

Variables aléatoires à densité

On étudie dans ce chapitre des variables aléatoires pouvant prendre une infinité *continue* de valeurs (par exemple $X(\Omega) = \mathbb{R}$ ou \mathbb{R}_+).

La distribution de probabilité est alors modélisée différemment : la plupart du temps, il ne s'agira plus de manipuler la probabilité qu'une variable X soit égale à une valeur donnée (intuitivement, on peut se convaincre qu'une telle probabilité sera presque toujours nulle) ; mais la probabilité que X appartienne à un intervalle donné. Cette information sera fournie par la *fonction de répartition*, qui caractérise complètement la loi d'une variable.

1 Rappels

1.1 Définitions : espace probabilisé, variable aléatoire

On dispose dans tout ce cours d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Dans cette notation, Ω est « l'univers » (on ne rentrera jamais dans les détails), \mathcal{A} est l'ensemble des événements (parties de l'univers dont on peut calculer la probabilité), et \mathbb{P} est la fonction probabilité, qui à un événement $A \in \mathcal{A}$ associe sa probabilité $\mathbb{P}(A)$.

Comme dans le cas des variables discrètes, on définit alors une *variable aléatoire* comme une fonction définie sur l'univers : à chaque issue $\omega \in \Omega$ est associé une valeur de la variable X , notée $X(\omega)$.

On demande quand même une propriété pour pouvoir effectuer des calculs :

Définition 1. On appelle variable aléatoire réelle (en abrégé : *var*) toute fonction X définie sur Ω , à valeurs dans \mathbb{R} telle que :

$$\forall x \in \mathbb{R}, \{ \omega \in \Omega \mid X(\omega) \leq x \} \text{ est un événement.}$$

On introduit traditionnellement des raccourcis de notation pour désigner les événements utilisés dans les calculs : si $x \in \mathbb{R}$, on note :

$$(X \leq x) = \{ \omega \in \Omega \mid X(\omega) \leq x \}$$

$$(X \geq x) = \{ \omega \in \Omega \mid X(\omega) \geq x \}$$

$$(X = x) = \{ \omega \in \Omega \mid X(\omega) = x \}$$

Dit en français : l'événement $(X \leq x)$ est l'ensemble des issues ω telles que la valeur prise par X pour cette issue est inférieure ou égale à x .

Les propriétés d'un espace probabilisé assurent que si tous les $(X \leq x)$ sont des événements, alors tous les $(X \geq x)$, $(X < x)$, $(X > x)$, $(X = x)$... plus généralement tous les $(X \in I)$ où I est un intervalle de \mathbb{R} sont aussi des événements.

Tout ceci pose les fondements de la *théorie de la mesure*, ce qui n'est pas vraiment le propos dans ce cours. Pour nous, cela assure simplement que :

Si X est une variable aléatoire, toutes les probabilités $\mathbb{P}(X \in I)$, avec I un intervalle de \mathbb{R} , sont bien définies.

La *fonction de répartition* est ici l'objet mathématique qui donne la loi d'une variable :

Définition 2. Soit X une variable aléatoire. On appelle fonction de répartition de X la fonction F_X définie sur \mathbb{R} par :

$$\forall x \in \mathbb{R}, F_X(x) = \mathbb{P}(X \leq x)$$

Remarque 1. Attention, une fonction de répartition est toujours définie sur \mathbb{R} , même si $X(\Omega)$ n'est pas égal à \mathbb{R} entier. Dans le cas très fréquent où $X(\Omega) = \mathbb{R}_+$, F_X est bel et bien définie sur \mathbb{R}_+ ; et on a : $\forall x < 0, F_X(x) = 0$. On peut plus généralement voir que si $X(\Omega) \subset [a, b]$:

- $\forall x < a, F_X(x) = 0$ (car si $x < a$, l'événement $(X \leq x)$ est impossible) ;
- $\forall x \geq b, F_X(x) = 1$ (car si $x \geq b$, l'événement $(X \leq x)$ est certain).

Ceci servira très souvent en pratique.

De manière rigoureuse, la loi d'une variable aléatoire est donnée par l'ensemble des $\mathbb{P}(X \in A)$, où A est une partie pouvant être écrite comme une réunion dénombrable d'intervalles. C'est assez peu maniable. Le théorème suivant montre que la fonction de répartition contient toute l'information pour définir la loi de X :

Théorème 1. *La donnée de la fonction de répartition F_X caractérise la loi de la variable aléatoire X . En particulier, pour tous $(a, b) \in \mathbb{R}^2$ tels que $a < b$, on a*

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

Remarque 2. Faire attention à la position des inégalités $<$ et \leq dans la formule précédent.

Démonstration. Le premier point est admis.

On a : $(X \leq b) = (X \leq a) \sqcup (a < X \leq b)$; d'où en passant aux probabilités : $\mathbb{P}(X \leq b) = \mathbb{P}(X \leq a) + \mathbb{P}(a < X \leq b)$. On obtient bien $\mathbb{P}(a < X \leq b) = \mathbb{P}(X \leq b) - \mathbb{P}(X \leq a) = F_X(b) - F_X(a)$. \square

Rappelons les propriétés d'une fonction de répartition. Certaines d'entre elles sont issues des *théorèmes de limite monotone* rappelés en annexe (mais hors-programme, ainsi que les démonstrations qui les utilisent).

Proposition 2. *Soit X une variable aléatoire de fonction de répartition F_X . Alors :*

- F_X est croissante sur \mathbb{R} .
 - F_X est continue à droite en tout point de \mathbb{R} (c'est-à-dire, en tout point x , $\mathbb{P}(X \leq x) = \lim_{t \rightarrow x^+} F_X(t)$).
 - $\lim_{x \rightarrow -\infty} F_X(x) = 0$.
 - $\lim_{x \rightarrow +\infty} F_X(x) = 1$.
 - En tout point x , $\mathbb{P}(X < x) = \lim_{t \rightarrow x^-} F_X(t)$.
 - En tout point x , $\mathbb{P}(X = x) = \lim_{t \rightarrow x^+} F_X(t) - \lim_{t \rightarrow x^-} F_X(t)$.
- En particulier, F_X est continue en x si et seulement si $\mathbb{P}(X = x) = 0$.*

Démonstration. En annexe. \square

1.2 Problème inverse : quand une fonction est-elle une fonction de répartition ?

Nous avons jusqu'ici considéré comme point de départ une variable aléatoire X ; puis défini sa fonction de répartition ; puis, à l'aide des propriétés d'une probabilité, obtenu des propriétés sur cette fonction. On souhaite ici effectuer l'opération inverse : une fonction de \mathbb{R} dans \mathbb{R} étant donnée, à quelle condition peut-on affirmer que c'est la fonction de répartition d'une certaine variable aléatoire ?

Ce résultat sera admis.

Théorème 3. Soit F une fonction de \mathbb{R} dans \mathbb{R} :

- croissante sur \mathbb{R} ;
- telle que $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$
- continue à droite en tout point de \mathbb{R} .

Alors F est une fonction de répartition, c'est-à-dire qu'il existe une variable aléatoire X dont F est la fonction de répartition.

1.3 Rappel : fonction de répartition d'une VAD

La notion de fonction de répartition n'est pas spécifique aux variables à densité : elle existe aussi pour les variables discrètes. Il est d'ailleurs utile de se souvenir de quelques résultats à ce propos :

- **Fonction de répartition d'une variable certaine égale à a .**

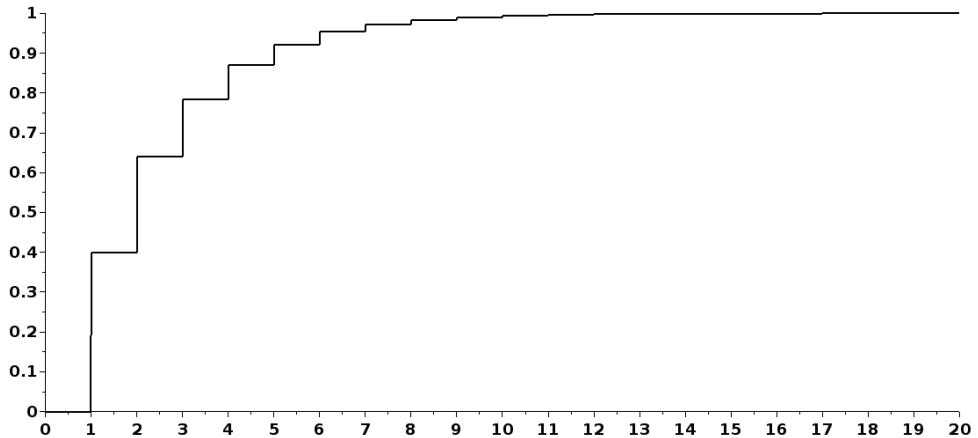
Si $(X = a)$ est certain, alors l'événement $(X \leq x)$ est de proba nulle si $x < a$, et de proba 1 si $x \geq a$. La fonction de répartition de X est alors un créneau en a :

- **Fonction de répartition d'une variable de Bernoulli :** si X vaut 0 avec une proba $1 - p$ et 1 avec une proba p , et $x \in \mathbb{R}$, alors on voit que :

- si $x < 0$, $\mathbb{P}(X \leq x) = 0$
- Si $0 \leq x < 1$, alors $\mathbb{P}(X \leq x) = \mathbb{P}(X = 0) = 1 - p$
- Si $x \geq 1$, alors $\mathbb{P}(X \leq x) = 1$

On obtient cette fois un « double créneau » :

- De manière générale, la fonction de répartition d'une variable discrète se présente comme un escalier, dont les marches sont situées aux valeurs de $X(\Omega)$ (donc éventuellement en nombre infini). Voici par exemple la représentation graphique de la fonction de répartition d'une variable suivant une loi géométrique ($p = 0.4$) :



Les tailles des marches correspondent à la loi de proba $\mathcal{G}(0.4)$:

- $\mathbb{P}(X = 1) = (0.6)^0 \times 0.4 = 0.4$
- $\mathbb{P}(X = 2) = (0.6)^1 \times 0.4 = 0.24$
- etc...

1.4 Variables à densité

Soit X une variable aléatoire, et F_X sa fonction de répartition.

La proposition 2 montre que si a est tel que $\mathbb{P}(X = a) > 0$, alors F_X est discontinue en a .

Si au contraire F_X est continue, on en déduit que pour tout réel x on a $\mathbb{P}(X = x) = 0$. Les $\mathbb{P}(X = x)$ ne sont donc pas, ici, les quantités pertinentes pour effectuer des calculs de probabilité.

On a par contre vu qu'on pouvait déduire de F_X la probabilité que X soit dans un intervalle $[x, y]$. Si cet intervalle est petit, on obtient une probabilité que X soit « très proche de x » : ce sera $F_X(x + h) - F_X(x)$, avec h « infiniment petit ».

Si on suppose enfin que F_X est dérivable, cette dernière probabilité peut s'approximer par $hF'_X(x)$: on voit donc qu'on peut estimer la probabilité que X soit autour de x à l'aide de F'_X . Lorsque cette dérivée existe, on la nomme *densité*.

Définition 3. Soit X une variable aléatoire. On dit que X est à densité (ou : X admet une densité)ssi sa fonction de répartition F_X est :

- continue sur \mathbb{R} ;
- de classe \mathcal{C}^1 sur \mathbb{R} , sauf éventuellement en un nombre fini de points.

On appelle alors densité de X toute fonction f définie sur \mathbb{R} telle que $F'_X = f$, sauf éventuellement en un nombre fini de points.

Méthode :

- Pour montrer qu'une fonction donnée est *une fonction de répartition*, il faudra montrer que cette fonction :
 - est croissante sur \mathbb{R} ;
 - vérifie $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$;
 - est continue à droite en tout point de \mathbb{R} .
- Pour montrer qu'une fonction donnée est *la fonction de répartition d'une variable à densité*, il faudra ajouter les propriétés supplémentaires mentionnées dans la définition 3. Donc :
 - croissante sur \mathbb{R} ;
 - $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$;
 - continue sur \mathbb{R} et de classe \mathcal{C}^1 sauf en un nombre fini de points.
- Pour montrer qu'une variable X donnée est une *variable à densité*, il faudra montrer que sa fonction de répartition est :
 - continue sur \mathbb{R} ;
 - de classe \mathcal{C}^1 sauf en un nombre fini de points.

(NB : pour gagner du temps, on pourra montrer qu'elle est \mathcal{C}^1 sur \mathbb{R} sauf en un nombre fini de points ; et qu'en ces points « problématiques » elle est continue.)

Remarque 3. Remarquer ce critère du « sauf éventuellement en un nombre fini de points ». Il interviendra souvent en pratique et ne doit pas être traité à la légère.

Définition 4. Soit X une variable aléatoire admettant une densité. On dit que f_X est une densité de X ssi :

- f_X est définie sur \mathbb{R} , à valeurs positives ;
- $F'_X = f_X$ sauf éventuellement en un nombre fini de points.

Remarque 4. On voit donc qu'une densité est continue sur \mathbb{R} , sauf éventuellement en un nombre fini de points.

Remarque 5. Une variable aléatoire admet une unique fonction de répartition (qui définit sa loi) ; mais une variable à densité admet plusieurs densités car on peut changer la valeur de f_X en un nombre fini de points (arbitraires) et obtenir une autre fonction qui vérifie aussi les propriétés demandées.

En pratique ceci intervient aux points où F_X n'est pas dérivable : on ne peut pas définir $f_X(a) = F'_X(a)$, et on **pose** alors $f_X(a) = \alpha$ où α est **un réel quelconque**. Sa valeur n'aura aucune influence sur les valeurs de diverses probabilités : il faut plutôt voir une densité comme un outil de calcul. Pour des raisons de régularité que nous évoquerons dans la remarque 8, on aura intérêt à prolonger par continuité (au moins d'un côté) la densité obtenue.

Méthode :

Pour déterminer une densité d'une variable dont on connaît la fonction de répartition, il faut :

- dériver cette fonction de répartition aux endroits où c'est possible (donc partout sauf en un nombre fini de points) ;
- et compléter en ces points par des valeurs de notre choix.

Exemple 1. Soit X une variable aléatoire dont la fonction de répartition est $F_X : t \mapsto \begin{cases} 0 & \text{si } t < 0 \\ 1 - e^{-t} & \text{si } t \geq 0 \end{cases}$. Montrer que X est une variable à densité, et en donner une densité.

L'existence d'une densité permet d'utiliser le calcul intégral pour déterminer les quantités utiles :

Proposition 4 (Propriétés des variables à densité).

Soit X une variable aléatoire admettant une densité f . On a les propriétés suivantes :

- Pour tout $x \in \mathbb{R}$, $\mathbb{P}(X = x) = 0$.
 - Si I est un intervalle de \mathbb{R} , la quantité $\mathbb{P}(X \in I)$ ne change pas si on ouvre ou ferme les crochets aux bornes de I .
 - Pour tout $x \in \mathbb{R}$,
- $$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X < x) = \int_{-\infty}^x f(t) dt$$
- Pour tout $x \in \mathbb{R}$,
- $$\mathbb{P}(X \geq x) = \mathbb{P}(X > x) = 1 - F_X(x) = \int_x^{+\infty} f(t) dt$$
- Pour tous $(a, b) \in \mathbb{R}^2$ tels que $a \leq b$:
- $$\mathbb{P}(a \leq x \leq b) = \mathbb{P}(a < x \leq b) = \mathbb{P}(a \leq x < b) = \mathbb{P}(a < x < b) = \int_a^b f(t) dt$$
- $\int_{-\infty}^{+\infty} f(t) dt$ converge, et est égale à 1.

Démonstration. En annexe. □

Remarque 6. On voit dans cet encadré justifie que dans beaucoup de calculs de probabilités faisant intervenir des variables à densité, on peut remplacer les inégalités larges par des strictes, et inversement. Attention par contre : si X et Y sont deux variables à densité, on n'a pas forcément $\mathbb{P}(X \leq Y) = \mathbb{P}(X < Y)$!! (contre-exemple : prendre $Y = X$).

Remarque 7. La dernière propriété $\int_{-\infty}^{+\infty} f(x) dx = 1$ est l'analogue continu de la propriété $\sum_{x \in X(\Omega)} \mathbb{P}(X = x) = 1$ vue pour des variables discrètes.

Remarque 8 (Subtile mais importante !). La propriété $F_X(x) = \int_{-\infty}^x f(t) dt$ laisse penser, en se souvenant du théorème fondamental de l'analyse, que $F'_X = f_X$. Mais on voit apparaître une contradiction : la fonction de répartition d'une variable aléatoire est unique, alors qu'une densité ne l'est pas (on peut la modifier arbitrairement en un nombre fini de points).

Il faut donc prendre plusieurs précautions :

- Si la densité f_X utilisée est continue sur \mathbb{R} , alors le TFA affirme bien que F_X est \mathcal{C}^1 sur \mathbb{R} ;
- Cela fonctionne en fait localement : en tout point x où la densité f_X est continue, la fonction de répartition est \mathcal{C}^1 , et $F'_X(x) = f_X(x)$.

Réiproquement, si notre point de départ est la fonction de répartition (\mathcal{C}^0 , \mathcal{C}^1 sauf en un nombre fini de points) : en tout point a où F_X est \mathcal{C}^1 on prendra $f_X(a) = F'_X(a)$; et en le nombre fini de points où on n'a pas le caractère \mathcal{C}^1 , on prendra la valeur de f_X la moins embêtante possible (voir exemples).

Remarque 9. Le troisième point de la proposition 4 montre que :

Si deux variables aléatoires admettent une même densité, alors elles suivent la même loi.

On a une réciproque de ce résultat : si une fonction f possède les propriétés d'une densité, on peut effectivement définir une variable X dont f sera une densité.

Proposition 5. (Définition / proposition)

On appelle **densité** toute fonction f définie sur \mathbb{R} :

- *positive* ;
- *continue sauf éventuellement en un nombre fini de points* ;
- *telle que $\int_{-\infty}^{+\infty} f(x) dx$ converge et est égale à 1*.

Si f est une densité, il existe une variable aléatoire X , admettant f pour densité.

Démonstration. Admis. Cette variable a alors pour fonction de répartition $F_X : x \mapsto \int_{-\infty}^x f(t) dt$. □

Méthode :

Pour montrer qu'une fonction donnée est *une densité de probabilité*, il faudra montrer que cette fonction vérifie les trois propriétés énoncées dans la proposition 5.

1.5 Parenthèse : intégrales de fonctions continues sauf en un nombre fini de points

On manipule dans la section précédente des intégrales de fonctions *non continues*, ce qui nous place hors du formalisme habituel du calcul intégral (ce sera en plus très fréquent en pratique). La solution est alors de découper l'intégrale : soient $a < c < b$ (a et b pouvant être infinis) et f continue sur $[a, b]$ (éventuellement crochets ouverts si les bornes sont improprees), sauf en c . On *défini* :

$$\int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt$$

La borne c est en fait impropre dans une des deux intégrales : par exemple si $f(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ e^{-x} & \text{si } x > 0 \end{cases}$, on écrira (les intégrales qui suivent convergent) :

$$\int_{-\infty}^{+\infty} f(x) dx = \int_{-\infty}^0 0 dx + \int_0^{+\infty} e^{-x} dx = \int_0^{+\infty} e^{-x} dx$$

Mais dans la deuxième intégrale, on intègre sur $]0, +\infty[$ ce qui pose le problème de la borne impropre $\int_0^{+\infty}$.

En fait dans ce cas¹, la fonction $x \mapsto e^{-x}$ définie sur \mathbb{R}_+^* se prolonge par continuité sur \mathbb{R}_+ car elle a une limite finie en 0^+ . L'intégrale à calculer est alors égale à celle du prolongement² : on fait comme si on intégrait sur $[0, +\infty[$.

Moralité : la manière « raisonnable » de traiter ces intégrales fonctionne. De plus vous pourrez passer tout cela sous silence sur une copie et faire vos calculs sans inquiétude.

Exemple 2. Soit $f : x \mapsto \begin{cases} 0 & \text{si } t < 1 \\ \frac{1}{t^2} & \text{si } t \geq 1 \end{cases}$.

Montrer que f est une densité de probabilité. Soit X une variable admettant f pour densité ; donner la fonction de répartition de X .

¹Et dans le cadre du programme ce sera toujours ainsi.

²Ce n'est pas évident.

1.6 Lois usuelles

Nous donnons ici les définitions et propriétés des lois à densité que nous renconterons le plus souvent :

Loi uniforme

Définition 5. Soient $a < b$ deux réels. On dit que X suit la loi uniforme sur $[a, b]$ si X est une variable à densité, dont une densité est donnée par :

$$f : x \mapsto \begin{cases} 0 & \text{si } x < a \text{ ou } x > b \\ \frac{1}{b-a} & \text{si } a \leq x \leq b \end{cases}$$

Dans ce cas, on note : $X \sim \mathcal{U}([a, b])$.

En intégrant cette dernière fonction, on retrouve la fonction de répartition :

Proposition 6. Si $X \sim \mathcal{U}([a, b])$, la fonction de répartition de X est

$$\mathbb{P}(X \leq x) = \begin{cases} 0 & \text{si } x \leq a \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ 1 & \text{si } x \geq b \end{cases}$$

Représentations graphiques :

Densité :

Fonction de répartition :

Loi exponentielle

Définition 6. Soit λ un réel > 0 . On dit que la variable aléatoire X suit la loi exponentielle de paramètre λ si X est une variable à densité, dont une densité est donnée par

$$f : x \mapsto \begin{cases} 0 & \text{si } x < 0 \\ \lambda e^{-\lambda x} & \text{si } x \geq 0 \end{cases}$$

Dans ce cas, on note : $X \sim \mathcal{E}(\lambda)$

La fonction de répartition s'obtient ici aussi en intégrant :

Proposition 7. Si $X \sim \mathcal{E}(\lambda)$, on a

$$\mathbb{P}(X \leq x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - e^{-\lambda x} & \text{si } x \geq 0 \end{cases}$$

Représentations graphiques :

Densité :

Fonction de répartition :

Loi normale

Définition 7. Soient $\mu \in \mathbb{R}$, $\sigma > 0$. On dit que la variable aléatoire X suit la loi normale de paramètres μ, σ^2 ssi X est une variable à densité, dont une densité est donnée par

$$f: x \mapsto \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

On note $X \sim \mathcal{N}(\mu, \sigma^2)$.

On dit que la variable est centrée si $\mu = 0$, et centrée réduite si de plus $\sigma = 1$. La loi normale centrée réduite a donc pour densité :

$$f_{\text{red}}: x \mapsto \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

Dans le cas d'une loi normale, la fonction de répartition ne peut pas s'écrire à l'aide de fonctions usuelles : on peut seulement utiliser la définition et écrire :

Proposition 8. Si $X \sim \mathcal{N}(\mu, \sigma^2)$, on a :

$$\forall x \in \mathbb{R}, \mathbb{P}(X \leq x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt$$

$$\forall (a, b) \in \mathbb{R}^2, \mathbb{P}(a \leq X \leq b) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_a^b \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt$$

Notamment, dans le cas de la loi centrée réduite, la fonction de répartition est notée Φ et est donnée par :

Corollaire 9. Si $X \sim \mathcal{N}(0, 1)$, on a :

$$\forall x \in \mathbb{R}, \Phi(x) = \mathbb{P}(X \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{t^2}{2}\right) dt$$

Représentations graphiques pour la loi normale

Densité :

Quelques indicateurs sur cette « cloche » :

- La proba que X s'écarte de son espérance d'au plus $\pm\sigma$ est $\simeq 68\%$
- La proba que X s'écarte de son espérance d'au plus $\pm 2\sigma$ est $\simeq 95\%$
- La proba que X s'écarte de son espérance d'au plus $\pm 3\sigma$ est $\simeq 99.7\%$
- La « largeur typique » de la cloche est donnée par σ : en fait la largeur est $\pm\sigma$ à 60% de la hauteur de la cloche.

Fonction de répartition :

1.7 Espérance

Le calcul intégral permet de définir (sous réserve d'existence) l'espérance d'une variable à densité :

Définition 8. Soit X une variable aléatoire de densité f .

X admet une espérance si et seulement si l'intégrale $\int_{-\infty}^{+\infty} xf(x) dx$ est **absolument** convergente, et on définit alors :

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} xf(x) dx$$

NB : même s'il faut examiner la convergence de $\int_{-\infty}^{+\infty} |x|f(x) dx$ pour établir l'existence de $\mathbb{E}(X)$, on a $\mathbb{E}(X) = \int_{-\infty}^{+\infty} xf(x) dx$.

Théorème 10 (Espérances des lois usuelles).

- Si $X \hookrightarrow \mathcal{U}([a, b])$, X admet une espérance, et on a $\mathbb{E}(X) = \frac{a+b}{2}$
- Si $X \hookrightarrow \mathcal{E}(\lambda)$, X admet une espérance, et on a $\mathbb{E}(X) = \frac{1}{\lambda}$
- Si $X \hookrightarrow \mathcal{N}(0, 1)$, X admet une espérance, et on a $\mathbb{E}(X) = 0$

Exercice : Retrouver ces valeurs.

2 Résultats additionnels : moments, transfert

2.1 Théorème de transfert

Le théorème suivant permet de calculer (sous réserve d'existence) l'espérance d'une fonction quelconque de la variable X :

Théorème 11 (Théorème de transfert).

Soit X une variable aléatoire admettant une densité f . On suppose que

- f est nulle en dehors d'un intervalle $]a, b[$ (a, b finis ou infinis) ;
- g est continue sur $]a, b[$, sauf éventuellement en un nombre fini de points.

Alors $g(X)$ admet une espérancessi $\int_a^b g(x)f(x) dx$ converge absolument, et dans ce cas :

$$\mathbb{E}(g(X)) = \int_a^b g(x)f(x) dx$$

Exemple 3.

1. Soit $X \hookrightarrow \mathcal{U}([0, 1])$. Calculer $\mathbb{E}(X^2)$.
2. Soit $Y \hookrightarrow \mathcal{N}(0, 1)$. Calculer $\mathbb{E}(Y^2)$.

2.2 Moments

On définit (sous réserve d'existence) de nouvelles quantités associées à une var :

Définition 9. Soit X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$. On dit que X admet un moment d'ordre $r \in \mathbb{N}^*$ si et seulement si l'intégrale $\int_{-\infty}^{+\infty} x^r f(x) dx$ est **absolument convergente**, et on a alors :

$$m_r(X) = \int_{-\infty}^{+\infty} x^r f(x) dx$$

Remarque 10. On a donc : $\forall r \in \mathbb{N}$, $m_r(X) = \mathbb{E}(X^r)$.

Concernant l'existence de ces moments, on a un résultat similaire à celui sur les VAD :

Proposition 12. Soit X une variable aléatoire de densité f . Si f admet un moment d'ordre $m \in \mathbb{N}^*$, alors elle admet des moments d'ordre p pour tout $p \in \llbracket 1, m \rrbracket$.

Notamment, si une variable admet un moment d'ordre 2, elle admet une espérance.

Démonstration. En annexe. □

Dans le cas de l'existence d'un moment d'ordre 2, on définit alors la variance d'une variable à densité :

Définition 10. Soit X une variable aléatoire de densité f admettant un moment d'ordre 2. Alors elle admet une espérance ; et la variable $(X - \mathbb{E}(X))^2$ admet une espérance. On note

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

$V(X)$ est appelée variance de X .

Sous les hypothèses ci-dessus, le calcul suivant est légitime :

$$\begin{aligned} \mathbb{E}((X - \mathbb{E}(X))^2) &= \int_{-\infty}^{+\infty} (x - \mathbb{E}(X))^2 f(x) dx \\ &= \int_{-\infty}^{+\infty} x^2 f(x) dx - 2\mathbb{E}(X) \int_{-\infty}^{+\infty} x f(x) dx + \mathbb{E}(X)^2 \int_{-\infty}^{+\infty} f(x) dx \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}(X)^2 \\ \mathbb{E}((X - \mathbb{E}(X))^2) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \end{aligned}$$

et on obtient :

Proposition 13 (Koenig-Huygens). Soit X une variable à densité admettant une variance. On a

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

Théorème 14 (Variances des lois usuelles).

- Si $X \hookrightarrow \mathcal{U}([a, b])$, X admet une variance, et on a $V(X) = \frac{(b-a)^2}{12}$
- Si $X \hookrightarrow \mathcal{E}(\lambda)$, X admet une variance, et on a $V(X) = \frac{1}{\lambda^2}$
- Si $X \hookrightarrow \mathcal{N}(0, 1)$, X admet une variance, et on a $V(X) = 1$

2.3 Valeurs d'intégrales usuelles

Les résultats précédents peuvent être invoqués pour justifier l'existence d'intégrales et donner leur valeur.

- Par propriété d'une densité, celle de la loi exponentielle vérifie :

$$\forall \lambda > 0, \int_0^{+\infty} \lambda e^{-\lambda x} dx = 1$$

On en déduit que :

$$\forall \lambda > 0, \text{ l'intégrale } \int_0^{+\infty} e^{-\lambda x} dx \text{ converge et vaut } \frac{1}{\lambda}$$

- $X \hookrightarrow \mathcal{E}(\lambda)$ admet une espérance égale à $\frac{1}{\lambda}$. On en déduit en effectuant le calcul que $\int_0^{+\infty} \lambda x e^{-\lambda x} dx = \frac{1}{\lambda^2}$, ou encore

$$\forall \lambda > 0, \text{ l'intégrale } \int_0^{+\infty} x e^{-\lambda x} dx \text{ converge et vaut } \frac{1}{\lambda^2}$$

- $X \hookrightarrow \mathcal{E}(\lambda)$ admet une variance égale à $\frac{1}{\lambda^2}$. On a donc $\mathbb{E}(X^2) = V(X) + \mathbb{E}(X)^2 = \frac{2}{\lambda^2}$. On en déduit en effectuant le calcul que $\int_0^{+\infty} \lambda x^2 e^{-\lambda x} dx = \frac{2}{\lambda^3}$, ou encore

$$\forall \lambda > 0, \text{ l'intégrale } \int_0^{+\infty} x^2 e^{-\lambda x} dx \text{ converge et vaut } \frac{2}{\lambda^3}$$

Remarque 11. Ces trois derniers résultats s'appliqueront assez souvent dans le cas $\lambda = 1$; il convient alors de donner la justification.

Exemple de rédaction :

« L'intégrale $\int_0^{+\infty} x e^{-x} dx$ donne l'espérance d'une variable $X \hookrightarrow \mathcal{E}(1)$; elle converge donc, et est égale à 1. »

- Soit $X \hookrightarrow \mathcal{N}\left(0, \frac{1}{2}\right)$. L'intégrale sur \mathbb{R} de la densité converge ; X admet une espérance (nulle) et une variance ; on en déduit que :

Les intégrales $\int_{-\infty}^{+\infty} e^{-x^2} dx$; $\int_{-\infty}^{+\infty} x e^{-x^2} dx$; $\int_{-\infty}^{+\infty} x^2 e^{-x^2} dx$ convergent

3 Densités associées à des transformations d'une var

Une problématique très fréquente est, connaissant la loi d'une variable X , de déterminer la loi d'une fonction de cette variable (exemple : $Y = 2X + 1$, $Y = X^2$, $Y = \ln(X)$, etc.)

La méthode à suivre est la suivante.

Méthode :

X est une variable aléatoire ; on définit une nouvelle variable $Y = f(X)$.

Pour déterminer la loi de Y , on calcule sa fonction de répartition.

- Pour $y \in \mathbb{R}$, on exprime l'événement $(Y \leq y)$ en fonction d'un événement sur X ;
- comme on connaît la fonction de répartition de X , on en déduit $F_Y(y) = \mathbb{P}(Y \leq y)$.
- (optionnel) si on demande une densité de Y , on vérifie que Y est bien à densité et on en trouve une densité par dérivation comme exposé dans la méthode page 5.

3.1 Exemple : transformation affine

Soit X une variable admettant une densité f_X ; on pose $Y = -2X + 1$.

Soit F_Y la fonction de répartition de Y : on a, pour $y \in \mathbb{R}$,

$$(Y \leq y) = (-2X + 1 \leq y) = \left(X \geq \frac{1-y}{2} \right)$$

On en déduit que $F_Y(y) = 1 - F_X\left(\frac{1-y}{2}\right)$ (comme X est à densité, $\mathbb{P}(X \geq x) = \mathbb{P}(X > x) = 1 - \mathbb{P}(X \leq x)$).

Si on veut repasser aux densités, il suffit de dériver la fonction de répartition aux points où c'est possible :

$$\forall y \in \mathbb{R}, F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}\left(X \geq \frac{1-y}{2}\right) = 1 - F_X\left(\frac{1-y}{2}\right)$$

donne en dérivant par rapport à y :

$$\forall y \in \mathbb{R}, F'_Y(y) = -\left(-\frac{1}{2}\right)F'_X\left(\frac{1-y}{2}\right)$$

et finalement

$$\forall y \in \mathbb{R}, f_Y(y) = \frac{1}{2}f_X\left(\frac{1-y}{2}\right)$$

donne une densité de Y en fonction d'une densité de X .

Ce résultat se généralise en :

Proposition 15. Soit X une variable aléatoire à densité f_X , et $Y = aX + b$, avec $a \neq 0$. Alors Y est à densité, et une densité de Y est donnée par :

$$\forall y \in \mathbb{R}, f_Y(y) = \frac{1}{|a|}f_X\left(\frac{y-b}{a}\right)$$

Cette formule n'est pas à connaître, elle se redémontre au cas par cas.

Transformée affine d'une loi uniforme

Proposition 16.

- Soit $X \hookrightarrow \mathcal{U}([0, 1])$, et $a < b$. Alors $a + (b - a)X \hookrightarrow \mathcal{U}([a, b])$.
- Notamment : si $X \hookrightarrow \mathcal{U}([a, b])$, alors $\frac{X-a}{b-a} \hookrightarrow \mathcal{U}([0, 1])$.

Démonstration. Par la formule précédente (avec $b - a > 0$), une densité de $a + (b - a)X$ est donnée par

$$f: x \mapsto \begin{cases} \frac{1}{b-a} & \text{si } 0 \leq \frac{x-a}{b-a} \leq 1 \\ 0 & \text{sinon} \end{cases} \quad \text{soit encore } f: x \mapsto \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{sinon} \end{cases}$$

ce qui est bien la densité de $\mathcal{U}([a, b])$. \square

Transformée affine d'une loi normale

Proposition 17. Soit $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$. Alors $aX + b \hookrightarrow \mathcal{N}(a\mu + b, a^2\sigma^2)$.

En particulier, on peut « réduire » une variable suivant une loi normale :

$$\text{Si } X \hookrightarrow \mathcal{N}(\mu, \sigma^2), \text{ alors } \frac{X - \mu}{\sigma} \hookrightarrow \mathcal{N}(0, 1)$$

Démonstration. C'est un cas particulier du théorème précédent; une densité de X est

$$f_X: x \mapsto \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

En notant $Y = aX + b$, on obtient :

$$\forall y \in \mathbb{R}, f_Y(y) = \frac{1}{|a|} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\left(\frac{y-b}{a} - \mu\right)^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi a^2\sigma^2}} \exp\left(-\frac{(y - (a\mu + b))^2}{2a^2\sigma^2}\right)$$

et on obtient bien que $Y \hookrightarrow \mathcal{N}(a\mu + b, a^2\sigma^2)$. \square

3.2 Un cas particulier important : simulation d'une loi exponentielle

On va démontrer le résultat suivant :

Proposition 18. Soit $X \hookrightarrow \mathcal{U}([0, 1])$, et $Y = -\frac{1}{\lambda} \ln(1 - X)$, avec $\lambda > 0$. Alors $Y \hookrightarrow \mathcal{E}(\lambda)$.

Démonstration. Soit $y \in \mathbb{R}$. On a : $Y \leq y \Leftrightarrow -\frac{1}{\lambda} \ln(1 - X) \leq y \Leftrightarrow X \leq 1 - e^{-\lambda y}$, et donc

$$\forall y \geq 0, F_Y(y) = F_X(1 - e^{-\lambda y})$$

Or

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}$$

Donc

$$F_Y(y) = \begin{cases} 0 & \text{si } 1 - e^{-\lambda y} < 0 \quad (1) \\ 1 - e^{-\lambda y} & \text{si } 0 \leq 1 - e^{-\lambda y} \leq 1 \quad (2) \\ 1 & \text{si } 1 - e^{-\lambda y} > 1 \quad (3) \end{cases}$$

Éclaircissons ces conditions :

- (1) : $1 - e^{-\lambda y} < 0 \Leftrightarrow e^{-\lambda y} > 1 \Leftrightarrow -\lambda y > 0 \Leftrightarrow y < 0$
- (2) : $0 \leq 1 - e^{-\lambda y} \leq 1 \Leftrightarrow 0 \leq e^{-\lambda y} \leq 1 \Leftrightarrow -\lambda y \leq 0 \Leftrightarrow y \geq 0$
- (3) : $1 - e^{-\lambda y} > 1 \Leftrightarrow e^{-\lambda y} < 0$: impossible.

Récapitulons : $F_Y(y) = \begin{cases} 0 & \text{si } y < 0 \\ 1 - e^{-\lambda y} & \text{si } y \geq 0 \end{cases}$. On reconnaît bien la fonction de répartition de $\mathcal{E}(\lambda)$; on peut donc conclure $Y \hookrightarrow \mathcal{E}(\lambda)$.

Remarque 12. Cette démo est très usuelle, tombe souvent aux concours, et doit alors être effectuée de manière très soigneuse, sans oublier d'étape... \square

4 Cas de plusieurs variables aléatoires sur le même univers

4.1 Indépendance

Comme dans le cas de variables discrètes, on examine des situations où on a défini plus d'une variable aléatoire sur un même espace probabilisé. Il se pose alors, notamment, des questions d'indépendance.

Définition 11. Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, et X et Y deux variables aléatoires réelles définies sur Ω . On dit que X et Y sont indépendantes si et seulement si :

$$\text{pour tous intervalles } I, J \text{ de } \mathbb{R}, \mathbb{P}([X \in I] \cap [Y \in J]) = \mathbb{P}(X \in I) \times \mathbb{P}(Y \in J)$$

Remarque 13. Ce critère **équivaut** à une équation similaire sur les fonctions de répartition.

$$X \text{ et } Y \text{ sont indépendantes si et seulement si : } \forall (t, t') \in \mathbb{R}^2, \mathbb{P}((X \leq t) \cap (Y \leq t')) = \mathbb{P}(X \leq t) \times \mathbb{P}(Y \leq t')$$

On peut immédiatement étendre cette définition au cas de n variables aléatoires :

Définition 12. Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, et X_1, \dots, X_n n variables aléatoires définies sur Ω . On dit que X_1, \dots, X_n sont mutuellement indépendantes si et seulement si :

$$\text{pour tous intervalles } I_1, \dots, I_n \text{ de } \mathbb{R}, \mathbb{P}\left(\bigcap_{i=1}^n [X_i \in I_i]\right) = \prod_{i=1}^n \mathbb{P}(X_i \in I_i)$$

Remarque 14. De même, X_1, \dots, X_n sont mutuellement indépendantes si et seulement si :

$$\forall (t_1, \dots, t_n) \in \mathbb{R}^n, \mathbb{P}(X_1 \leq t_1, X_2 \leq t_2, \dots, X_n \leq t_n) = \prod_{k=1}^n \mathbb{P}(X_k \leq t_k)$$

On voit que si (X_1, \dots, X_n) sont indépendantes, tout sous-ensemble de variables aléatoires prises dans cet ensemble vérifie la même propriété. Par contre, si pour tous $i \neq j$, X_i et X_j sont indépendantes, on ne peut pas conclure que (X_1, \dots, X_n) sont mutuellement indépendantes.

On dispose aussi du *lemme des coalitions* :

Proposition 19. Si (X_1, \dots, X_n) sont mutuellement indépendantes, et $f : \mathbb{R}^k \rightarrow \mathbb{R}$, $g : \mathbb{R}^{n-k} \rightarrow \mathbb{R}$ sont deux fonctions, alors les variables aléatoires $f(X_1, \dots, X_k)$ et $g(X_{k+1}, \dots, X_n)$ sont indépendantes.

Remarque 15. On retrouve comme corollaire usuel que si X et Y sont indépendantes, et f et g sont deux fonctions quelconques de \mathbb{R} dans \mathbb{R} , alors $f(X)$ et $g(Y)$ sont aussi indépendantes.

Toutes les formules suivantes sont similaires au cas des VAD. Leurs démonstrations nécessitent de disposer du formalisme des *intégrales doubles*, qui est très loin du programme : nous les admettrons.

Théorème 20 (Propriétés diverses). *On désigne par X, Y, X_1, \dots, X_n des variables aléatoires à densité sur Ω .*

- **Linéarité de l'espérance** : si X et Y admettent des espérances, et $\lambda \in \mathbb{R}$, alors $\lambda X + Y$ admet une espérance, et

$$\mathbb{E}(\lambda X + Y) = \lambda \mathbb{E}(X) + \mathbb{E}(Y)$$

On peut étendre au cas de n variables aléatoires admettant des espérances :

$$\forall (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n, \mathbb{E}\left(\sum_{i=1}^n \lambda_i X_i\right) = \sum_{i=1}^n \lambda_i \mathbb{E}(X_i)$$

- **Croissance de l'espérance** : Si $\mathbb{P}(X \leq Y) = 1$, alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$.

(Remarque : ce sera notamment le cas si on a toujours $X \leq Y$).

- **Propriété de l'espérance d'un produit de variables aléatoires indépendantes** :

Si X et Y sont indépendantes et admettent des espérances, alors XY admet une espérance, et $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

Grâce au lemme des coalitions, on peut étendre au cas de n variables aléatoires mutuellement indépendantes admettant des espérances :

$$\text{si } (X_1, \dots, X_n) \text{ sont mut. indép., } \mathbb{E}\left(\prod_{i=1}^n X_i\right) = \prod_{i=1}^n \mathbb{E}(X_i)$$

- **Propriété de la variance d'une somme de variables indépendantes** :

Si X et Y sont indépendantes et admettent des variances, alors $X + Y$ admet une variance, et $V(X + Y) = V(X) + V(Y)$.

Grâce aux lemme des coalitions, on peut étendre au cas de n variables aléatoires mutuellement indépendantes admettant des variances :

$$\text{si } (X_1, \dots, X_n) \text{ sont mut. indép., } V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i)$$

Si les λ_i sont des réels, le caractère quadratique de la variance donne, pour des variables aléatoires X_i mutuellement indépendantes :

$$V\left(\sum_{i=1}^n \lambda_i X_i\right) = \sum_{i=1}^n \lambda_i^2 V(X_i)$$

5 Propriétés additionnelles

5.1 Compléments sur la loi normale

Propriétés de la fonction de répartition

Proposition 21. Soit X une variable aléatoire suivant la loi normale **centrée réduite**. On rappelle que sa fonction de répartition Φ vérifie : $\forall x \in \mathbb{R}$, $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{t^2}{2}\right) dt$. On a :

- $\forall x \in \mathbb{R}$, $\Phi(-x) = 1 - \Phi(x)$.
- En particulier, $\Phi(0) = \frac{1}{2}$.

Démonstration. Par définition de la fonction de répartition :

$$1 - \Phi(x) = 1 - \mathbb{P}(X \leq x) = \mathbb{P}(X > x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} \exp\left(-\frac{t^2}{2}\right) dt$$

Le changement de variable $u = -t$ donne :

$$1 - \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} \exp\left(-\frac{t^2}{2}\right) dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-x} \exp\left(-\frac{u^2}{2}\right) du = \Phi(-x)$$

□

Remarque 16. Ce résultat subsiste si on suppose seulement la loi normale centrée ($\mu = 0$, $\sigma > 0$ quelconque).

Représenter graphiquement ce résultat (pour $x \geq 0$ par exemple) :

Stabilité de la loi normale Il existe un résultat de stabilité pour une somme de variables indépendantes suivant des lois normales. Ce théorème est admis.

Théorème 22. Soient $X \hookrightarrow \mathcal{N}(m_1, \sigma_1^2)$ et $Y \hookrightarrow \mathcal{N}(m_2, \sigma_2^2)$ deux variables aléatoires indépendantes. Alors $X + Y \hookrightarrow \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$.

On peut étendre à une combinaison linéaire de n variables : si, pour tout $i \in \llbracket 1, n \rrbracket$, $X_i \hookrightarrow \mathcal{N}(m_i, \sigma_i^2)$, $\lambda_i \in \mathbb{R}$, et les X_i sont mutuellement indépendantes, alors

$$\sum_{i=1}^n \lambda_i X_i \hookrightarrow \mathcal{N}\left(\sum_{i=1}^n \lambda_i m_i, \sum_{i=1}^n \lambda_i^2 \sigma_i^2\right)$$

5.2 Loi du min, loi du max de variables indépendantes

La méthode suivante intervient dans beaucoup d'exercices. Elle est similaire à celle vue dans le cas de variables discrètes.

Soient X_1, \dots, X_n des variables aléatoires mutuellement indépendantes, de fonctions de répartition respectives F_1, \dots, F_n .

Soit $S = \max(X_1, X_2, \dots, X_n)$ et $I = \min(X_1, X_2, \dots, X_n)$. On peut obtenir les fonctions de répartition de S et I de la manière suivante :

- Soit $t \in \mathbb{R}$. On a $(S \leq t) = \bigcap_{k=1}^n (X_k \leq t)$, soit par indépendance

$$F_S(t) = \mathbb{P}(S \leq t) = \mathbb{P}\left(\bigcap_{k=1}^n (X_k \leq t)\right) = \prod_{k=1}^n \mathbb{P}(X_k \leq t) = F_1(t) \times F_2(t) \times \dots \times F_n(t)$$

Dans le cas où toutes les variables X_i suivent la même loi, et possèdent donc la même fonction de répartition notée F , cette dernière formule donne :

$$\forall t \in \mathbb{R}, F_S(t) = (F(t))^n$$

- Soit $t \in \mathbb{R}$. On a $(I > t) = \bigcap_{k=1}^n (X_k > t)$, soit par indépendance

$$1 - F_I(t) = \mathbb{P}(I > t) = \mathbb{P}\left(\bigcap_{k=1}^n (X_k > t)\right) = \prod_{k=1}^n \mathbb{P}(X_k > t) = (1 - F_1(t)) \times (1 - F_2(t)) \times \dots \times (1 - F_n(t))$$

Ici aussi, dans le cas où toutes les X_i ont pour fonction de répartition F , on trouve :

$$\forall t \in \mathbb{R}, 1 - F_I(t) = (1 - F(t))^n$$

Annexe (hors-programme) : théorèmes de limite monotone

Théorème 23.

- Soit (A_n) une suite croissante d'événements : alors

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} P(A_n)$$

- Soit (B_n) une suite décroissante d'événements : alors

$$P\left(\bigcap_{n=0}^{+\infty} B_n\right) = \lim_{n \rightarrow +\infty} P(B_n)$$

Démonstration. On rappelle qu'on dit qu'une suite d'événements (A_n) est croissante (pour l'inclusion) ssi :

$$\forall n \in \mathbb{N}, A_n \subset A_{n+1}$$

Soit une telle suite (A_n) : pour $n \geq 1$, on pose $B_n = A_n \setminus A_{n-1}$ et $B_0 = A_0$. On vérifie alors que les B_i sont 2 à 2 disjoints, et que

$$\forall n \in \mathbb{N}, A_n = \bigsqcup_{k=0}^n B_k \quad \text{et} \quad \bigcup_{n=0}^{+\infty} A_n = \bigsqcup_{k=0}^{+\infty} B_k$$

Par propriété d'une probabilité, $P\left(\bigcup_{n=0}^{+\infty} A_n\right) = P\left(\bigsqcup_{k=0}^{+\infty} B_k\right) = \sum_{k=0}^{+\infty} P(B_k)$; et $P(A_N) = P\left(\bigsqcup_{k=0}^N B_k\right) = \sum_{k=0}^N P(B_k)$. En faisant tendre N vers $+\infty$,

$$\lim_{N \rightarrow +\infty} P(A_N) = \sum_{k=0}^{+\infty} P(B_k) = P\left(\bigcup_{n=0}^{+\infty} A_n\right)$$

□

En conséquence :

Proposition 24. Pour toute suite d'événements (A_n) , on a :

$$P\left(\bigcup_{k=0}^{+\infty} A_k\right) = \lim_{n \rightarrow +\infty} P\left(\bigcup_{k=0}^n A_k\right)$$

$$P\left(\bigcap_{k=0}^{+\infty} A_k\right) = \lim_{n \rightarrow +\infty} P\left(\bigcap_{k=0}^n A_k\right)$$

Démonstration. On applique les propositions précédentes :

- à la suite croissante d'événements $\left(\bigcup_{k=0}^n A_k\right)_{n \in \mathbb{N}}$.

- à la suite décroissante d'événements $\left(\bigcap_{k=0}^n A_k\right)_{n \in \mathbb{N}}$.

□

Démonstrations

Proposition 2

Soit X une variable aléatoire de fonction de répartition F_X . Alors :

- F_X est croissante sur \mathbb{R} .
 - F_X est continue à droite en tout point de \mathbb{R} (c'est-à-dire, en tout point x , $\mathbb{P}(X \leq x) = \lim_{t \rightarrow x^+} F_X(t)$).
 - $\lim_{x \rightarrow -\infty} F_X(x) = 0$.
 - $\lim_{x \rightarrow +\infty} F_X(x) = 1$.
 - En tout point x , $\mathbb{P}(X < x) = \lim_{t \rightarrow x^-} F_X(t)$.
 - En tout point x , $\mathbb{P}(X = x) = \lim_{t \rightarrow x^+} F_X(t) - \lim_{t \rightarrow x^-} F_X(t)$.
- En particulier, F_X est continue en x si et seulement si $\mathbb{P}(X = x) = 0$.

Démonstration. • Pour tous $(x, y) \in \mathbb{R}^2$ tels que $x \leq y$: $F(y) - F(x) = \mathbb{P}(X \leq y) - \mathbb{P}(X \leq x) = \mathbb{P}(x < X \leq y) \geq 0$. F est donc croissante.

- F_X est monotone, et minorée par 0, donc admet une limite à droite en tout point réel. On a

$$\lim_{t \rightarrow x^+} F_X(t) = \lim_{n \rightarrow +\infty} F_X(x + \frac{1}{n}) = \lim_{n \rightarrow +\infty} \mathbb{P}(X \leq x + \frac{1}{n}) = \mathbb{P}\left(\bigcap_{n=1}^{+\infty} (X \leq x + 1/n)\right) = \mathbb{P}(X \leq x) = F_X(x)$$

par décroissance de la suite d'événements $\left(X \leq x + \frac{1}{n}\right)$.

- Par monotonie et minoration $\lim_{x \rightarrow -\infty} F_X(x)$ existe et est comprise entre 0 et 1. On a alors

$$\lim_{x \rightarrow -\infty} F_X(x) = \lim_{n \rightarrow +\infty} F_X(-n) = \lim_{n \rightarrow +\infty} \mathbb{P}(X \leq -n) = \mathbb{P}\left(\bigcap_{n=0}^{+\infty} (X \leq -n)\right) = \mathbb{P}(\emptyset) = 0$$

par décroissance de la suite d'événements $(X \leq -n)$.

- Par monotonie et majoration $\lim_{x \rightarrow +\infty} F_X(x)$ existe et est comprise entre 0 et 1. On a alors

$$\lim_{x \rightarrow +\infty} F_X(x) = \lim_{n \rightarrow +\infty} F_X(n) = \lim_{n \rightarrow +\infty} \mathbb{P}(X \leq n) = \mathbb{P}\left(\bigcup_{n=0}^{+\infty} (X \leq n)\right) = \mathbb{P}(\Omega) = 1$$

par croissance de la suite d'événements $(X \leq n)$.

- Comme précédemment, $\lim_{t \rightarrow x^-} F_X(t)$ existe ; elle est donc égale à $\lim_{n \rightarrow +\infty} F_X(x - 1/n)$.

Or

$$\lim_{n \rightarrow +\infty} F_X(x - 1/n) = \lim_{n \rightarrow +\infty} \mathbb{P}(X \leq x - \frac{1}{n}) = \mathbb{P}\left(\bigcup_{n=1}^{+\infty} X \leq x - \frac{1}{n}\right) = \mathbb{P}(X < x)$$

par croissance de la suite d'événements $(X \leq x - 1/n)$.

-

$$\mathbb{P}(X = x) = \mathbb{P}(X \leq x) - \mathbb{P}(X < x) = F_X(x) - \lim_{t \rightarrow x^-} F_X(t) = \lim_{t \rightarrow x^+} F_X(t) - \lim_{t \rightarrow x^-} F_X(t)$$

□

Proposition 4

Soit X une variable aléatoire admettant une densité f . On a les propriétés suivantes :

- Pour tout $x \in \mathbb{R}$, $\mathbb{P}(X = x) = 0$.
 - Si I est un intervalle de \mathbb{R} , la quantité $\mathbb{P}(X \in I)$ ne change pas si on ouvre ou ferme les crochets aux bornes de I .
 - Pour tout $x \in \mathbb{R}$,
- $$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X < x) = \int_{-\infty}^x f(t) dt$$
- $\int_{-\infty}^{+\infty} f(t) dt = 1$.
 - Pour tout $x \in \mathbb{R}$,
- $$\mathbb{P}(X \geq x) = \mathbb{P}(X > x) = 1 - F_X(x) = \int_x^{+\infty} f(t) dt$$
- Pour tous $(a, b) \in \mathbb{R}^2$ tels que $a \leq b$:

$$\mathbb{P}(a \leq x \leq b) = \mathbb{P}(a < x \leq b) = \mathbb{P}(a \leq x < b) = \mathbb{P}(a < x < b) = \int_a^b f(x) dx$$

Démonstration. L'existence de points accidentels de non-continuité / non-dérivabilité rendant les discussions complexes, nous allons supposer que F_X est \mathcal{C}^1 sur \mathbb{R} , et que $f = F'_X$. f est donc continue sur \mathbb{R} .

- On a vu que si F_X est continue en $x \in \mathbb{R}$, alors $\mathbb{P}(X = x) = 0$. Ici, F_X étant \mathcal{C}^1 sur \mathbb{R} entier, on a :

$$\forall x \in \mathbb{R}, \mathbb{P}(X = x) = 0$$

- Cette dernière propriété montre que $\mathbb{P}(X \in I)$ ne change pas si les crochets de I sont ouverts ou fermés.

- Comme $F'_X = f$, on a l'existence de $a \in \mathbb{R}$ tel que : $\forall x \in \mathbb{R}, F_X(x) = a + \int_0^x f(t) dt$. De plus, d'après la proposition 2, $\lim_{x \rightarrow -\infty} F_X(x) = a + \int_0^{-\infty} f(t) dt = 0$ (l'existence de cette limite justifiant la convergence de l'intégrale impropre), ce qui donne $a = \int_{-\infty}^0 f(t) dt$ et finalement $F_X(x) = \int_{-\infty}^x f(t) dt$.
- La proposition 2 donne $\lim_{x \rightarrow +\infty} F_X(x) = 1 = \int_{-\infty}^{+\infty} f(t) dt$ d'après le point précédent. (et on a prouvé au passage la cv de cette intégrale).
- On utilise les résultats précédents pour écrire :

$$\begin{aligned} \mathbb{P}(X \geq x) &= \mathbb{P}(X > x) = \mathbb{P}(\overline{X \leq x}) \\ &= 1 - \mathbb{P}(X \leq x) \\ &= 1 - \int_{-\infty}^x f(t) dt \\ &= \int_{-\infty}^{+\infty} f(t) dt - \int_{-\infty}^x f(t) dt \\ &= \int_x^{+\infty} f(t) dt. \end{aligned}$$

- L'égalité de ces quatre probas est donnée par le second point ; ensuite l'égalité entre événements

$$(X \leq b) = (X \leq a) \sqcup (a < X \leq b)$$

donne $\mathbb{P}(X \leq b) = \mathbb{P}(X \leq a) + \mathbb{P}(a < X \leq b)$, puis

$$\mathbb{P}(a < X \leq b) = \mathbb{P}(X \leq b) - \mathbb{P}(X \leq a) = \int_{-\infty}^b f(t) dt - \int_{-\infty}^a f(t) dt = \int_a^b f(t) dt$$

□

Proposition 12

Soit X une variable aléatoire de densité f . Si f admet un moment d'ordre $m \in \mathbb{N}^*$, alors elle admet des moments d'ordre p pour tout $p \in \llbracket 1, m \rrbracket$.

Démonstration. Soit $p \in \llbracket 1, m \rrbracket$.

On sait que l'intégrale $\int_{-\infty}^{+\infty} |x^m f(x)| dx$ est convergente ; il s'agit de montrer que l'intégrale $\int_{-\infty}^{+\infty} |x^p f(x)| dx$ est convergente.

On a vu (cf cours sur les variables discrètes) : $\forall x \in \mathbb{R}$, $|x|^p \leq 1 + |x|^m$; d'où en multipliant par $f(x) \geq 0$:

$$\forall x \in \mathbb{R}, |x|^p f(x) \leq f(x) + |x|^m f(x)$$

ou encore

$$\forall x \in \mathbb{R}, 0 \leq |x^p f(x)| \leq f(x) + |x|^m f(x)$$

Par hypothèse les intégrales $\int_{-\infty}^{+\infty} f(x) dx$ (intégrale de la densité !) et $\int_{-\infty}^{+\infty} |x^m f(x)| dx$ convergent ; par majoration on en déduit que $\int_{-\infty}^{+\infty} |x^p f(x)| dx$ converge. □