$$

C.1.4 Somme de v.a. discrètes indépendantes

Proposition C.4 (Cas discret)

Soient X et Y deux v.a.r. indépendantes. On pose $Z = X + Y$. On suppose que X et Y prennent leurs valeurs dans $\{x_i, i \in \mathbb{N}\}$ et $\{y_j, j \in \mathbb{N}\}$ respectivement. On note $Z(\Omega)$ l'ensemble (dénombrable) de toutes les valeurs possibles de $x_i + y_j$ quand $(i, j) \in \mathbb{N}^2$. Alors, pour tout $z \in Z(\Omega)$,

$$\mathbb{P}[Z = z] = \sum_{i \in \mathbb{N}} \mathbb{P}[X = x_i] \mathbb{P}[Y = z - x_i]$$

C.2 Lois usuelles discrètes.

C.2.1 Lois, espérances, variances.

Mesure de Dirac en $x_0 \in \mathbb{R}$. On dit que X suit une mesure de Dirac en $x_0 \in \mathbb{R}$ et on note $X \rightsquigarrow \delta_{x_0}$ si X ne prend que la valeur x_0 . On a alors

$$\mathbb{P}[X = x_0] = 1, \quad \mathbb{E}[X] = x_0, \quad \text{Var}[X] = 0.$$

Loi de Bernoulli. On dit que X suit une loi de Bernoulli de paramètre $p \in [0, 1]$ et on note $X \rightsquigarrow \mathcal{B}(p)$ si X prend les valeurs 0 et 1 et

$$\mathbb{P}[X = 1] = p.$$

On a

$$\mathbb{E}[X] = p \text{ et } \text{Var}[X] = p(1 - p).$$

Remarque, si $p = 0$, $X \rightsquigarrow \delta_0$ et si $p = 1$, $X \rightsquigarrow \delta_1$.

Loi uniforme sur $\{1, \dots, n\}$. On dit que X suit une loi uniforme sur $\{1, \dots, n\}$ et on note $X \rightsquigarrow \mathcal{U}(\{1, \dots, n\})$ si X prend ses valeurs dans $\{1, \dots, n\}$ et

$$\mathbb{P}[X = k] = \frac{1}{n} \text{ pour tout } k \in \{1, \dots, n\}.$$

On a

$$\mathbb{E}[X] = \frac{n+1}{2} \text{ et } \text{Var}[X] = \frac{n^2-1}{12}.$$

Loi binomiale. On dit que X suit une loi binomiale de paramètre $n \geq 1$ et $p \in [0, 1]$ et on note $X \rightsquigarrow \mathcal{B}(n, p)$ si X prend ses valeurs dans $\{0, \dots, n\}$ et

$$\mathbb{P}[X = k] = \binom{n}{k} p^k (1-p)^{n-k} \text{ pour tout } k \in \{0, \dots, n\}.$$

On a

$$\mathbb{E}[X] = np \text{ et } \text{Var}[X] = np(1-p).$$

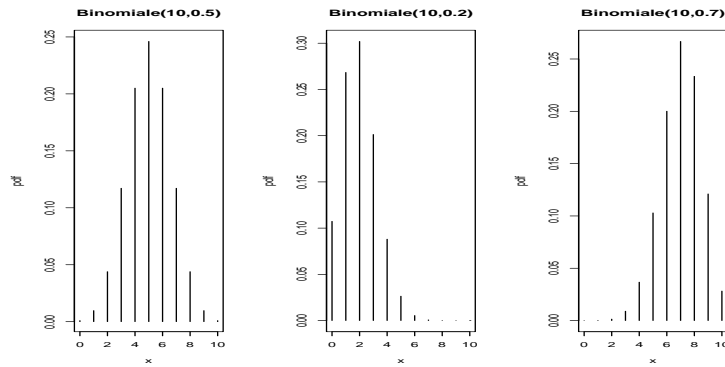


FIGURE 14 – Loi Binomiale

Loi hypergéométrique. On dit que X suit une loi hypergéométrique de paramètres $N \geq 1, n \geq 1, p \in [0, 1]$ tels que $pN \in \mathbb{N}$ et $n \leq \min(pN, (1-p)N)$, et on note

$$X \rightsquigarrow \mathcal{H}(N, n, p)$$

si X prend ses valeurs dans $\{0, \dots, n\}$ et

$$\mathbb{P}[X = k] = \frac{\binom{pN}{k} \binom{(1-p)N}{n-k}}{\binom{N}{n}}.$$

On a

$$\mathbb{E}[X] = np \text{ et } \text{Var}[X] = \frac{N-n}{N-1} np(1-p).$$

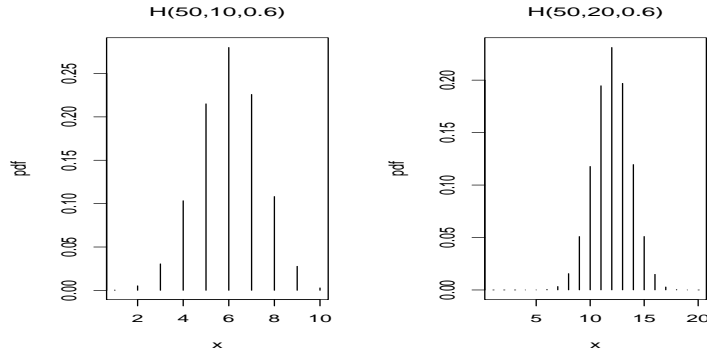


FIGURE 15 – Loi Hypergémométrique

Loi de Poisson. On dit que X suit une loi de Poisson de paramètre $\lambda > 0$ et on note

$$X \rightsquigarrow \mathcal{P}(\lambda)$$

si X prend ses valeurs dans \mathbb{N} et

$$\mathbb{P}[X = k] = e^{-\lambda} \frac{\lambda^k}{k!} \text{ pour tout } k \in \mathbb{N}.$$

On a

$$\mathbb{E}[X] = \lambda \text{ et } \text{Var}[X] = \lambda.$$

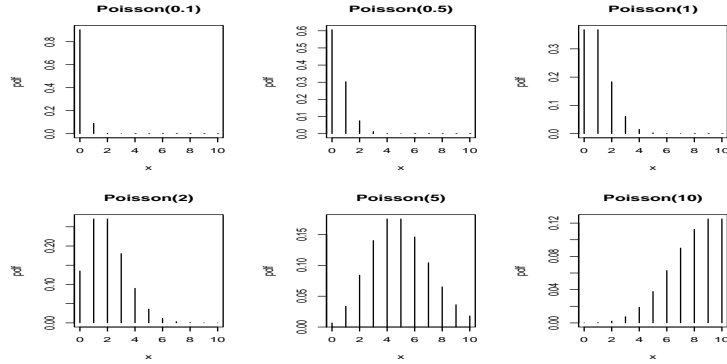


FIGURE 16 – Loi de Poisson

Loi géométrique. On dit que X suit une loi géométrique de paramètre $p \in [0, 1]$ et on note

$$X \rightsquigarrow \mathcal{G}(p)$$

si X prend ses valeurs dans \mathbb{N}^* et

$$\mathbb{P}[X = k] = p(1 - p)^{k-1} \text{ pour tout } k \in \mathbb{N}^*.$$

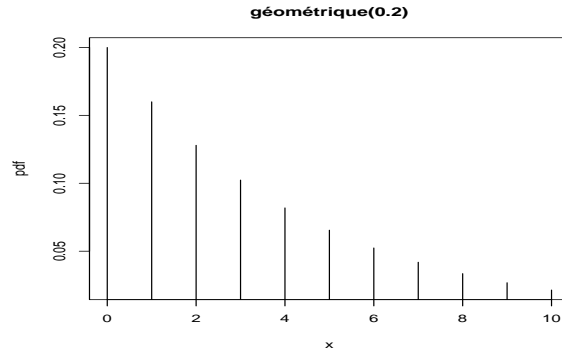


FIGURE 17 – Loi Géométrique

On a

$$\mathbb{E}[X] = \frac{1}{p} \text{ et } \text{Var}[X] = \frac{1-p}{p^2}.$$

Loi de Pascal. On dit que X suit une loi de Pascal de paramètres $n \in \mathbb{N}^*$ et $p \in [0, 1]$ et on note

$$X \rightsquigarrow Pa(n, p)$$

si X prend ses valeurs dans $\{k \in \mathbb{N} / k \geq n\}$ et

$$\mathbb{P}[X = k] = \binom{n-1}{k-1} p^n (1-p)^{k-n} \text{ pour tout } k \geq n.$$

On a

$$\mathbb{E}[X] = \frac{n}{p} \text{ et } \text{Var}[X] = \frac{n(1-p)}{p^2}.$$

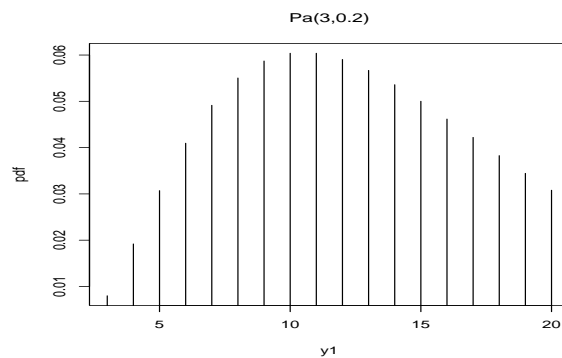


FIGURE 18 – Loi de Pascal

C.2.2 Propriétés.

Une loi de Bernoulli de paramètre p est une loi binomiale de paramètres 1 et p .

La somme de 2 v.a. indépendantes suivant des lois binomiales de paramètres (n_1, p) et (n_2, p) respectivement suit une loi binomiale de paramètres $n_1 + n_2$ et p .

La somme de deux v.a. indépendantes suivant respectivement des lois de Poisson de paramètre λ_1 et λ_2 suit une loi de Poisson de paramètre $\lambda_1 + \lambda_2$.

Une loi géométrique de paramètre p est une loi de Pascal de paramètres 1 et p .

La somme de 2 v.a. indépendantes suivant des lois de Pascal de paramètres (n_1, p) et (n_2, p) respectivement suit une loi de Pascal de paramètres $n_1 + n_2$ et p .

Proposition C.5

Convergence de la loi hypergéométrique vers la loi binomiale. On suppose que, pour tout $k \geq 1$, X_k suit une loi hypergéométrique de paramètres N_k, n, p_k (avec $p_k N_k \in \mathbb{N}^*$) et que

$$\lim_{k \rightarrow +\infty} N_k = +\infty, \quad \lim_{k \rightarrow +\infty} p_k = p \in]0, 1[.$$

Alors

$$\mathcal{L}(X_k) \rightarrow \mathcal{B}(n, p) \text{ quand } k \rightarrow +\infty.$$

Application. Approximation de la loi hypergéométrique par la loi binomiale. Dans la pratique, si $X \rightsquigarrow \mathcal{H}(N, n, p)$, on peut approcher sa loi par une $\mathcal{B}(n, p)$ dès que $10n < N$.

Proposition C.6

Convergence de la loi binomiale vers la loi de Poisson. On suppose que, pour tout $n \geq 1$, X_n suit une loi binomiale de paramètres n, p_n et que

$$\lim_{n \rightarrow +\infty} np_n = \lambda > 0.$$

Alors

$$\mathcal{L}(X_n) \rightarrow \mathcal{P}(\lambda) \text{ quand } n \rightarrow +\infty.$$

Application. Approximation de la loi binomiale par la loi de Poisson. Dans la pratique, si $X \rightsquigarrow \mathcal{B}(n, p)$, on peut approcher sa loi par une $\mathcal{P}(np)$ dès que $n \geq 30$ et $p \leq 0, 1$.

C.3 Simulation de variables aléatoires discrètes

C.3.1 Simulation d'une v.a. de Bernoulli

Pour simuler une v.a. de Bernoulli de paramètre $p \in]0, 1[$, il suffit de simuler une v.a. U uniforme sur $]0, 1[$ et de poser $X = \mathbf{1}_{[U \leq p]}$ (i.e. $X = 1$ si $U \leq p$ — donc avec probabilité p — et 0 sinon).

C.3.2 Simulation d'une v.a. binomiale

Pour simuler une v.a. binomiale de paramètres $n \geq 1$ et $p \in]0, 1[$, il suffit de sommer n v.a. de Bernoulli de paramètre p indépendantes.

C.3.3 Simulation d'une v.a. discrète

Pour simuler une v.a. discrète à valeurs dans $\{x_i, i \in \mathbb{N}\}$ de loi $p_i, i \in \mathbb{N}$, il suffit de simuler une v.a. U uniforme sur $]0, 1[$ et de poser $X = x_i$ si $p_1 + \cdots + p_{i-1} \leq U < p_1 + \cdots + p_i$ où $p_1 + P_0 = 0$.