Lois de probabilité à densité (I)

1 Généralités sur les lois à densité

1.1 Définition

Soit I un intervalle (quelconque) de $\mathbb{R}$ et f une fonction définie sur I.

Définition 1
f est une **densité de probabilité** sur I si f est continue, positive, et si $\displaystyle\int_I f(t)\mathrm{d}t = 1$.

L'expression $\displaystyle\int_I f(t)\mathrm{d}t$ signifie :

- $\displaystyle\int_a^b f(t)\mathrm{d}t$ si $I = [a\,;b]$,
- $\displaystyle\lim_{x\to+\infty}\int_a^x f(t)\mathrm{d}t$ si $I = [a\,;+\infty[$,
- $\displaystyle\lim_{x\to-\infty}\int_x^0 f(t)\mathrm{d}t + \lim_{y\to+\infty}\int_0^y f(t)\mathrm{d}t$ si $I =]-\infty\,;+\infty[$, etc.

On peut traduire cette condition par : l'aire sous la courbe représentative de f vaut 1.

Définition 2
Soit f une densité de probabilité sur I et soit X une variable aléatoire (continue).
On dit que X suit la loi de densité f, ou que la loi de X admet pour densité f si pour tous a, $b \in I$, $\displaystyle P(a \leq X \leq b) = \int_a^b f(t)\mathrm{d}t$.

Important : D'après la définition, $P(X = a) = 0$; $P(X \leq b) = P(X < b)$, etc. Que ce soit $<$ ou $\leq$ n'a pas d'importance. Attention, ce n'est **pas vrai pour les lois discrètes** (comme la loi binomiale).

1.2 Espérance - Variance - Écart-type

Par analogie avec le cas discret, on a la définition suivante (avec la convention d'écriture de la définition 1) :

Définition 3
Soit X une variable aléatoire dont la loi de probabilité admet la densité f sur I.
Alors l'espérance de X est : $\displaystyle E(X) = \mu = \int_I tf(t)\mathrm{d}t$.
La variance de X est : $\displaystyle Var(X) = \int_I (t-\mu)^2 f(t)\mathrm{d}t$.
L'écart-type de X est la racine carrée de la variance : $\sigma(X) = \sqrt{Var(X)}$.

1.3 Exemple de la loi uniforme

Propriété 1
La densité de la loi uniforme sur $[a\,;b]$ est $\dfrac{1}{b-a}$.

Si c et d appartiennent à $[a\,;b]$, on a : $\displaystyle P(c \leq X \leq d) = \int_c^d \frac{1}{b-a}\mathrm{d}t = \left[\frac{t}{b-a}\right]_c^d = \frac{d-c}{b-a}$.

Propriété 2
L'espérance d'une variable aléatoire suivant la loi uniforme sur $[a\,;b]$ est : $\dfrac{a+b}{2}$.
La variance est : $\dfrac{(b-a)^2}{12}$. On en déduit l'écart-type : $\sigma(X) = \dfrac{(b-a)}{2\sqrt{3}}$.

2 La loi normale d'espérance μ et d'écart-type σ

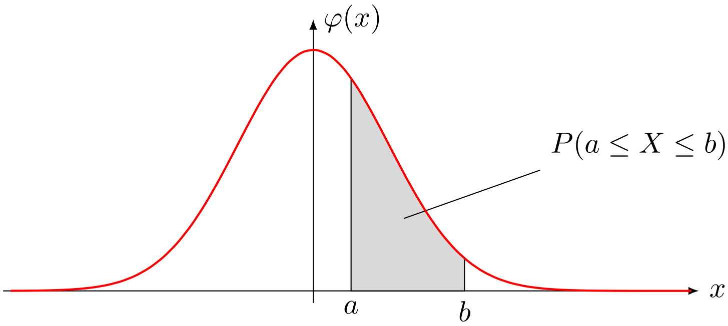
2.1 Définition

Soient μ, σ deux réels avec $\sigma > 0$.

Définition 4
La densité φ de la loi normale de paramètres μ et σ est définie sur $\mathbb{R}$ par : $\varphi(t) = \dfrac{1}{\sigma\sqrt{2\pi}}\,\mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

La v.a. X suit la loi normale de paramètres μ et σ si X est définie sur $\mathbb{R}$ et suit la loi de densité φ.

D'après la définition, on a alors : $P(a \leq X \leq b) = \displaystyle\int_a^b \frac{1}{\sigma\sqrt{2\pi}}\,\mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\,\mathrm{d}t$



2.2 Espérance, variance et écart-type

Propriété 3
Si X suit la loi normale de paramètre μ et σ, alors $E(X) = \mu$, $Var(X) = \sigma^2$, et l'écart-type est σ.

On peut donc parler de loi normale d'espérance μ et d'écart-type σ.

Vocabulaire : la loi normale d'espérance 0 et d'écart-type 1 s'appelle loi normale centrée-réduite.

2.3 Calcul numérique

1. Pour calculer $P(a \leq X \leq b)$:

Sur T.I :	Sur Casio :
• Taper `distrib`. • Choisir `normalFrép` ou `normalcdf`. • Taper les arguments dans l'ordre : a, b, μ, σ. • Fermer la parenthèse et valider.	• Taper `OPTN, F5, F3, F1, F2` pour obtenir `Ncd` (ou utiliser le catalogue). • Taper les arguments dans l'ordre : a, b, σ, μ. • Fermer la parenthèse et valider.

2. Pour déterminer t tel que $P(X \leq t) = p$:

Sur T.I :	Sur Casio :
• Taper `distrib`. • Choisir `FracNormale` ou `invNorm`. • Taper les arguments dans l'ordre : t, μ, σ. • Fermer la parenthèse et valider.	• Taper `OPTN, F5, F3, F1, F3` pour obtenir `InvN` (ou utiliser le catalogue). • Taper les arguments dans l'ordre : t, σ, μ. • Fermer la parenthèse et valider.

2.4 Valeurs approchées de certaines probabilités

Pour tout $k \in \mathbb{R}^+$, $P(X \in [\mu - k\sigma; \mu + k\sigma])$ est un nombre indépendant de μ et σ.

Par exemple :

$P(X \in [\mu-\sigma;\mu+\sigma]) \simeq 0{,}683.$ $\quad P(X \in [\mu-2\sigma;\mu+2\sigma]) \simeq 0{,}954$ $\quad P(X \in [\mu-3\sigma;\mu+3\sigma]) \simeq 0{,}997.$

2.5 Approximation d'une loi binomiale par une loi normale

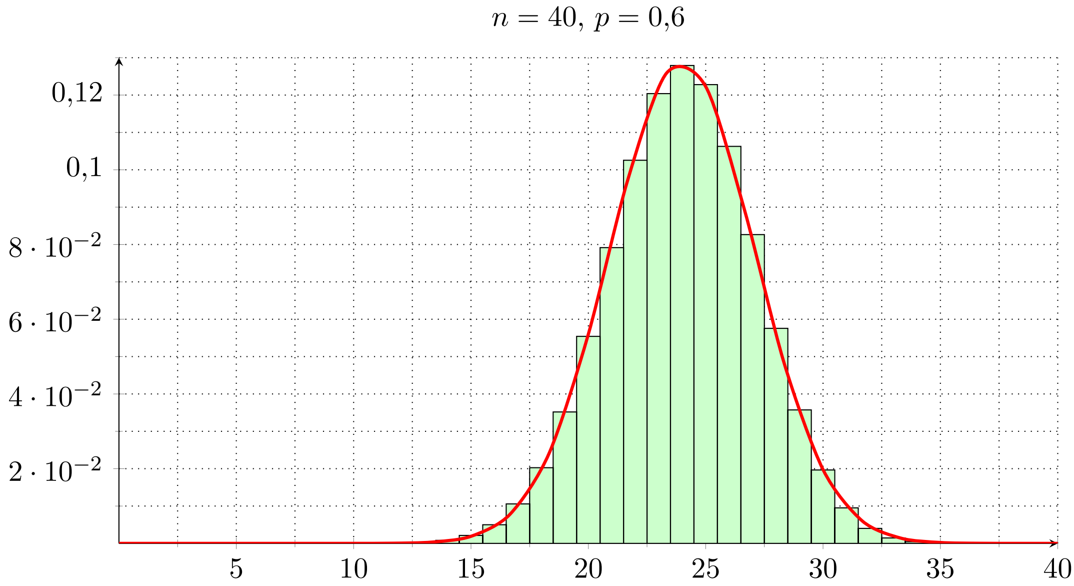
Lorsque le paramètre n est grand, et que p est ni trop proche de 0, ni trop proche de 1, on peut approcher la loi binomiale de paramètres n et p par la loi normale qui a même espérance et même écart-type.

Rappel :

Si X est une v.a. qui suit une loi binomiale de paramètres n et p, alors $E(X) = np$ et $\sigma(X) = \sqrt{np(1-p)}$.

Exemple : Si X suit la loi binomiale de paramètres $n = 40$ et $p = 0{,}6$, alors $E(X) = 24$ et $\sigma(X) \simeq 3{,}098$. Considérons alors la v.a. Y qui suit la loi normale d'espérance $\mu = 24$ et d'écart-type$\sigma = 3{,}098$.

Dans le graphique ci-dessous, les valeurs k sont en abscisse, le diagramme en bâton représente les $P(X = k)$ et la courbe en rouge la densité de la loi de Y.



On voit que $P(X = k)$ est voisin de $\varphi(k)$ où φ est la densité de la loi de Y.

On considère que l'approximation est suffisamment bonne lorsque $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$, mais cette condition n'est pas à savoir.

Calcul pratique :

Chaque bâton est centré en k et a pour largeur 1. Sa base est donc le segment $[k - 0{,}5\ ;k + 0{,}5]$. L'aire $\mathcal{A}_R$ de ce rectangle, est voisine de l'aire $\mathcal{A}_C$ sous la courbe de φ entre $k - 0{,}5$ et $k + 0{,}5$.

$\mathcal{A}_R = P(X = k) \times 1 = P(X = k)$; $\mathcal{A}_C = \displaystyle\int_{k-0{,}5}^{k+0{,}5} \varphi(t)\mathrm{d}t$. Par conséquent : $P(X = k) \simeq \displaystyle\int_{k-0{,}5}^{k+0{,}5} \varphi(t)\mathrm{d}t$.

Si on se donne deux valeurs entières r et s (avec $r < s$), la probabilité $P(r \leq X \leq s)$ correspond à la somme des $P(X = k)$, k prenant les valeurs de r à s.

On obtient : $P(r \leq X \leq s) \simeq \displaystyle\int_{r-0{,}5}^{r+0{,}5} \varphi(t)\mathrm{d}t + \int_{r+1-0{,}5}^{r+1+0{,}5} \varphi(t)\mathrm{d}t + ... + \int_{s-0{,}5}^{s+0{,}5} \varphi(t)\mathrm{d}t = \int_{r-0{,}5}^{s+0{,}5} \varphi(t)\mathrm{d}t$

Autrement dit :

Propriété 4
Si X suit la loi binomiale de paramètres $(n\,;p)$, $P(r \leq X \leq s) \simeq P(r - 0{,}5 \leq Y \leq s + 0{,}5)$, où Y suit la loi normale de même espérance et de même écart-type que X.

Exemple :

Soit X une v.a. suivant la loi binomiale de paramètres $n = 50$ et $p = 0{,}3$.

On cherche une valeur approchée de $P(8 \leq X \leq 12)$.

Soit Y la v.a. suivant $\mathcal{N}(\mu\,;\sigma^2)$ avec $\mu = 50 \times 0{,}3 = 15$ et $\sigma = \sqrt{50 \times 0{,}3 \times 0{,}7} \simeq 3{,}2404$.

On trouve : $P(7{,}5 \leq Y \leq 12{,}5) \simeq 0{,}210$. On peut vérifier à la calculatrice que $P(8 \leq X \leq 12) \simeq 0{,}215$

3 Opérations sur les variables aléatoires

1) Cas discret :

Soit X une variable aléatoire discrète d'espérance mathématique $E(X)$, de variance $V(X)$ et soit a un réel.

On définit la variable aléatoire $aX + b$ en appliquant aux valeurs prises la fonction $x \to ax + b$: si X prend la valeur x_i avec la probabilité p_i, alors $aX + b$ prend la valeur $ax_i + b$ avec la même probabilité p_i.

Propriété 5
$E(aX + b) = aE(X) + b$ et $V(aX + b) = a^2V(X)$.

On en déduit que : $\sigma(aX + b) = |a|\sigma(X)$.

Soient X et Y deux variables aléatoires prenant respectivement les valeurs x_1, x_2, ..., x_n et y_1, y_2, ..., y_p. On suppose que X et Y sont **indépendantes**, ce qui signifie : $P(X = x_i \text{ et } Y = y_j) = P(X = x_i) \times P(Y = y_j)$.

On considère l'ensemble $\mathscr{S}$ de toutes les sommes $x_i + y_j$; on définit une variable aléatoire Z en associant à chaque valeur z_k de $\mathscr{S}$ la somme des probabilités $P(X = x_i \text{ et } Y = y_j)$ pour lesquelles $x_i + y_j = z_k$.

De même, on définit une variable aléatoire W sur l'ensemble $\mathscr{D}$ des différences $x_i - y_j$ en associant à chaque valeur w_k de $\mathscr{D}$ la somme des probabilités $P(X = x_i \text{ et } Y = y_j)$ pour lesquelles $x_i - y_j = w_k$.

Propriété 6
$E(X + Y) = E(X) + E(Y)$ $\quad$ $E(X - Y) = E(X) - E(Y)$
$V(X + Y) = V(X) + V(Y)$ $\quad$ $V(X - Y) = V(X) + V(Y)$

2) Cas continu :

Soient X (resp. Y) une variable aléatoire continue d'espérance mathématique $E(X)$ (resp. $E(Y)$), de variance $V(X)$ (resp. $V(Y)$) et soit a un réel. On suppose X et Y indépendantes (les valeurs prises par l'une n'ont pas d'influence sur les valeurs prises par l'autre.) Sous réserve de définir les variables aléatoires $aX + b$, $X + Y$ et $X - Y$, on montre que les deux propriétés précédentes restent vraies.

Attention ! Les variances s'ajoutent, mais pas les écarts-type.

3) Cas particulier de la loi normale :

Si X suit la loi normale d'espérance μ_1 et d'écart-type σ_1 et si Y suit la loi normale d'espérance μ_2 et d'écart-type σ_2, alors pour a, $b \in \mathbb{R}$:

La v.a. $aX + b$ suit la loi normale d'espérance $a\mu_1 + b$ et d'écart-type $|a|\sigma_1$.

La v.a.aléatoire $Z = X + Y$ suit la loi normale d'espérance $\mu_1 + \mu_2$ et d'écart-type $\sqrt{\sigma_1^2 + \sigma_2^2}$.

La variable aléatoire $W = X - Y$ suit la loi normale d'espérance $\mu_1 - \mu_2$ et d'écart-type $\sqrt{\sigma_1^2 + \sigma_2^2}$.

4 Théorème de la limite centrée (ou théorème central-limite)

Théorème 1
Soient X_1, X_2, ..., X_n : n variables aléatoires deux à deux **indépendantes** et de **même loi**.
La variable aléatoire somme : $X_1 + X_2 + ... + X_n$ suit approximativement une loi normale.

En appliquant plusieurs fois la propriété 6, on voit que si la loi de chacune des variables X_1, X_2, ..., X_n a pour espérance m et pour écart-type σ, alors la variable aléatoire $X_1 + X_2 + ... + X_n$ suit approximativement la loi normale d'espérance $n \times m$ et d'écart-type $\sigma \times \sqrt{n}$.