Lois de probabilité à densité

I- Densité sur un intervalle

1- Définition

Soit f une fonction définie, continue et positive sur un intervalle $[a\,;\,b]$.
On dit que f est une fonction à densité sur $[a\,;\,b]$ lorsque son intégrale (sur $[a\,;\,b]$) est égale à 1.

Exemple : Soit f la fonction définie sur $[2\,;\,3]$ par : $f(x) = \dfrac{1}{8}x^2 + \dfrac{1}{12}x$.
Montrer que f est une fonction à densité.

2- Définition

Soit X une variable aléatoire continue sur $[a\,;\,b]$ et munie d'une fonction de densité f.
On dit que P est la loi de probabilité de densité f lorsque pour tout intervalle $[c\,;\,d]$ inclus dans $[a\,;\,b]$, on a : $P(X \in [c\,;\,d]) = \displaystyle\int_c^d f(x)\,\mathrm{d}x$.

3- Propriétés

- Pour tout réel c de l'intervalle $[a\,;\,b]$, $P(X = c) = P(\{c\}) = 0$;
- Pour tout réel c de l'intervalle $[a\,;\,b]$, $P(X \in [a\,;\,c[) + P(X \in [c\,;\,b]) = 1$.

II- Lois à densités

1- Loi uniforme

a- Définition

Une variable aléatoire X suit une loi uniforme sur $[a\,;\,b]$ lorsque sa densité est constante sur ce même intervalle.

b- Propriété

- La fonction de densité f d'une loi uniforme sur $[a\,;\,b]$ est définie par : $f(x) = \dfrac{1}{b-a}$.
- Pour tout intervalle $[c\,;\,d]$ inclus dans $[a\,;\,b]$, on a : $P(X \in [c\,;\,d]) = \dfrac{d-c}{b-a}$.
- Si X suit une loi uniforme, alors $E(X) = \displaystyle\int_a^b xf(x)\,\mathrm{d}x = \dfrac{b+a}{2}$.

2- Loi exponentielle

a- Définition

Soit λ un réel strictement positif.
La fonction f définie sur $[0\,;\,+\infty[$ par $f(t) = \lambda e^{-\lambda t}$ est la densité d'une loi de probabilité, appelée loi exponentielles de paramètre λ.

b- Propriétés

Soit X une variable aléatoire suivant une loi exponentielle de paramètre λ.

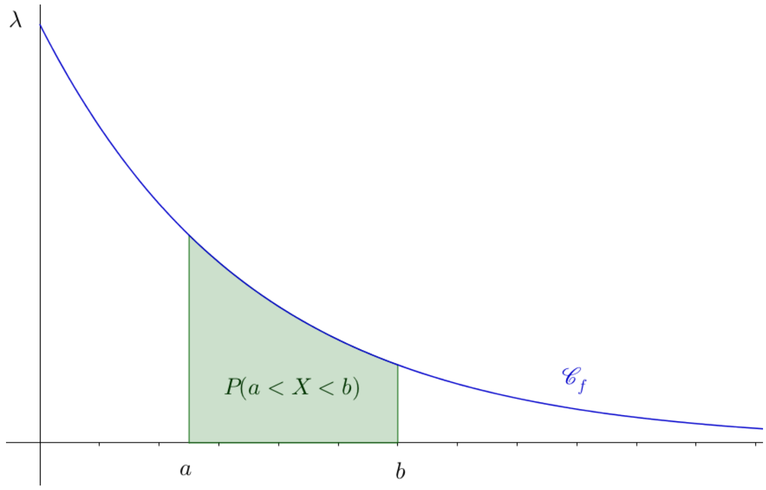
1. Pour tout réel t positif : $P(X \leqslant t) = \displaystyle\int_0^t \lambda e^{-\lambda t}\,\mathrm{d}x = 1 - e^{-\lambda t}$ et $P(X \geqslant t) = e^{-\lambda t}$.
2. Pour tous réels positif t et h, $P_{(X \geqslant t)}(X \geqslant t+h) = P(X \geqslant h)$.

3. X admet une espérance $E(X)$ définie par : $E(x) = \lim\limits_{x \to +\infty} \displaystyle\int_0^x t\lambda e^{-\lambda t}\,\mathrm{d}t = \frac{1}{\lambda}$.

Remarque :
Cette loi porte aussi le nom de loi de durée de vie sans vieillissement ou loi sans mémoire.
On peut interpréter l'espérance comme étant la durée de vie moyenne.

Démonstration : Propriété 2.

$$P_{(X\geqslant t)}(X \geqslant t+h) = \frac{P(X \geqslant t \text{ et } X \geqslant t+h)}{P(X \geqslant t)} = \frac{P(X \geqslant t+h)}{P(X \geqslant t)} = \frac{e^{-\lambda(t+h)}}{e^{-\lambda t}} = e^{-\lambda h} = P(X \geqslant h)$$

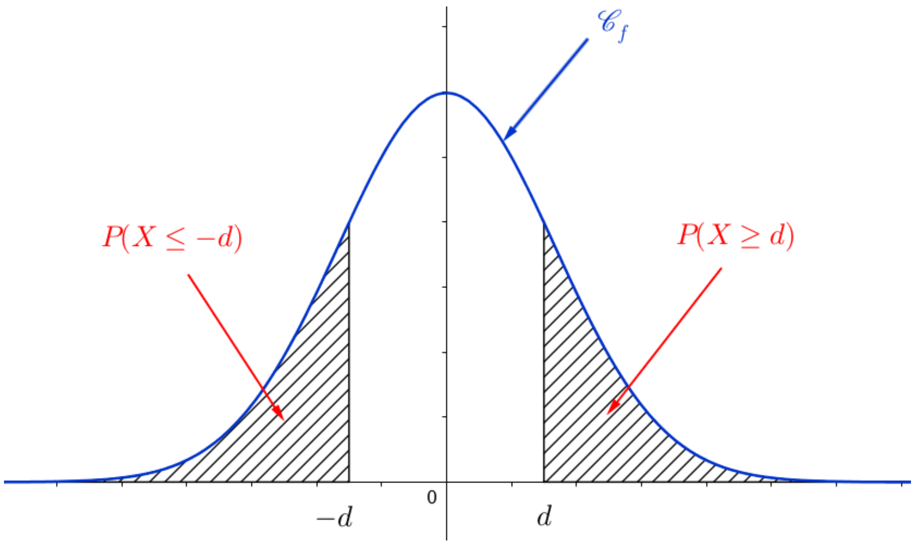


Courbe de la fonction densité d'une variable aléatoire suivant une loi exponentielle.

3- Loi normale $\mathcal{N}(0\,;\,1)$

a- Définition

Soit X une variable aléatoire, d'espérance $\mu = 0$ et d'écart-type $\sigma = 1$.
X suit un loi normale $\mathcal{N}(0\,;\,1)$ lorsqu'elle admet pour fonction de densité la fonction f définie sur $\mathbb{R}$ par : $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.



La courbe de la fonction densité d'une variable qui suit une loi $\mathcal{N}(0\,;\,1)$ est symétrique par rapport à l'axe des ordonnées.

b- Propriétés

- $P(x=c)=0$, $\forall c \in [a\,;\,b]$;
- $P(-d \leqslant X \leqslant d) = P(X \leqslant d) - P(X \leqslant -d)$;
- $P(X \leqslant -d) = P(X \geqslant d)$. On a alors :
  $P(-d \leqslant X \leqslant d) = 1 - 2P(X \leqslant -d)$ ou
  $P(-d \leqslant X \leqslant d) = 1 - 2P(X \geqslant d)$
- $P(-1,96 \leqslant X \leqslant 1,96) = 0,95$.

4- Loi normale $\mathcal{N}(\mu\,;\,\sigma^2)$

a- Définition

Si X suit une loi normale $\mathcal{N}(\mu\,;\,\sigma^2)$, alors la variable aléatoire définie par $Z = \dfrac{X-\mu}{\sigma}$ suit une loi normale $\mathcal{N}(0\,;\,1)$.

b- Propriétés

Soit X une variable aléatoire qui suit une loi normale $\mathcal{N}(\mu\,;\,\sigma^2)$.
On a (avec la calculatrice) alors :

- $P(X \in [\mu-\sigma\,;\,\mu+\sigma]) = 0,68$;
- $P(X \in [\mu-2\sigma\,;\,\mu+2\sigma]) = 0,95$;
- $P(X \in [\mu-3\sigma\,;\,\mu+3\sigma]) = 0,997$.

5- Passage de la loi binomiale discrète à la loi normale continue

a- Définition : Loi binomiale

Soit $\mathscr{E}$ une épreuve comportant deux issus (succès et échec) et p la probabilité de succès.
On repère n fois, de façon identiques et indépendantes, l'épreuve $\mathscr{E}$.
Soit X, la variable aléatoire associée au nombre total de succès.
Alors X suit une loi binomiale de paramètre n et p. On note $\mathscr{B}(n\,;\,p)$.

b- Propriétés

Soit X une variable aléatoire qui suit une loi binomiale $\mathscr{B}(n\,;\,p)$.
L'espérance, la variance et l'écart-type de X sont respectivement donnés par :
$E(X) = np$, $V(X) = np(1-p)$ et $\sigma_X = \sqrt{np(1-p)}$.

c- Théorème de Moivre-Laplace

Soit n un entier naturel non nul et p un nombre réel dans $[0;1]$. Soit X une variable aléatoire suivant une loi binomiale $\mathscr{B}(n\,;\,p)$. Alors, pour tous réels a et b, on a :

$$\lim_{n\to+\infty} P\left(a \leqslant \frac{X-np}{\sqrt{np(1-p)}} \leqslant b\right) = P(a \leqslant Z \leqslant b) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,\mathrm{d}x$$

où Z suit la loi normale centrée réduite $\mathcal{N}(0\,;\,1)$.

Remarque : En pratique, on considère que la limite dans le théorème de Moivre-Laplace est pratiquement atteinte lorsqu'on a simultanément $n \geqslant 30$, $np \geqslant 5$ et $n(1-p) \geqslant 5$.

On a ainsi pour tous réels a et b : $P\left(a \leqslant \dfrac{X-np}{\sqrt{np(1-p)}} \leqslant b\right) \approx P(a \leqslant Z \leqslant b)$,
où Z suit la loi normale centrée réduite $\mathcal{N}(0\,;\,1)$.

d- Corollaire : Intervalle de fluctuation

Si X_n suit une loi binomiale $\mathscr{B}(n\,;\,p)$, alors la fréquence de succès $F_n = \dfrac{X_n}{n}$ vérifie :

$$\lim_{n\to+\infty} P\left(p - u_\alpha \frac{\sqrt{p(1-p)}}{\sqrt{n}} \leqslant F_n \leqslant p + u_\alpha \frac{\sqrt{p(1-p)}}{\sqrt{n}}\right) = 1-\alpha$$

où u_α vérifie $P(-u_\alpha \leqslant Z \leqslant u_\alpha) = 1-\alpha$ pour Z suivant la loi $\mathcal{N}(0\,;\,1)$.

e- Intervalle de fluctuation

Pour $\alpha = 0,05$, $u_{0,05} = 1,96$. On en déduit que lorsque $n \geqslant 30$, $np \geqslant 5$ et $n(1-p) \geqslant 5$, la fréquence de succès F_n fluctue avec une probabilité de $0,95$ dans l'intervalle :

$$\left[p - 1,96\frac{\sqrt{p(1-p)}}{\sqrt{n}}\,;\,p + 1,96\frac{\sqrt{p(1-p)}}{\sqrt{n}}\right]$$

C'est l'intervalle de fluctuation asymptotique à 95 % de F_n.

III- Estimation du paramètre p d'une loi $\mathscr{B}(n\,;\,p)$

On souhaite déterminer la probabilité de succès p d'une épreuve de Bernoulli à partir de la réalisation de n expériences identiques et indépendantes.
Soit X_n le nombre de succès parmi les n expériences. Un estimateur naturel de p est donné par : $F_n = \dfrac{X_n}{n}$.
Quelle est alors la fiabilité de l'estimateur F_n.
La réponse donnée généralement est de proposer un *intervalle de confiance* de la probabilité p.

1- Définition

Soit $\alpha \in]0\,;\,1]$. Un intervalle de confiance au niveau $1-\alpha$ pour l'estimateur de p est un intervalle, noté IC, qui ne s'exprime qu'en fonction de F_n et n, tel que $P(p \in IC) \geqslant 1-\alpha$.

2- Proposition

Supposons que les les conditions suivantes sont vérifiées : $n \geqslant 30$, $np \geqslant 5$ et $n(1-p) \geqslant 5$.
L'intervalle de confiance au niveau asymptotique 95 % pour l'estimation de p est donné par :

$$IC = \left[F_n - \frac{1}{\sqrt{n}}\,;\,F_n + \frac{1}{\sqrt{n}}\right]$$