Module : Techniques d'estimation pour l'ingénieur **Classe :** $3^{\text{ème}}$

Variables aléatoires continues & lois usuelles continues

1 Variable aléatoire continue

1.1 Notion de variable aléatoire continue

Définition 1.

Une **variable aléatoire continue** (v.a.c) est une fonction X, allant d'un univers Ω à valeurs dans un intervalle de $\mathbb{R}$. i.e.

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

Exemple 1 *Exemples de variables aléatoires qui ne sont pas discrètes :*

1. *Variable T correspondant à la taille d'un étudiant.*
2. *Variable L correspondant à la longueur d'un train.*
3. *Variable A correspondant au temps d'attente à une caisse .*

Définition 2.

Soit X une v.a. continue. On appelle densité de probabilité de X, une application positive et intégrable $f : \mathbb{R} \to \mathbb{R}_+$, vérifiant :

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

Exemple 2 *Voici quelques exemples de densités de probabilités. Vérifier que $f_1(x)$ et $f_2(x)$ sont des densités de probabilités et tracer leurs courbes représentatives :*

$$f_1(x) = \begin{cases} \textit{0 si } x < 0 \\ \textit{1 si } 0 \leq x \leq 1 \\ \textit{0 si } x > 1 \end{cases} \qquad f_2(x) = \begin{cases} \textit{0 si } x < -1 \\ \textit{x+1 si } -1 \leq x \leq 0 \\ \textit{-x+1 si } 0 \leq x \leq 1 \\ \textit{0 si } x > 1 \end{cases}$$

1. — $f_1(x) \geq 0$
   — La fonction f_1 est continue par morceaux sur $\mathbb{R}$
   — $\displaystyle\int_{\mathbb{R}} f_1(x)dx = \int_{-\infty}^{0} f_1(x)dx + \int_{0}^{1} f_1(x)dx + \int_{1}^{+\infty} f_1(x)dx = 0 + [x]_0^1 + 0 = 1$

2. — $f_2(x) \geq 0$
   — La fonction f_1 est continue par morceaux sur $\mathbb{R}$

$$\begin{aligned} \int_{\mathbb{R}} f_2(x)dx &= \int_{-\infty}^{-1} f_2(x)dx + \int_{-1}^{0} f_2(x)dx + \int_{0}^{1} f_2(x)dx + \int_{1}^{+\infty} f_2(x)dx \\ &= 0 + [\frac{x^2}{2} + x]_{-1}^{0} + [-\frac{x^2}{2} + x]_0^1 + 0 = 1 \end{aligned}$$

1.2 Loi de probabilité d'une v.a.c.

On dit que X est une variable aléatoire continue de densité f si pour tout intervalle A de $\mathbb{R}$ on a :

$$\mathbb{P}[X \in A] = \int_A f(x)\,dx\ .$$

La loi de la variable aléatoire X est la loi continue sur $\mathbb{R}$, de densité f
Pour déterminer la loi d'une variable aléatoire continue, il faut donc calculer sa densité. De manière équivalente, on détermine la loi d'une variable continue en donnant la probabilité qu'elle appartienne à un intervalle A quelconque.

Une variable aléatoire continue X, de densité f, tombe entre a et b avec une probabilité égale à :

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

La probabilité pour une variable aléatoire continue, de tomber sur un point quelconque est nulle :

$$P[X = a] = \int_{\{a\}} f(x)\,dx = 0\ .$$

Par conséquent :

$$\mathbb{P}[\,X \in [a,b]\,] = \mathbb{P}[\,X \in [a,b[\,] = \mathbb{P}[\,X \in]a,b]\,] = \mathbb{P}[\,X \in]a,b[\,]\,.$$

1.3 Fonction de répartition

Définition 3.

Soit X une variable aléatoire, on appelle **fonction de répartition** de X la fonction définie sur $\mathbb{R}$ par

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X \in]-\infty, x]) = \int_{-\infty}^{x} f(t)dt$$

Remarque 1.

Dans le cas où F est dérivable,

$$\forall x \in \mathbb{R} \quad , F'(x) = f(x)$$

Propriété 1.

La définition nous permet d'écrire :

1. $\mathbb{P}(a \leq X \leq b) = \mathbb{P}(X \leq b) - P(X \leq a) = F(b) - F(a)$.
2. $\mathbb{P}(X > b) = 1 - F(b)$.

Propriété 2.

La fonction de répartition F d'une variable aléatoire continue X a les propriétés suivantes :

1. F est une fonction croissante, définie et continue sur $\mathbb{R}$
2. Pour tout $x \in \mathbb{R}$, $0 \leq F(x) \leq 1$
3. $lim_{x \to -\infty} F(x) = 0$ et $lim_{x \to +\infty} F(x) = 1$

Exemple 3 *Soit X une v.a.c dont la fonction de répartition est donnée par :*

$$F(x) = \begin{cases} 1 - x^{-\theta} & si\ x > 1 \\ 0 & sinon, \end{cases}$$

Déterminer la densité de X.

1.4 Espérance et variance

Définition 4.

Soit X une variable aléatoire continue et f sa densité.

— On appelle **espérance** de X le réel, noté $\mathbf{E}(X)$, défini par la relation

$$\mathbf{E}(X) = \int_{-\infty}^{+\infty} x f(x) dx.$$

— On appelle variance de X le réel, noté $\mathbf{V}(X)$, qui, s'il existe, est défini par la relation

$$\mathbf{V}(X) = \int_{-\infty}^{+\infty} (x - \mathbf{E}(X))^2 f(x) dx.$$

— On appelle **écart-type** de X le réel, noté $\sigma(X)$, défini par la relation

$$\sigma(X) = \sqrt{\mathbf{V}(X)}.$$

Exemple 4 *Calculer l'espérance, la variance et l'écart-type pour la fonction f_2 définie dans l'exemple précédent.*

Propriété 3.

Soit X une variable aléatoire continue admettant une espérance et une variance, alors pour tous a et $b \in \mathbb{R}$:

— $E(aX + b) = aE(X) + b$.

— $V(aX + b) = a^2 V(X)$.

— $\sigma(aX + b) = |a|\sigma(X)$.

— $E(aX + Y) = aE(X) + E(Y)$.

Si de plus X et Y sont indépendantes,

— $V(X + Y) = V(X) + V(Y)$.

— $V(X - Y) = V(X) + V(Y)$.

2 Lois usuelles continues :

2.1 La loi uniforme

La loi uniforme est la loi exacte de phénomènes continus uniformément répartis sur un intervalle.
La variable aléatoire X suit une loi uniforme sur le segment $[a;b]$ avec $a < b$ si sa densité de probabilité est donnée par :

$$f(x) = \begin{cases} \frac{1}{b-a}\,; & \text{si } x \in [a;b] \\ 0 & \text{si non.} \end{cases}$$

Exercice 1 (Fonction de répartition)**.**

1. Déterminer la fonction de répartition $F(x)$de la variable aléatoire $X \sim \mathcal{U}([a;b])$
2. Tracer les courbes de $f(x)$ et $F(x)$.

Propriété 4.

L'espérance de la loi uniforme continue vaut :

$$E(X) = \frac{b+a}{2}.$$

La variance de la loi uniforme continue vaut :

$$V(X) = \frac{(b-a)^2}{12}.$$

Exercice 2.

Pour générer un nombre aléatoire de l'intervalle $[0,1]$, Python utilise la commande random().
X est la variable aléatoire égale au nombre affiché avec cette commande.

1. Calculer $P(0,33 \leq X \leq 0,59)$.
2. Quel est le nombre moyen qu'on s'attend à trouver si on repète plusieurs fois cette expérience ?

2.2 La loi Exponentielle :

Une loi exponentielle modélise la durée de vie d'un phénomène sans mémoire, ou sans vieillissement.

Définition 5.

Soit λ un réel strictement positif. Une variable aléatoire X suit une loi exponentielle de paramètre λ lorsque sa densité de probabilité est la fonction f définie sur $[0, +\infty[$ par :

$$f(x) = \begin{cases} \lambda e^{-\lambda x}\,; & \text{si } x \geq 0 \\ 0 & \text{si x ¡0.} \end{cases}$$

2.2.1 Propriétés (à démontrer)

Si X suit la loi exponentielle $\mathcal{E}(\lambda)$, de parmètre $\lambda > 0$ on a :

Fonction de répartition $F(x) = \begin{cases} 1 - e^{-\lambda x}\,; & \text{si } x \geq 0 \\ 0 & \text{si x ¡0.} \end{cases}$

Espérance : $E(X) = 1/\lambda$.

Variance $V(X) = 1/\lambda^2$.

Exercice 3.

A un standard téléphonique, on entend « Votre moyenne de temps d'attente est estimé à 5 minutes».
Quelle est la probabilité d'attendre plus de 10 minutes au total ?

Définition 6.

Une variable aléatoire X, à valeur dans $\mathbb{R}_+$, est dite **sans mémoire** si elle vérifie la propriètè suivante : pour tous $s, t > 0$,on a :

$$P(X > (t+s)|X > t) = P(X > s).$$

Exercice 4 (Loi sans mémoire).

X une variable aléatoire qui suit la loi exponentielle de paramétre 0.2. Montrer que la loi de X est sans mémoire. (i.e : Montrer que pour tous $s, t > 0$, on a : $P(X > (t+s)|X > t) = P(X > s)$.)

2.3 La loi Normale :

2.3.1 Définition

Cette loi est celle qui rend compte de diverses mesures d'une grandeur donnée, opérées à diverses reprises, chaque mesure étant sujette à des erreurs. La loi normale (ou de Laplace-Gauss, appelée « normale » par Pearson en 1893) est la loi de certains phénomènes continus qui fluctuent autour d'une valeur moyenne μ, de manière aléatoire, résultante d'un grand nombre de causes indépendantes dont les effets s'ajoutent sans que l'un d'eux soient dominant : par exemple la taille d'un individu en cm, influencée par le sexe, la nourriture, l'environnement, l'hérédité, le lieu géographique . . .

*

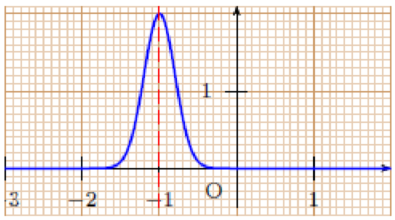
Définition 7.

On appelle loi Normale de paramètres $m \in \mathbb{R}$ (la moyenne) et $\sigma > 0$ (l'écart-type) la loi d'une variable aléatoire continue X prenant toutes les valeurs réelles, de densité de probabilité la fonction définie pour tout $x \in \mathbb{R}$ par

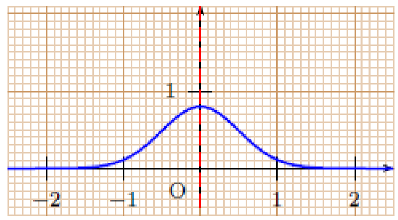
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}(\frac{x-m}{\sigma})^2}$$

On note $X \sim N(m; \sigma)$.

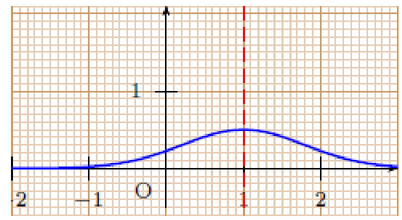
Exemple 5 *Voici des exemples de courbes pour quelques valeurs de m et σ :*



$m = -1$ et $\sigma = 0,2$



$m = 0$ et $\sigma = 0,5$



$m = 1$ et $\sigma = 0,8$

Remarque 2.

Dans l'exemple précédent, on peut observer :

— que la courbe admet comme axe de symétrie la droite d'équation x = m,

— que le maximum de la courbe est atteint en m, espérance de la variable X (ce maximum valant $\frac{1}{\sigma\sqrt{2\pi}}$)

— et que plus σ est grand, plus la courbe « s'étale » autour de la moyenne, en accord avec la signification de l'écart-type.

2.3.2 La loi normale centrée réduite

Définition 8.

La variable aléatoire Z qui suit la loi normale de paramètres $m = 0$ et $\sigma = 1$ est dite **variable aléatoire centrée réduite**. Sa densité de probabilité est définie sur $\mathbb{R}$ par

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}z^2}$$

On note Z la variable aléatoire suivant la loi $N(0;1)$. Tracer la densité de probabilité $f(z)$

— $P(Z \geq z) = 1 - F(z)$.

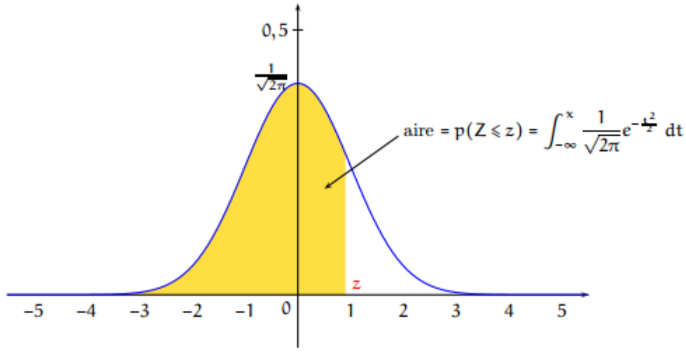
— Si z est positif : $F(-z) = 1 - F(z)$.

— Pour tous $a; b \in \mathbb{R}$, avec $a \leq b$: $P(a \leq Z \leq b) = F(b) - F(a)$.

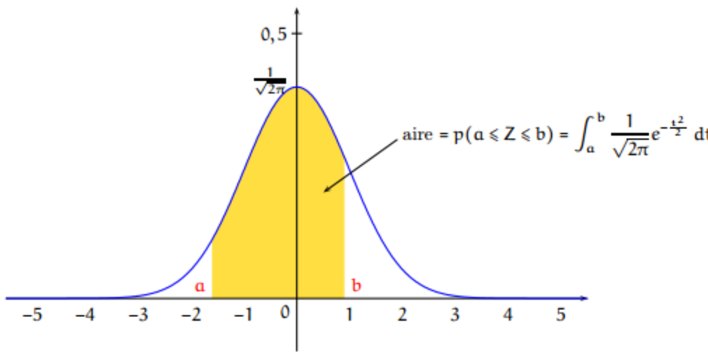
— Pour tout $z \geq 0$, $P(-z \leq Z \leq z) = 2F(z) - 1$.

Démonstrations et interprétations graphiques :

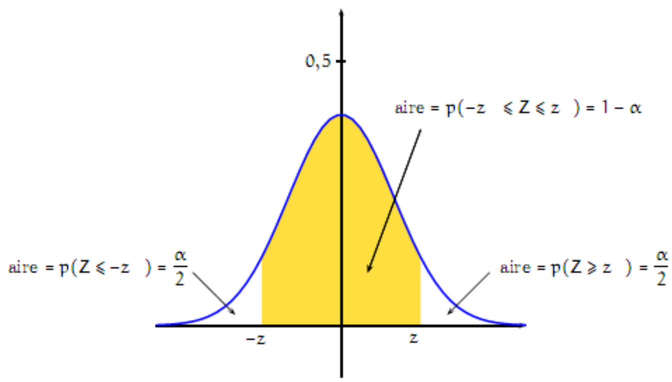
— $F(z) = P(Z < z) = \int_{-\infty}^{z} f(t)dt$



— $P(Z \geq z) = 1 - P(Z < z) = 1 - F(z)$
— $P(a \leq Z \leq b) = F(b) - F(a)$.



— $F(-z) = P(Z < -z) = P(Z > z) = 1 - P(Z < z) = 1 - F(z)$.
— $P(-z \leq Z \leq z) = F(z) - F(-z) = F(z) - (1 - F(z)) = 2F(z) - 1$



2.3.3 Utilisation de la table de la loi normale

Le formulaire ne donne que les valeurs de la loi normale centrée réduite et pour des valeurs positives. En voici quelques exemples pour comprendre la méthode de lecture :

— Calcul de $P(Z \geq 1.25)$: Le nombre situé à l'intersection de la colonne $0,05$ et de la ligne 1.2 est la valeur de $1 - F(t)$ pour $z = 1.2 + 0.05 = 1.25$. Ainsi $P(Z \geq 1.25) = 0.1056$.
— Calcul de $P(Z \leq 1.36)$: $P(Z \leq 1.36) = 1 - P(Z \geq 1.36) = 1 - 0.08691 = 0.9131$.

— Calcul de $P(Z \geq -1.17)$: $P(Z \geq -1.17) = 1 - P(Z \geq 1.17) = 1 - 0.121 = 0.879$

— Calcul de $P(1.15 \leq Z \leq 1.37)$: $P(1.15 \leq Z \leq 1.37) = F(1.37) - F(1.15) = P(Z \leq 1.37) - P(Z \leq 1.15) = 1 - P(Z \geq 1.37) - (1 - P(Z \geq 1.15)) = 0.0398$.

Remarque 3.

Si une variable aléatoire X suit la loi normale $N(m; \sigma)$, alors la variable aléatoire $Z = \frac{X-m}{\sigma}$ suit la loi normale centrée réduite $N(0; 1)$, avec $E(Z) = 0$ et $\sigma(Z) = 1$.

Ce résultat est très important, puisqu'alors il nous suffit d'étudier la loi normale centrée réduite puis de procéder à un changement de variable pour obtenir n'importe quelle loi normale !

Exemple 6 *Une variable X suit la loi normale de paramètres $m = 12$ et $\sigma = 3$. Calculer $P(X \leq 16)$.*

2.3.4 Stabilité de la loi normale

Soit X une variable aléatoire suivant la loi normale $N(m; \sigma)$. Alors, pour tous $a; b \in \mathbb{R}$:

— La variable aléatoire $aX + b$ suit la loi normale $N(am + b; |a|\sigma)$,

Si de plus Y suit une loi normale $N(m'; \sigma')$ (X et Y sont indépendantes), alors

— La variable aléatoire $X + Y$ suit une loi normale, $N(m + m'; \sqrt{\sigma^2 + \sigma'^2})$.

— La variable aléatoire $X - Y$ suit une loi normale $N(m - m'; \sqrt{\sigma^2 + \sigma'^2})$.

Exercice 5.

Soit X une v.a qui suit la loi $N(1; 3)$ et Y une v.a qui suit la loi $N(-1; 1)$. Sachant que X et Y sont indépendantes, determiner la loi de :

1. $-2X + 1$.
2. $X + Y$.
3. $X - Y$.

2.4 La loi du χ^2 (Khi-deux)

La loi de Pearson ou loi de χ^2 (Khi-deux) trouve de nombreuses applications dans le cadre de la comparaison de proportions, des tests de conformité d'une distribution observée à une distribution théorique et le test d'indépendance de deux caractères qualitatifs. Ce sont les test du khi-deux.
Soit $X_1, X_2, \ldots, X_i, \ldots, X_n$, n variables aléatoires indépendantes normales centrées réduites, on appelle Y la variable aléatoire définie par

$$Y = X_1^2 + X_2^2 + \ldots + X_n^2 = \sum_{i=1}^{n} X_i^2$$

On dit que Y suit une loi de χ^2 à n degrés de liberté (d.d.l.). et on note

$$Y \sim \chi_n^2.$$

Propriété 5.

Si U suit une loi de Pearson à n d.d.l., V suit une loi de Pearson à m d.d.l., et si U et V sont indépendantes alors $U + V$ suit une loi de Pearson à $(n + m)$ ddl.

— **Calcul des probabilités dans la distribution khi-deux**

Définition 9.

Soit $0 \leq \alpha \leq 1$, le quantile d'ordre $1 - \alpha$, noté par $y_{n,\alpha}$, est défini par :

$$P(T > y_{n,\alpha}) = \alpha.$$

Si $Y \sim \chi_n^2$ alors quelle est la valeur de la probabilité suivante :

$$P(Y > y) = \alpha$$

À l'aide de la table de la loi du Khi-deux (*Voir Annexe*), chercher la valeur située à la première ligne dans la colonne qui s'intersecte avec la $n^{iéme}$ (n=d.l.l) ligne en y.

Exemple 7 .

1. *Si* $Y \sim \chi^2_{20}$ *et* $P(Y > y) = 0.01$ *chercher la valeur de* y.
2. *Si* $Y \sim \chi^2_{11}$ *et* $P(Y > y) = 0.1$ *chercher la valeur de* y.

2.5 La loi de Student

La loi de Student (ou loi de Student-Fisher) est utilisée lors des tests de comparaison de paramètres comme la moyenne et dans l'estimation de paramètres de la population à partir de données sur un échantillon (Test de Student).
Soit U une variable aléatoire suivant une loi normale réduite $N(0,1)$ et V une variable aléatoire suivant une loi du Khi-deux à n degrés de liberté χ^2_n , U et V étant indépendantes, on dit alors que

$$T = \frac{U}{\sqrt{\frac{V}{n}}}$$

suit une loi de Student à n degrés de liberté.

Propriété 6.

— L'espérance de la variable de Student est : $E(T) = 0$ si $n > 1$
— La variance de la variable de Student est : $V(T) = \frac{n}{n-2}$.

Définition 10.

Soit $0 \leq \alpha \leq 1$, le quantile d'ordre $1-\alpha$, noté par $t_{n,\alpha}$, est défini par :

$$P(T > t_{n,\alpha}) = \alpha.$$

— **Calcul des probabilités dans la distribution de student.**

Si $Y \sim T_n$ alors quelle est la valeur de la probabilité suivante :

$$P(Y > y) = \alpha$$

À l'aide de la table de la loi de student (*Voir Annexe*), chercher la valeur située à la première ligne dans la colonne qui s' intersecte avec la $n^{\text{iéme}}$ (n=d.l.l) ligne en $t_{n,\alpha}$.

Exemple 8

1. *Si* $T \sim T_{10}$, *chercher* $P(T > 1.372)$.
2. *Si* $T \sim T_{20}$, *chercher* $P(T > 2.197)$.
3. *Si* $T \sim T_{20}$ *et* $P(T > t) = 0.005$ *chercher la valeur de* y.
4. *Si* $T \sim T_{11}$ *et* $P(T > t) = 0.1$ *chercher la valeur de* t.