CNAM MULHOUSE . Ramm Algebra Center

Technicien Développeur. Année 2.

Mathématiques

Cours Michel GOZE

Chapitre 5

Variables aléatoires continues. La loi normale. Approximation

Considérons un disque sur lequel un rayon a été tracé. Faisons tourner le disque (comme dans les tombolas des champs de foire) autour d'un axe horizontal et notons X la valeur de l'angle, une fois le disque arrêté, entre le rayon et un repère fixé à l'avance. Cette variable aléatoire sera un nombre compris entre 0 et 2π et ne peux être considérer comme une variable discrète. On va donc dans ce chapitre introduire la notion de variables aléatoires continues. L'application la plus importante sera un résultat d'approximation : nous avons vu qu'il était parfois difficile et même impossible de faire du calcul numériques pour des variables aléatoires discrètes, par exemple en raison des calculs difficiles des coefficients binomiaux. Ces calculs ou ces lois vont être approchées par des lois de variables continues, pour lesquelles les calculs seront "facilités" par l'usage de technique d'analyse mathématique, comme le calcul intégral. Nous allons donc dans ce chapitre faire un rappel sur le calcul intégral avant d'introduire la notion de variables aléatoires continues et présenter celle que nous utiliserons pour le calcul d'approximation à savoir la variable aléatoire "normale".

1. Rappel sur le calcul intégral

1.1. **Primitives d'une fonction continue.** Le point de départ est le théorème fondamental de l'analyse :

Théorème 1. *Soit $f : [a,b] \rightarrow \mathbb{R}$ une fonction continue sur l'intervalle $[a,b]$ à valeurs dans $\mathbb{R}$. La fonction F définie sur l'intervalle $[a,b]$ par*

$$F(x) = \int_a^x f(t)dt$$

est dérivable sur $[a,b]$ et vérifie $\left\{ \begin{array}{l} F(a) = 0 \\ \forall x \in [a,b], \ F'(x) = f(x) \end{array} \right.$

La fonction $F(x)$ est appelée la primitive de $f(x)$ qui s'annule en a. De manière générale, on appellera primitive de la fonction $f(x)$ toute fonction $F(x)$ telle que

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

Le théorème ci-dessus nous dit que si f est continue il existe une seule fonction primitive $F(x)$ avec la condition $F(a) = 0$. Le calcul de la fonction $F(x)$ à partir de sa dérivée $f(x)$ est souvent très difficile ou parfois impossible. On commence à établir un tableau des principaux résultats connus, comme les primitives des fonctions polynômes, des fonctions circulaires, de l'exponentielle. Les premières difficultés apparaissent lorsque l'on veut chercher une primitive de la fonction $\frac{1}{x}$. On introduit pour cela la fonction logarithme népérien qui est donc défini par

$$(lnx)' = \frac{1}{x}, \quad ln1 = 0.$$

En classe de terminale ou en première année d'études supérieures, on présente certaines techniques, dites d'intégration. Citons l'intégration par partie ou les changements de variables. Cela permet d'élargir le spectre des fonctions pour lesquelles on sait calculer une primitive. Par exemple les fonctions données par des fractions rationnelles ont des primitives que l'on sait calculer. Pour ceux qui ont "subi" cette étude, ils en ont certainement gardé quelques mauvais souvenirs, car si le calcul est "faisable" il en reste compliqué.

Soit $f(x)$ une fonction continue sur $\mathbb{R}$. On appelle intégrale définie entre deux points a et b le nombre noté

$$\int_a^b f(x)dx$$

qui est égal à

$$F(b) - F(a)$$

où $F(x)$ est une primitive de $f(x)$, c'est-à-dire donnée par $F'(x) = f(x)$.

Des techniques de calcul numériques permettent parfois de pouvoir calculer ce nombre (ou d'en donner une très bonne approximation) sans savoir calculer explicitement $F(x)$.

1.2. **Intégrales généralisées.** Rappelons la notion d'intégrale généralisée (parfois appelée aussi intégrale impropre).

1.2.1. *Intégrales généralisées de première espèce.* Soit f une fonction continue d'une variable réelle définie sur l'intervalle $[a, +\infty[$. L'intégrale $\int_a^{+\infty} f(t)dt$ n'est pas une intégrale classique, on dit que c'est une intégrale généralisée. On dit que cette intégrale généralisée *converge* si la fonction d'une variable réelle F définie sur $[a, +\infty[$ par

$$F(x) = \int_a^x f(t)dt$$

admet une limite quand x tend vers $+\infty$:

$$\int_a^{+\infty} f(t)dt = \lim_{x\to+\infty} \int_a^x f(t)dt.$$

Si F n'a pas de limite en $+\infty$ on dit que l'intégrale $\int_a^{+\infty} f(t)dt$ *diverge.*

On définit de même les intégrales généralisées $\int_{-\infty}^{a} f(t)dt$. Soit f une fonction continue d'une variable réelle définie sur l'intervalle $]-\infty, a]$. L'intégrale $\int_{-\infty}^{a} f(t)dt$ est généralisée. Elle converge si la fonction d'une variable réelle F définie sur $]-\infty, a]$ par

$$F(x) = \int_{-\infty}^{a} f(t)dt$$

admet une limite quand x tend vers $-\infty$:

$$\int_{-\infty}^{a} f(t)dt = \lim_{x \to -\infty} \int_{x}^{a} f(t)dt.$$

Si F n'a pas de limite en $-\infty$ on dit que l'intégrale $\int_{-\infty}^{a} f(t)dt$ diverge.

Remarque. Les intégrales généralisées $\int_{-\infty}^{+\infty} f(t)dt$. Soit f une fonction continue sur $\mathbb{R}$. On dit que l'intégrale généralisée $\int_{-\infty}^{+\infty} f(t)dt$ converge si pour tout $a \in \mathbb{R}$, les deux intégrales généralisées $\int_{-\infty}^{a} f(t)dt$ et $\int_{a}^{+\infty} f(t)dt$ convergent. Dans ce cas on écrit

$$\int_{-\infty}^{+\infty} f(t)dt = \int_{-\infty}^{a} f(t)dt + \int_{a}^{+\infty} f(t)dt$$

Critères de convergence des intégrales généralisées $\int_{a}^{+\infty} f(t)dt$. En L1 et L2, dans le cours d'analyse, des critères de convergence ont été présentés : critères de convergence pour les fonctions positives, absolue convergence, semi-convergence). On s'y rapportera dès que nécessaire. Mais dans ce cours, nous ne serons pas amener à de tels calculs. On admettra, si nécessaire, certains résultats.

1.2.2. *Intégrales généralisées de deuxième espèce.* On considère également des intégrales généralisées associées à des fonctions qui ne sont pas définies sur des intervalles fermés de $\mathbb{R}$. Soit f une fonction continue sur l'intervalle semi-ouvert $[a, b[$. Pour tout $x \in [a, b[$ on considère $F(x) = \int_{a}^{x} f(t)dt$. L'intégrale $\int_{a}^{b} f(t)dt$ est alors une intégrale généralisée qui converge si la fonction F admet une limite quand x tend vers b

$$\int_{a}^{b} f(t)dt = \lim_{x \to b} \int_{a}^{x} f(t)dt.$$

Si F n'a pas de limite en b on dit que l'intégrale généralisée $\int_{a}^{b} f(t)dt$ diverge.

2. Variables aléatoires continues

2.1. **Fonctions de densité.** On dira qu'une fonction est continue par morceaux si elle est continue sauf en un nombre fini de points, mais en ces points elle admet une limite à droite et une limite à gauche.

Définition 1. *On appelle densité de probabilité, une fonction* $f : \mathbb{R} \to [0, +\infty[$ *continue par morceaux telle que*

$$\int_{-\infty}^{+\infty} f(t)dt = 1.$$

Exemples.

(1) Montrons que la fonction définie sur l'intervalle $[0,4]$ par
$$f(x) = \frac{x^3}{4}$$
et nulle en dehors de cet intervalle est une fonction de densité. Cette fonction est bien continue par morceaux. On a
$$\int_{-\infty}^{+\infty} f(t)dt = \int_0^4 \frac{t^3}{4}dt = 1.$$
Or une primitive de $f(x)$ est
$$F(x) = \frac{1}{4}^3 \frac{x^4}{4}.$$
Ainsi
$$\int_0^4 \frac{t^3}{4}dt = F(4) - F(0) = \frac{1}{4}^3 \frac{4^4}{4} = 1.$$
La fonction $f(x)$ est bien une fonction de densité.

(2) On admet le résultat suivant
$$\int_{-\infty}^{+\infty} e^{-t^2}dt = \sqrt{\pi}.$$
Ceci permet de voir que
$$\frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$$
est une fonction de densité.

2.2. **Variables aléatoires continues.** Soit Ω un ensemble éventuellement infini. La notion d'espace probabilisé est à peu près analogue à la définition donnée dans le cas infini, à l'exception de l'ensemble des évènements. Dans le cas fini nous avons considéré $\mathcal{P}(\omega)$ comme ensemble des évènements, ce qui revenait à dire que tout sous-ensemble de Ω était un évènement. Ce n'est plus du tout le cas dans le cas infini, il faut considérer comme ensemble des évènements des familles de sous-ensembles contenant Ω, l'ensemble vide, les complémentaires et qui sont fermés pour la réunion et l'intersection finies. Par exemple, si $\Omega = \mathbb{R}$, l'ensembles des évènements est formé des intervalles et des réunions finies et des intersections d'intervalles. On ne va pas développer ce concept, il suffira de se souvenir que si $\Omega = \mathbb{R}$, les évènements sont des intervalles ou des réunions ou des intersections.

Soit X une application de Ω à valeurs dans $\mathbb{R}$. Soient a, b deux réels. On note
$$(a \leq X \leq b) = \{\omega \in \Omega,\ a \leq X(\omega) \leq b\}.$$

Définition 2. *Soit (Ω, P) un espace probabilisé infini. Une application $X : \Omega \to \mathbb{R}$ est une variable aléatoire continue de densité f si pour tout couple de réels a, b, on a*
$$P(a \leq X \leq b) = \int_a^b f(t)dt.$$

Ainsi, la fonction de répartition joue le rôle de loi de probabilité de loi de probabilité de la variable aléatoire X.

2.3. **Fonction de répartition. Espérance. Variance.** Soit X une variable aléatoire de densité f. La fonction de répartition de X est la fonction réelle

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt.$$

Ainsi, la fonction de densité est la dérivée de $F(x)$. On a

$$P(a \leq X \leq b) = F(b) - F(a).$$

On en déduit immédiatement que

$$P(X = x) = 0.$$

Cette dernière propriété est fausse dans le cas fini. Les évènements élémentaires ont dans le cas fini des probabilités non nulles.

Définition 3. *L'espérance et la variance de la variable aléatoire X de densité f sont données par*

$$E(X) = \int_{-\infty}^{+\infty} tf(t)dt$$

$$V(X) = \int_{-\infty}^{+\infty} (t - E(X))^2 f(t)dt = E(X^2) - (E(X))^2.$$

3. La loi normale

3.1. La loi normale centrée réduite.

Définition 4. *On dit qu'une variable aléatoire continue suit la loi normale centrée réduite si sa densité est*

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

La loi normale centrée rédute est notée $\mathcal{N}(0,1)$.

On admettra les résultats suivants, bien que leur démonstration ne soit pas inabordable, reposant essentiellement sur des changements de variable et des intégrations par partie

Proposition 1. *Si X est une variable aléatoire qui suit la loi centrée réduite alors son espérance et sa variance sont données par*

$$E(X) = 0, \quad V(X) = 1.$$

La courbe représentative de la fonction $\Phi(x)$ est la célèbre courbe "en cloche" : Pour une telle loi, on a donc

$$P(a \leq X \leq b) = \frac{1}{\sqrt{2\pi}} \int_{a}^{b} e^{-\frac{x^2}{2}} dx.$$

Remarques

(1) $P(X \leq 0) = P(X \geq 0) = 0,5$.
(2) $P(X \geq a) = 1 - P(X \leq a)$.
(3) Si $a \geq 0$ alors $P(X \geq a) = 0,5 - P(0 \leq X \leq a)$, $P(X \leq a) = 0,5 + P(0 \leq X \leq a)$.
(4) Si $a \leq 0$ alors $P(X \geq a) = 0,5 + P(a \leq X \leq 0)$, $P(X \leq a) = 0,5 - P(a \leq X \leq 0)$.

Théorème 2. *Si X est une variable aléatoire suivant la loi normale centrée réduite $\mathcal{N}(0,1)$, alors, pour tout $\alpha \in]0,1[$, il existe un unique réel positif u_α tel que $P(-u_\alpha \leq X \leq u_\alpha) = 1 - \alpha$.*

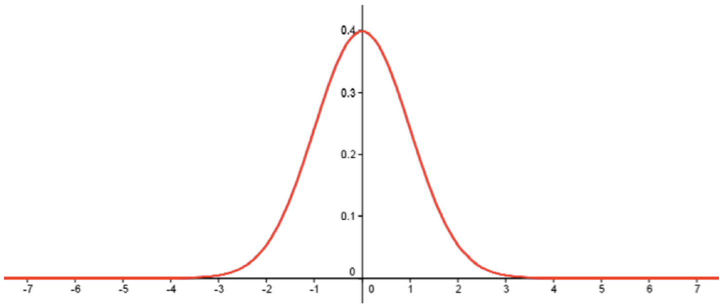

FIGURE 1. $\Phi(x)$

3.2. **La loi normale.**

Définition 5. *On dit qu'une variable aléatoire continue suit la loi normale* $\mathcal{N}(\mu, \sigma^2)$ *si sa densité est*

$$\Phi_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\dfrac{(x-\mu)^2}{2\sigma^2}}.$$

Ceci est équivalent à dire que $Y = \dfrac{X-\mu}{\sigma}$ suit la loi centrée réduite $\mathcal{N}(0,1)$.

3.3. **Application : Loi des grands nombres et théorème central limite.** On lance un dé. Le premier jeu consiste à deviner le résultat. On a une chance sur 6 de gagner. Par contre, considérons le leu un peu plus compliqué : on lance plusieurs fois le dé et il faut annoncer la somme des résultats. Par exemple si on lance 10 fois, on peut estimer que cette somme sera proche de 35 (le milieu du segment $(30, 40)$). Plus on lancera le dé, plus cette somme se rapprochera de $n \times 0,5$. Ceci peut se formaliser en disant :

Soit (X_n) une suite de variables aléatoires indépendantes ayant toutes la même moyenne μ et la même variance finie σ^2. Si on pose

$$Y_n = \frac{X_1 + \cdots + X_n}{n}$$

alors la suite (Y_n) "converge" (on ne précise pas ici cette notion, mais on peut l'imaginer) vers la variable aléatoire constante $Y = \mu$.

Le théorème central limite précise comment cette suite converge vers la moyenne.

Théorème 3. *Si on pose*

$$Z_n = \sqrt{n}(Y_n - \mu)$$

alors la suite Z_n converge vers une variable qui suit la loi normale centrée réduite $\mathcal{N}(0,1)$.

4. Approximation de la loi binomiale par la loi normale

Rappelons, dans un premier temps, la formule de Stirling, qui est la base de notre théorie d'approximation.

Formule de Stirling. Lorsque n devient grand, alors

$$n! \sim n^n e^{-n}\sqrt{2\pi n}.$$

De cette formule se déduit l'approximation suivante : si $npq \gg 1$, alors

$$\binom{k}{n} p^k q^{n-k} \simeq \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k-np)^2}{2npq}}.$$

La notation $npq \gg 1$ signifie que le nombre npq est grand par rapport à 1, sans trop préciser ce que signifie grand !

Proposition 2. *Soit X une variable aléatoire discrète suivant la loi de Bernoulli. Alors $P(X=k) = \binom{k}{n} p^k q^{n-k}$ et si $npq \gg 1$ alors*

$$P(X=k) \simeq \frac{1}{\sqrt{npq}} \varphi_{0,1}\left(\frac{k-np}{\sqrt{npq}}\right).$$

Exemple. On lance 1000 fois une pièce de monnaie. Quelle est la probabilité de faire 500 pile ? Cette probabilité est

$$P(X=500) = \binom{500}{1000} (0,5)^{500} (0,5)^{500}.$$

D'après le théorème d'approximation ci-dessus, comme $npq = 1000 \times 0,5 \times 0,5 = 250 \gg 1$,

$$P(X=500) \simeq \frac{1}{\sqrt{250}} \varphi_{0,1}\left(\frac{500-500}{\sqrt{250}}\right) = \frac{1}{\sqrt{250}} \frac{1}{\sqrt{2\pi}} \simeq 0,0252.$$

5. Fonction caractéristique d'une variable aléatoire continue

Dans le chapitre précédent, nous avons défini la notion de fonction caractéristique d'une variable aléatoire finie ou discrète.

Définition 6. *Soit X une variable aléatoire continue de fonction de densité $f_X(x)$. Sa fonction génératrice est la fonction*
$$\varphi_X(u) = \int_{-\infty}^{+\infty} e^{iux} f_X(x)dx.$$

Ce n'est rien d'autre que la transformée de Fourier de la fonction densité. Elle définit entièrement la loi de probabilité de la variable X. Ceci signifie que deux variables aléatoires ayant même fonction caractéristique ont des fonctions de répartition égales ou égales presque partout. Si m_k sont les moments d'ordre k de X, on a alors,
$$\varphi_X(u) = 1 + ium_1 + \cdots + \frac{(iu)^k}{k!} m_k + \cdots$$

Proposition 3. *Soit X une variable aléatoire suivant la loi normale $\mathcal{N}(\mu, \sigma)$. Sa fonction génératrice est la fonction*
$$\varphi_X(u) = e^{i\mu u - \frac{u^2\sigma^2}{2}}.$$

6. L'inégalité de Bienaymé-Tchebychev

6.1. **L'inégalité.** Etant donnée une variable aléatoire X, son écart-type mesure la tendance qu'à X à s'écarter de sa moyenne. L'inégalité qui suit mesure l'erreur que l'on peut commettre en remplaçant X par sa moyenne.

Théorème 4. *Soit X une variable aléatoire sur l'espace probabilité $(\Omega, \mathcal{F}, P)$ telle que $E(X^2)$ soit finie. Si on note par μ son espérance et par σ son écart-type, alors pour tout $\varepsilon \geq 0$,*
$$P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\varepsilon^2}.$$

Par exemple, la probabilité pour que le résultat d'une expérience $X(\omega)$ diffère de μ de plus de 8σ est
$$P(|X - \mu| \geq 8\sigma) \leq \frac{\sigma^2}{(8\sigma^2} = \frac{1}{64}.$$

6.2. **Calcul de π et l'aiguille de Buffon.** On lance une aiguille de longueur a sur un parquet dont les lattes sont de largeur b, $b \geq a$. Buffon calcula la probabilité pour que l'aiguille coupe la rainure entre deux lattes et trouva
$$p = \frac{2a}{\pi b}.$$
Lorsque $a = b$, on trouve
$$p = \frac{2}{\pi}$$

d'où
$$\pi = 2p.$$
On peut espérer calculer, par une bonne estimation de p, une valeur approchée de π. Soit k le nombre de lancers. La loi de probabilité de la variable aléatoire X correspondant à cette expérience est la loi de Bernoulli $\mathcal{B}(k,p)$. Son espérance est donc
$$E(X) = kp$$
et l'écart-type
$$\sigma(X) = \sqrt{kpq}.$$
L'inégalité de Bienaymé-Tchebychev s'écrit
$$P(|X-\mu| \geq \epsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$
ou bien
$$P(|X-\mu| \geq \epsilon\sigma) \leq \frac{1}{\varepsilon^2}$$
soit
$$P(|X-kp| \geq \epsilon\sigma) \leq \frac{1}{\varepsilon^2}.$$
Supposons que $\dfrac{1}{\varepsilon^2} = \dfrac{1}{100}$ soit $\varepsilon = 10$. Soit x le nombre d'aiguilles ayant coupée la rainure. Le calcul approché $\widetilde{\pi}$ de π est alors $\widetilde{\pi} = 2k/x$. Supposons que $\widetilde{\pi} = \pi \pm 10^{-3}$. Alors
$$\frac{2k}{\pi + 10^{-3}} \leq x \leq \frac{2k}{\pi - 10^{-3}}.$$
Si $|X - kp| = \varepsilon\sigma$, alors
$$k = \frac{\pi^2}{\varepsilon} 2.10^{-3}\sqrt{kpq}$$
et donc
$$k = \frac{\pi^{4}}{\varepsilon}^{2} 4.10^{-6} pq$$
ce qui donne à peu près $k = 560000000$.