Variables aléatoires - Lois usuelles

G. Marot-Briend
guillemette.marot@univ-lille.fr

2021-2022

Plan

Introduction

Exemple introductif

Considérons une famille de 3 enfants tirée au hasard (expérience aléatoire). On s'intéresse au nombre de garçons dans la famille.

L'ensemble fondamental Ω associé à cette expérience aléatoire est :

$$\Omega = \{FFF, FFG, FGF, FGG, GFF, GFG, GGF, GGG\},\ card(\Omega) = 8$$

Associons à chaque élément ω de Ω le nombre de garçons :

ω	FFF	FFG	FGF	FGG	GFF	GFG	GGF	GGG
Nb de garçons	0	1	1	2	1	2	2	3

Introduction

Remarque : 2 événements élémentaires ω de Ω peuvent conduire au même nombre.

Définition : On appelle variable aléatoire (v.a.) X, l'application :

$$\begin{aligned} X : & \Omega \to E \\ & \omega \to X(\omega) \end{aligned}$$

Dans notre exemple :

$$E \subseteq \mathbb{N}$$

$$X(\omega) = \text{ Le nb de garçons pour une configuration possible } \omega$$

$x = X(\omega)$ est une réalisation de X

Introduction

Autrement dit, une variable aléatoire (v.a.) est une règle qui, à tout événement associe un nombre réel

Une variable aléatoire peut être de deux types :

1. **Discrète** si E est fini ou dénombrable ($E \subseteq \mathbb{N}$) (Nombre de garçons...)
2. **Continue** si $E \subseteq \mathbb{R}$ (Taille de l'individu)

Variables aléatoires discrètes

On appelle variable aléatoire discrète (v.a.d.) X l'application :

$$\begin{aligned} X : & \Omega \to E \subseteq \mathbb{N} \\ & \omega \to X(\omega) \end{aligned}$$

L'exemple avec le nombre de garçons dans une famille de 3 enfants est un exemple de v.a.d..

Afin de modéliser l'expérience aléatoire, on cherche à déterminer la probabilité de chaque valeur possible de X. Dans l'exemple sur le nombre de garçons, l'ensemble des valeurs possibles de X est donné par :

$$E = \{0, 1, 2, 3\}$$

Variables aléatoires discrètes

Calculons la probabilité pour chaque valeur possible de X (Nombre de garçons) :

ω	FFF	FFG	FGF	FGG	GFF	GFG	GGF	GGG
X	0	1	1	2	1	2	2	3

$$P(X=0) = \frac{card(X=0)}{card(\Omega)} = 1/8$$

$$P(X=1) = 3/8$$
$$P(X=2) = 3/8$$
$$P(X=3) = 1/8$$

Si on connait $P(X=x)$ $\forall x \in E$ alors on peut prévoir l'ensemble du phénomène aléatoire caractérisé par X ! !

On a défini la **loi de probabilité** de la variable aléatoire X

Loi de probabilité

Définition : On appelle loi de probabilité d'une variable aléatoire discrète la donnée de

- l'ensemble des valeurs possibles $x_i, i \in I \subseteq \mathbb{N}$
- la probabilité de chaque valeur $P(X = x_i)$

On note $f(x_i) = P(X = x_i)$

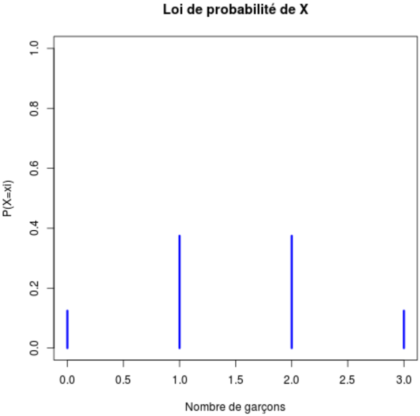
Propriétés :

- $\forall i,\ \ f(x_i) \geqslant 0$
- $\sum_i f(x_i) = 1$

Variables aléatoires discrètes

Représentations de la loi de probabilité de X :
tableau ou graphique

x_i	$P(X = x_i)$
0	$1/8$
1	$3/8$
2	$3/8$
3	$1/8$



Remarque : $1/8 + 3/8 + 3/8 + 1/8 = 1$

Variables aléatoires discrètes

La loi de probabilité de X nous permet donc de connaitre

$$\forall x \in E,\ P(X = x_i) = f(x_i)$$

Elle permet également de répondre à la question $P(X \in]a, b])$:

$$P(X \in]a, b]) = \sum_{x_i \in]a,b]} f(x_i)$$

Exemple avec le nombre de garçons :

$$P(X \in]1, 3]) = \sum_{x_i \in]1,3]} P(X = x_i) = f(2) + f(3) = 3/8 + 1/8 = 1/2$$

Variables aléatoires discrètes

Définition : On appelle Fonction de répartition d'une variable aléatoire l'application F telle que :

$$\begin{aligned} F : & E \to [0, 1] \\ & x \to F(x) = P(X \leq x) \end{aligned}$$

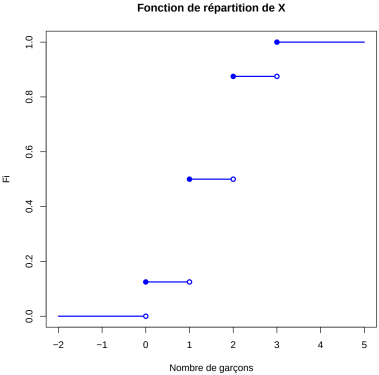
Dans le cas d'une variable aléatoire discrète $E = \{x_1, x_2, \ldots, x_k\}$, la fonction de répartition a pour expression :

$$F(x_j) = P(X \leq x_j) = \sum_{i=1}^{j} P(X = x_i)$$

Variables aléatoires discrètes

Retour à l'exemple :

x	$F(x) = P(X \leq x)$
0	$1/8$
1	$4/8$
2	$7/8$
3	1



$$F(2) = P(X = 0) + P(X = 1) + P(X = 2) = 7/8$$

Utile pour calculer un intervalle : $P(X \in]1, 2]) = F(2) - F(1)$

Variables aléatoires discrètes

Soit une variable aléatoire discrète X prenant les valeurs dans l'ensemble $E = \{x_1, x_2, \ldots, x_k\}$

Définition : On appelle espérance mathématique d'une variable aléatoire discrète X, la quantité notée $E[X]$:

$$E[X] = \sum_{i=1}^{k} x_i P(X = x_i)$$

Exemple : *Nombre moyen de garçons dans une famille de 3 enfants*

$$E[X] = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} = 3/2$$

Variables aléatoires discrètes

Propriétés

- $E[\lambda X] = \lambda E[X]$ avec λ une constante
- $E[X + Y] = E[X] + E[Y]$
- $E[\lambda X + \delta Y] = \lambda E[X] + \delta E[Y]$ avec λ et δ des constantes
- $E[\lambda] = \lambda$

Variables aléatoires discrètes

Exemple : *Un jeu de hasard*

On jette un dé équilibré. On regarde le résultat :

- 1 : on perd 175 euros
- 2 : on perd 50 euros
- 3 : on ne gagne rien
- 4 on gagne 25 euros
- 5 : on gagne 50 euros
- 6 : on gagne 120 euros

Est-il intéressant de jouer à ce jeu ?

Variables aléatoires discrètes

Modélisation mathématique

Soit une expérience aléatoire (lancé de dé équilibré) d'ensemble fondamental équiprobable $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Soit X une v.a.d. qui associe à un lancé de dé le gain d'argent.
$E = \{-175, -50, 0, 25, 50, 120\}$

$$E[X] = -175 \times \frac{1}{6} + (-50) \times \frac{1}{6} + 0 \times \frac{1}{6} + 25 \times \frac{1}{6} + 50 \times \frac{1}{6} + 120 \times \frac{1}{6}$$

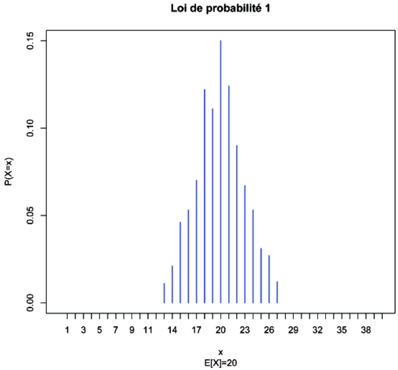
$$E[X] = -5$$

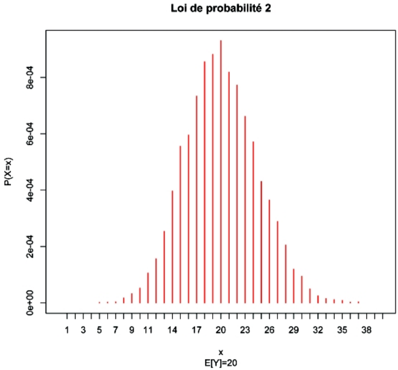
"En moyenne" nous aurons un gain négatif.

Variables aléatoires discrètes

Introduction à la variance

L'espérance mathématique permet de résumer une loi de probabilité, mais seulement **en partie** :





Variables aléatoires discrètes

Les v.a.d. X et Y présentent les mêmes espérances $E[X] = E[Y]$.

Cependant, elles ne sont pas dispersées de la manière : Y est plus dispersée par rapport à sa moyenne que X.

$\Rightarrow$ la variance est un indicateur de dispersion, en complément de l'espérance, pour résumer une loi de probabilité !

Variables aléatoires discrètes

Définition : Soit X une variable aléatoire définie sur un espace probabilisé (Ω, P) et d'espérance mathématique $E[X]$. On appelle variance de X, la quantité notée $V[X]$:

$$V[X] = E[(X - E[X])^2]$$

Si X est discrète et prend ses valeurs dans l'ensemble $E = \{x_1, x_2, \ldots, x_k\}$ alors :

$$V[X] = \sum_{i=1}^{k} (x_i - E[X])^2 P(X = x_i)$$

Note : $V[X] \geqslant 0$ et $\sqrt{V[X]} = \sigma(X)$ écart-type

Variables aléatoires discrètes

Propriétés

- $V[X] = E[X^2] - (E[X])^2$ (Huygens)
- $V[aX] = a^2 V[X]$ avec a constante
- $V[a] = 0$ avec a constante
- $V[X + a] = V[X]$ avec a constante

Variables aléatoires discrètes

Exemple : *famille de 3 enfants*

$$E[X] = \frac{3}{2}$$

$$V[X] = \sum_{i=1}^{4} (x_i - E[X])^2 P(X = x_i)$$

$$V[X] = (0-3/2)^2 \times \frac{1}{8} + (1-3/2)^2 \times \frac{3}{8} + (2-3/2)^2 \times \frac{3}{8} + (3-3/2)^2 \times \frac{1}{8}$$

$$V[X] = \frac{3}{4} = 0.75$$

Variables aléatoires continues

Exemple introductif

Soit X une variable aléatoire qui associe à un bébé son poids à la naissance.

X peut prendre toutes les valeurs possibles dans un intervalle.

X est une **variable aléatoire continue**.

Définition : On appelle variable aléatoire continue l'application X telle que :

$$X : \Omega \to E \subseteq \mathbb{R}$$
$$\omega \to X(\omega)$$

Variables aléatoires continues

Comme pour les variables aléatoires discrètes, on résume une variable aléatoire continue par sa loi de probabilité, appelée dans ce cas densité de probabilité.

Définition : une variable aléatoire continue X a pour densité de probabilité f si la fonction vérifie :

$$P(X \in [a,b]) = \int_a^b f(x).dx$$

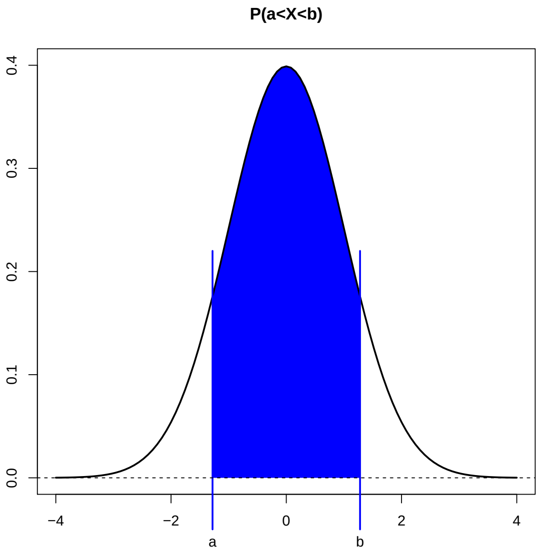
Dans ce cas,

$$\int_{-\infty}^{+\infty} f(x).dx = 1 \text{ et } \forall x \in \mathbb{R},\ f(x) \geqslant 0$$

Et

$$\forall x \in E,\ P(X = x) = 0$$

Variables aléatoires continues



Variables aléatoires continues

De manière générale, on appelle densité de probabilité toute fonction f telle que

- $\int_{-\infty}^{+\infty} f(x).dx = 1$
- $\forall x \in \mathbb{R},\ f(x) \geqslant 0$

Exercice :
Soit une fonction f définie par :

$$\forall x \in \mathbb{R},\ f(x) = \begin{cases} kx & \text{si } x \in [0,1] \\ 0 & \text{sinon.} \end{cases}$$

Trouver k tel que f soit une densité de probabilité.

Variables aléatoires continues

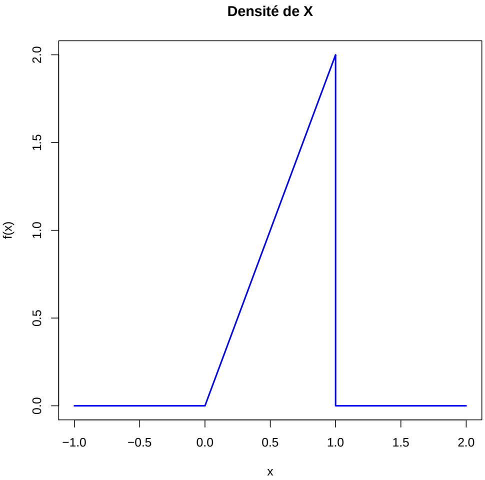
$\forall x \in [0,1],\ f(x) \geqslant 0$ donc $k \geqslant 0$

$$\int_{\mathbb{R}} f(x).dx = \int_0^1 f(x).dx$$
$$= \left[k.\frac{x^2}{2}\right]_0^1$$
$$= \frac{k}{2}$$

On veut que $\int_{\mathbb{R}} f(x).dx = 1$ donc $k = 2$

Variables aléatoires continues

$$f(x) = \begin{cases} 2x & \text{si } x \in [0,1] \\ 0 & \text{sinon.} \end{cases}$$



Exercice : Calculer $P(X \in [1/2, 1])$

Variables aléatoires continues

$$
\begin{aligned}
P(X \in [1/2, 1]) &= \int_{1/2}^{1} f(x).dx \\
&= \int_{1/2}^{1} 2x.dx \\
&= \left[x^2\right]_{1/2}^{1} \\
&= 1 - \frac{1}{4} \\
&= \frac{3}{4}
\end{aligned}
$$

Note : il peut devenir fastidieux de calculer $P(X \in [a, b])$ à cause du calcul d'intégrale

On préfère utiliser la **fonction de répartition de** X

Variables aléatoires continues

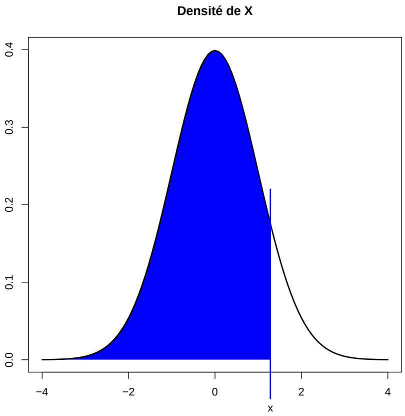
Définition : Soit X une variable aléatoire continue sur un espace probabilisé (Ω, P). On appelle fonction de répartition de X l'application F telle que :

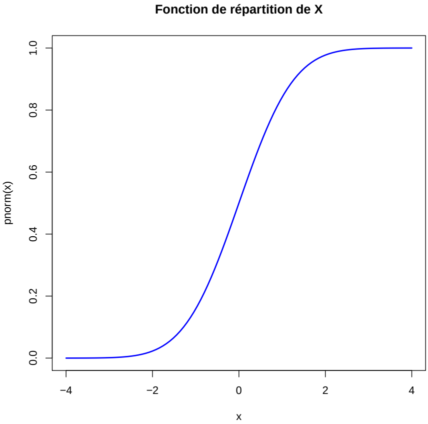
$$F : \mathbb{R} \to \mathbb{R}^+$$
$$x \to F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t).dt$$

$$P(X \in [a, b]) = \int_{a}^{b} f(t).dt = F(b) - F(a)$$

Remarque : on pourra évidemment utiliser aussi les notations suivantes : $F(u) = \int_{-\infty}^{u} f(x).dx$

Variables aléatoires continues





Variables aléatoires continues

Propriétés

F est croissante

$$\lim_{x \to -\infty} F(x) = 0$$

$$\lim_{x \to +\infty} F(x) = 1$$

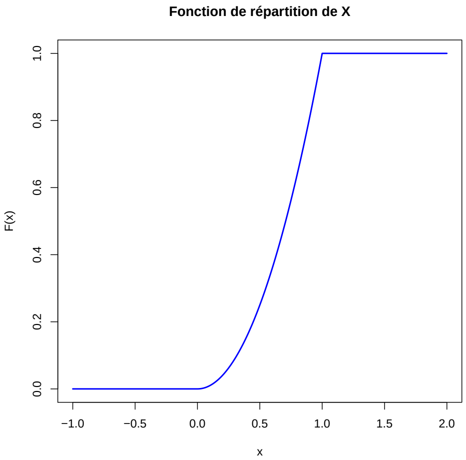
$$F' = f$$

Variables aléatoires continues

Retour à l'exemple

$$f(x) = \begin{cases} 2x & \text{si } x \in [0,1] \\ 0 & \text{sinon.} \end{cases}$$

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ x^2 & \text{si } x \in [0,1] \\ 1 & \text{si } x > 1 \end{cases}$$



Variables aléatoires continues

$$\begin{aligned} P(X \in [1/2, 1]) &= F(1) - F(1/2) \\ &= 1 - 1/4 \\ &= 3/4 \end{aligned}$$

Le calcul est plus rapide qu'en passant à chaque fois par la densité (calcul d'intégrale).

Variables aléatoires continues

Espérance mathématique

Motivation : *Quel est le poids moyen d'un bébé à la naissance ?*

Définition : Soit X une variable aléatoire continue sur un espace probabilisé (Ω, P). On appelle espérance mathématique de X, la quantité notée $E[X]$:

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

Variables aléatoires continues

Retour à l'exemple

$$f(x) = \begin{cases} 2x & \text{si } x \in [0,1] \\ 0 & \text{sinon.} \end{cases}$$

$$E[X] = \int_{-\infty}^{+\infty} xf(x).dx = \int_0^1 xf(x).dx = \int_0^1 x2x.dx = \int_0^1 2x^2.dx$$

$$\int_0^1 2x^2.dx = \left[\frac{2x^3}{3}\right]_0^1 = 2/3 - 0 = 2/3$$

$$E[X] = 2/3$$

Variables aléatoires continues

Variance

Définition : Soit X une variable aléatoire continue sur un espace probabilisé (Ω, P) et d'espérance $E[X]$. On appelle variance de X, la quantité notée $V[X]$:

$$V[X] = E[(X - E[X])^2] = E[X^2] - E[X]^2$$

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

Variables aléatoires continues

Retour à l'exemple

$$f(x) = \begin{cases} 2x & \text{si } x \in [0,1] \\ 0 & \text{sinon.} \end{cases}$$

Et $E[X] = 2/3$

$$V[X] = \int_{-\infty}^{+\infty} x^2 f(x).dx - E[X]^2 = \int_0^1 2x^3.dx - (2/3)^2$$

$$V[X] = \left[2\frac{x^4}{4}\right]_0^1 - 4/9 = \left[\frac{x^4}{2}\right]_0^1 - 4/9 = 1/2 - 0 - 4/9$$

$$V[X] = 1/18$$

Récapitulatif

Variable aléatoire (v.a.) : une règle qui, à tout événement, associe un nombre réel

- **v.a. discrète**

⇒ la loi de probabilité est caractérisée par les valeurs possibles et les probabilités associées.

- **v.a. continue**

⇒ la loi de probabilité associe une probabilité à chaque ensemble de valeurs définies dans un **intervalle donné**. Lorsque cet intervalle tend vers 0, la valeur prise par X tend alors vers une fonction que l'on appelle densité de probabilité.

La fonction de répartition correspond aux probabilités cumulées associées à la variable aléatoire.

- **v.a. discrète** ⇒ somme
- **v.a. continue** ⇒ primitive de la densité de probabilité.

Rappels

En pratique... calcul de la loi de probabilité

Loi des grands nombres

Si on répète N fois une expérience dans laquelle la probabilité d'apparition d'un événement A est p, la fréquence de cet événement au cours des N expériences tend vers p lorsque N tend vers l'infini.

Lorsque le nombre d'épreuves augmente indéfiniment, les **fréquences observées** tendent vers les **probabilités** et les **distributions observées** vers les lois de probabilité.

Exercices

Exercice 1 :

On suppose que la probabilité d'être immunisé contre la tuberculose est égale à 0.8

1. Si on s'intéresse à 1000 personnes, quelle est la proportion attendue de personnes atteintes de la tuberculose ?
2. On s'intéresse aux 10 premières personnes qu'on rencontre. Aucune d'elles n'est atteinte de la tuberculose. Expliquez pourquoi.

Exercices

Exercice 1 :

On suppose que la probabilité d'être immunisé contre la tuberculose est égale à 0.8

1. Si on s'intéresse à 1000 personnes, quelle est la proportion attendue de personnes atteintes de la tuberculose ?
2. On s'intéresse aux 10 premières personnes qu'on rencontre. Aucune d'elles n'est atteinte de la tuberculose. Expliquez pourquoi.

Correction :

1. La proportion attendue est de 20%.
2. On ne peut pas appliquer la loi des grands nombres car l'effectif est trop petit. Il n'y a donc aucune raison d'observer la proportion attendue.

Plan

Lois discrètes

Par définition, les variables aléatoires discrètes prennent des valeurs entières discontinues sur un intervalle donné. Ce sont généralement le résultat de dénombrement.

Exemples de lois discrètes :

- loi uniforme
- loi de Bernoulli
- loi binomiale
- loi de Poisson

Loi uniforme

Une distribution de probabilité suit une loi uniforme lorsque toutes les valeurs prises par la variable aléatoire sont équiprobables.
Si n est le nombre de valeurs différentes prises par la variable aléatoire,

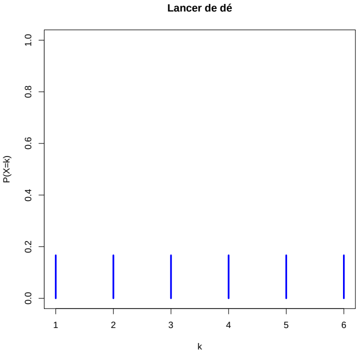
$$\forall i, P(X = x_i) = \frac{1}{n}$$

Exemple : lancer de dé (non pipé)
Donner la distribution de probabilité des chiffres obtenus au lancer de dé et la représenter graphiquement.

Exercices

Correction :

k	1	2	3	4	5	6
$P(X=k)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$



Loi de Bernoulli

$\Omega = \{E, S\}$ (S pour succès et E pour échec)

X v.a. discrète : « nombre de succès »

$$\begin{cases} X = 1 \text{ si S est réalisé} \\ X = 0 \text{ si E est réalisé} \end{cases}$$

On appelle variable de Bernoulli ou variable indicatrice, la variable aléatoire X telle que :

$$X : \Omega \to R \qquad\qquad X(\Omega) = 0, 1$$

Loi de Bernoulli

La loi de probabilité associée à la variable de Bernoulli X telle que

$$\begin{cases} P(X=0)=q \\ P(X=1)=p \quad \text{avec } p+q=1 \end{cases}$$

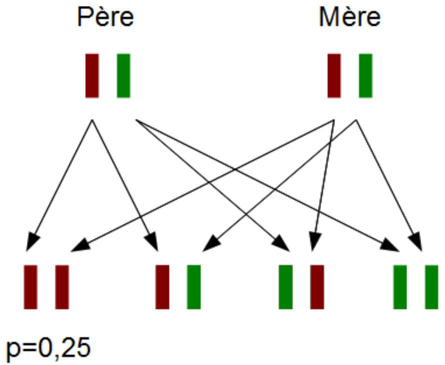
est appelée loi de Bernoulli notée $\mathcal{B}(1,p)$

Exercice : Maladie génétique.

Dans la mucoviscidose, l'allèle responsable de la maladie est récessif. Pour un enfant dont les parents sains portent cet allèle, son statut M ("atteint de la mucoviscidose") définit une épreuve de Bernoulli. Calculer p et q.

Exercices

Correction :



p=0,25 donc q=1-p=0,75

Loi de Bernoulli

Exercice : Notons X une variable aléatoire qui suit une loi de Bernoulli. Calculer E(X) et V(X).

Loi de Bernoulli

Exercice : Notons X une variable aléatoire qui suit une loi de Bernoulli. Calculer E(X) et V(X).

Réponses :

$$\begin{aligned} \mathrm{E}(X) &= \sum x_i P(X = x_i) \\ &= 1 * p + 0 * q \\ &= p \end{aligned}$$

$$\begin{aligned} \mathrm{Var}(X) &= E(X^2) - E(X)^2 \\ &= 1^2 * p + 0^2 * q - p^2 \\ &= p(1 - p) \end{aligned}$$

A RETENIR

$$\boxed{\mathrm{E}(X) = p\,; \quad \mathrm{Var}(X) = pq}$$

Loi binomiale

Variable binomiale $\Rightarrow S_n$: nombre de succès obtenus lors de la répétition de n épreuves de Bernoulli identiques et indépendantes.

$$S_n = \sum_{i=1}^{n} X_i \sim B(n, p)$$

Probabilité de l'obtention de k succès au cours de n épreuves indépendantes :

$$P(S_n = k) = \binom{n}{k} p^k q^{n-k}$$

avec $$\binom{n}{k} = C_n^k = \frac{A_n^k}{k!} = \frac{n!}{(n-k)!k!}$$

Loi binomiale

$$\begin{aligned}
\mathrm{E}(S_n) &= \mathrm{E}(\sum X_i) \\
&= \sum \mathrm{E}(X_i) \\
&= np
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}(S_n) &= \mathrm{Var}(\sum X_i) \\
&= \sum \mathrm{Var}(X_i) \text{ car les } X_i \text{ sont indépendantes} \\
&= \sum p(1-p) \\
&= np(1-p)
\end{aligned}$$

A RETENIR

$$\boxed{\mathrm{E}(S_n) = np, \quad \mathrm{Var}(S_n) = npq}$$

Loi binomiale

Exercice :

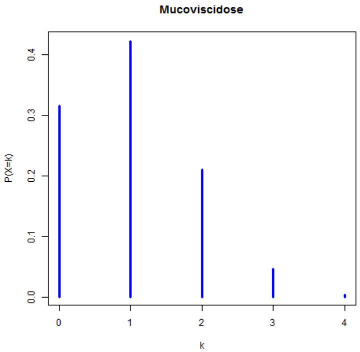
On s'intéresse à une famille de 4 enfants dont les parents sont porteurs de l'allèle récessif responsable de la mucoviscidose. On étudie la probabilité que k enfants soient atteints de cette maladie dans cette famille.

1- Donner la loi de probabilité et tracer la distribution de probabilités (diagramme en batons)
2- Quel est en moyenne le nombre attendu d'enfants malades ?
3- Donner la variance et l'écart type de cette variable.

Loi binomiale

k	0	1	2	3	4
$\binom{n}{k}$	1	4	6	4	1
$P(X=k)$	0.316	0.422	0.211	0.047	0.004

$$\binom{4}{3} = \frac{4!}{1!3!} = \frac{4*3*2*1}{1*3*2*1} = 4$$



$$\begin{aligned} E(X) &= np = 4*0.25 \\ &= 1 \end{aligned}$$

$$\begin{aligned} \mathrm{Var}(X) &= npq = 4*0.25*(1-0.25) \\ &= 0.75 \\ \sigma_X &= \sqrt{\mathrm{Var}(X)} \\ &= 0.87 \end{aligned}$$

Loi binomiale

Exercice (suite) :

La prévalence de la mucoviscidose dans la population française est autour de $p = 2/10000$.
On s'intéresse maintenant à un échantillon de 10000 personnes.

Quelle est la probabilité que dans cet échantillon on ait 2 personnes malades ?

Loi binomiale

Exercice (suite) :

La prévalence de la mucoviscidose dans la population française est autour de $p = 2/10000$.
On s'intéresse maintenant à un échantillon de 10000 personnes.

Quelle est la probabilité que dans cet échantillon on ait 2 personnes malades ?

$$P(X = 2) = \binom{10^4}{2}.(2.10^{-4})^2.(1 - 2.10^{-4})^{10^4-2}$$

n grand $\Rightarrow$ calcul des probabilités d'une loi binomiale fastidieux

Loi de Poisson

Comportement asymptotique :

si $n \to \infty$ et $p \to 0$
alors $X : B(n, p) \to P(\lambda)$ avec $np \to \lambda$
Approximation de la loi binomiale par la loi de Poisson correcte quand $n \geq 30$ *et* $np \leq 10$

loi de Poisson : loi des "événements rares"

$$P(X = k) = \frac{\exp^{-\lambda} \lambda^k}{k!}$$

A RETENIR

$$\boxed{\mathrm{E}(X) = \lambda\,; \quad \mathrm{Var}(X) = \lambda}$$

Loi de Poisson

Exemple : Maladie rare précédente $p = 2/10000$

1- Quelle est la probabilité que dans un échantillon de 10000 personnes, on ait 0,1 ou 2 malades ?
2- Quelle est la probabilité qu'il y ait au moins 3 malades ?

Loi de Poisson

<u>Exemple :</u> Maladie rare précédente $p = 2/10000$

1- Quelle est la probabilité que dans un échantillon de 10000 personnes, on ait 0,1 ou 2 malades ?
2- Quelle est la probabilité qu'il y ait au moins 3 malades ?

Réponses : $\lambda = np = 2$

- $P(X = 0) = 0.13; \quad P(X = 1) = 0.27; \quad P(X = 2) = 0.27$
- $P(X \geq 3) = 1 - P(X \leq 2) = 1 - (0.13 + 0.27 * 2) = 0.33$

Lois usuelles continues

X v.a. continue si elle peut prendre toutes les valeurs dans un intervalle donné.

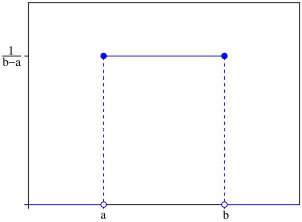
Lois de base :

- loi uniforme
- loi normale

loi uniforme

La v.a. X suit une loi uniforme sur le segment [a,b] avec $a < b$ si sa densité de probabilité est donnée par :

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



loi uniforme

Espérance :
Montrer que
$$E(X) = \frac{b+a}{2}$$

Variance :
On admettra que
$$\mathrm{Var}(X) = \frac{(b-a)^2}{12}$$

Exercice : En soirée, un métro part toutes les 6 minutes de la station CHU - Eurasanté. On note X le temps d'attente à la station. On suppose que X suit une loi uniforme sur $[0, 6]$. Quelle est la probabilité qu'une personne attende entre 3 et 5 minutes ?

loi uniforme

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} xf(x)dx \\
&= \underbrace{\int_{-\infty}^{a} xf(x)dx}_{0} + \int_{a}^{b} xf(x)dx + \underbrace{\int_{b}^{\infty} xf(x)dx}_{0} \\
&= \int_{a}^{b} x(\frac{1}{b-a}).dx \\
&= \frac{1}{(b-a)}\left[\frac{x^2}{2}\right]_a^b = \frac{1}{(b-a)}\frac{1}{2}(b^2-a^2) \\
&= \frac{1}{2(b-a)}(b-a)(b+a) \\
E(X) &= \frac{b+a}{2}
\end{aligned}$$

Loi uniforme

Rappel :

Pour toute loi continue, on peut calculer la loi de probabilité associée à chaque ensemble de valeurs définies dans un intervalle $[a, b]$

- via la densité de probabilité

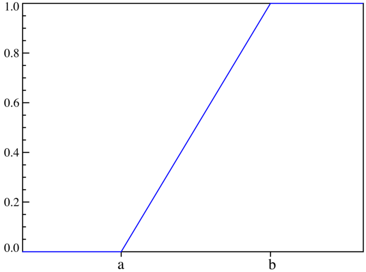
$$P(X \in [c, d]) = \int_c^d f(x)\, dx$$

- via la fonction de répartition $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt$

$$P(X \in [c, d]) = F(d) - F(c)$$

loi uniforme

Fonction de répartition



- si $x \leq a$, $F(x) = 0$
- si $x \geq b$, $F(x) = 1$

- si $x \in [a, b]$

$$\begin{aligned} F(x) &= P(X \leq x) \\ &= \int_{-\infty}^{x} f(t)\, dt \\ &= \int_{a}^{x} \frac{1}{b-a}\, dt \\ &= \frac{x-a}{b-a} \end{aligned}$$

$$P(X \in [3,5]) = F(5) - F(3) = \frac{5-0}{6-0} - \frac{3-0}{6-0} = \frac{1}{3} \approx 0.33$$

Loi normale

On parle de loi normale (ou loi de Laplace-Gauss) lorsque l'on a affaire à une variable aléatoire continue dépendant d'un grand nombre de causes indépendantes dont les effets s'additionnent et dont aucune n'est prépondérante.

- de nombreux phénomènes suivent une loi très proche de la loi normale
- sous certaines conditions, la moyenne sur un échantillon suit une telle loi.
- certaines lois tendent vers une loi normale sous certaines conditions.

Loi normale

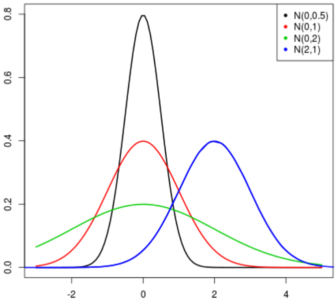
X suit une loi normale de paramètres (μ,σ) si sa densité de probabilité est donnée par :

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

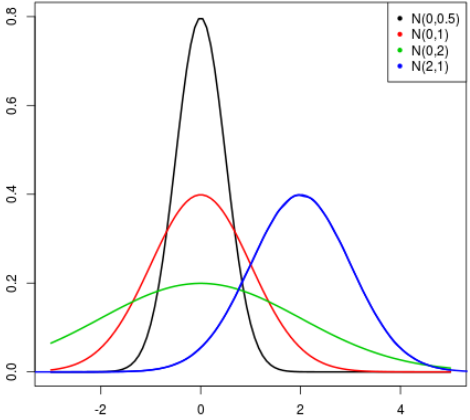
Notation : $X \sim \mathcal{N}(\mu, \sigma)$

<u>Propriété :</u>
symétrie par rapport à μ.
μ : tendance centrale
σ^2 : dispersion autour de la valeur centrale

$E(X) = \mu$
$\mathrm{Var}(X) = \sigma^2$



Loi normale



Densité de probabilité

Loi normale

Fonction de répartition de la loi normale ?

$$F(u) = \int_{-\infty}^{u} f(x)\, dx$$

Loi normale

Fonction de répartition de la loi normale ?

$$F(u) = \int_{-\infty}^{u} f(x)\, dx$$

$$F(u) = \int_{-\infty}^{u} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Le calcul de la fonction de répartition (FDR) n'est pas possible analytiquement, d'où la nécessité d'utiliser des tables.

Méthode des trapèzes $\Rightarrow$ autant de calculs que de lois différentes

Loi normale

Nécessité de normalisation pour utiliser une seule table :

Définition :
Soit X une va continue de moyenne μ et d'écart type σ. On appelle variable centrée réduite la va X^* définie par

$$X^* = \frac{X - \mu}{\sigma}$$

$E(X^*) = 0$ (centrée) $V(X^*) = 1$ (réduite)
Si $X \sim \mathcal{N}(\mu, \sigma)$ alors $X^* \sim \mathcal{N}(0, 1)$. La FDR de X^* est tabulée.
X^* est aussi souvent notée Z ou U.

Loi normale

Présentation des tables

- table de la fonction de répartition $\pi(u)$ de la loi normale centrée réduite
  Cette table indique les probabilités des intervalles sous la forme $]-\infty, u]$ pour des valeurs positives de u.
  C'est la fonction de répartition F telle que $F(u) = \int_{-\infty}^{u} f(z)\, dz$
  Pour des valeurs négatives de u, on utilise la symétrie de la courbe.
- table de distribution de U, loi normale centrée réduite
  On veut déterminer u tel que
  $P(U \leq u) = F(u) = 1 - \alpha$ ou $P(U \geq u) = G(u) = \alpha$
  La table donne u_α en fonction de α

Loi normale

<u>Exercice :</u> Poids des bébés à la naissance

Le poids des bébés à la naissance suit une loi normale de moyenne 3,3 Kg et d'écart-type 0,6 Kg. On note X la variable aléatoire "poids des bébés à la naissance".

1- Calculer la probabilité que $2,12 \leq X \leq 4,48$
2- Déterminer la limite du poids correspondant aux 10% des bébés les plus gros.

Loi normale

Exercice : Poids des bébés à la naissance

Le poids des bébés à la naissance suit une loi normale de moyenne 3,3 Kg et d'écart-type 0,6 Kg. On note X la variable aléatoire "poids des bébés à la naissance".

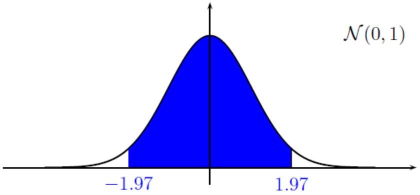
1- Calculer la probabilité que $2,12 \leq X \leq 4,48$
2- Déterminer la limite du poids correspondant aux 10% des bébés les plus gros.

Loi normale

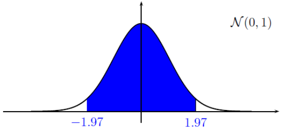
X : poids des bébés à la naissance $X \sim \mathcal{N}(\mu = 3.3, \sigma = 0.6)$
Z : va centrée réduite

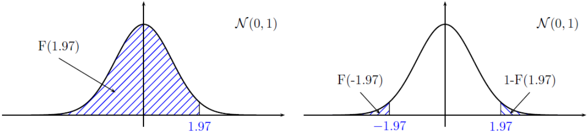
$$Z = \frac{X - \mu}{\sigma}$$

$$P(2.12 < X < 4.48) = P\left(\frac{2.12 - 3.3}{0.6} < Z < \frac{4.48 - 3.3}{0.6}\right)$$
$$= P(-1.97 < Z < 1.97)$$



Loi normale





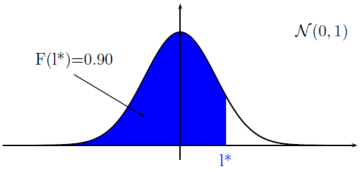
$$
\begin{aligned}
P(2.12 < X < 4.48) &= P(-1.97 < Z < 1.97) \\
&= F(1.97) - F(-1.97) \\
&= F(1.97) - [1 - F(1.97)] \\
&= 2 * F(1.97) - 1 \\
&= 2 * 0.97558 - 1 \\
&\approx 0.95
\end{aligned}
$$

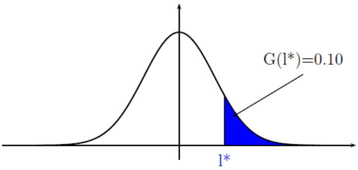
Loi normale

l : limite du poids correspondant aux 10% des bébés les plus gros.

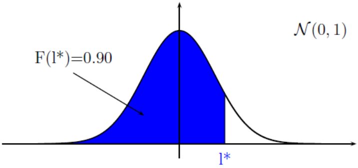
$$P(X \geq l) = 0.10$$
$$P(X < l) = 1 - 0.10 = 0.90$$

$$P\left(\underbrace{\frac{X-3.3}{0.6}}_{Z} < \underbrace{\frac{l-3.3}{0.6}}_{l^*}\right) = 0.90$$





Loi normale



La valeur de l^* telle que $P(Z < l^*) = 0.90$ est l*=1.2816

$$l = 0.6 * 1.2816 + 3.3$$
$$l \approx 4.1\text{kg}$$

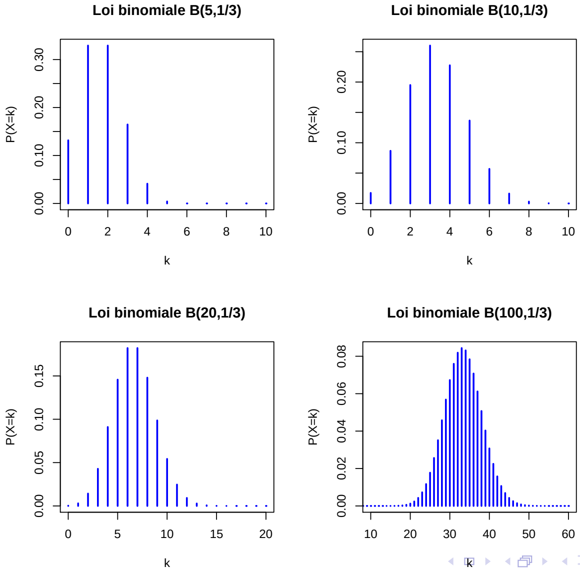
Théorèmes de convergence

Rappel : Approximation de la loi binomiale par la loi de Poisson correcte quand $n \geq 30$ et $np \leq 10$

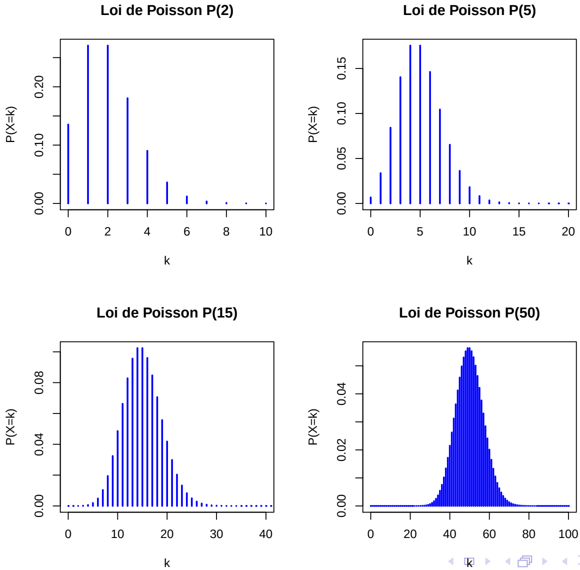
La loi binomiale et la loi de Poisson peuvent être approchées par une loi normale quand :

- $np \geq 10$ et $n(1-p) \geq 10$ pour la loi binomiale ; la loi normale correspondante est $\mathcal{N}(np, \sqrt{npq})$
- $np = \lambda \geq 10$ pour la loi de Poisson ; la loi normale correspondante est $\mathcal{N}(\lambda, \sqrt{\lambda})$

Théorèmes de convergence



Théorèmes de convergence



Théorèmes de convergence

Théorème central limite

Contexte : Epreuves répétées caractérisées par une suite de $X_1, X_2, \ldots, X_n$ va indépendantes et de même loi
$E(X_1) = \mu$ et $Var(X_1) = \sigma^2$

On définit $S_n = \sum_{i=1}^{n} X_i$ et Z_n la va centrée réduite

$$Z_n = \frac{S_n - \mu}{\sigma\sqrt{n}}$$

Théorème

$\forall x$, la fonction de répartition $Fn(x) = P(Z_n \leq x)$ est telle que

$$\lim_{n\to\infty} Fn(x) = \Phi$$

avec Φ fonction de répartition de $\mathcal{N}(0, 1)$

Théorème central limite

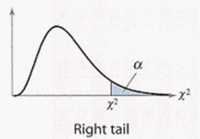
Autrement dit, une variable aléatoire résultant de la somme de plusieurs variables aléatoires ayant même loi et paramètres est distribuée selon une loi normale centrée réduite lorsque le nombre d'épreuves tend vers l'infini.

Le TCL s'applique quelque soit la loi de probabilité suivie par les va discrètes ou continues, pourvu que les épreuves soient indépendantes, reproductibles et en très grand nombre

Lois du χ^2 de Pearson

Loi du χ^2 de Pearson

Soient $X_1, X_2, \ldots, X_n$ n variables normales centrées réduites. On appelle χ^2 la variable $\sum_{i=1}^{n} X_i^2$. On dit que χ^2 suit une loi de Pearson à n degrés de liberté.



Degrees of freedom	α									
	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	—	—	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
30	13.787	14.954	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169

Remarque : si n=1, la variable du χ^2 correspond au carré d'une va $\mathcal{N}(0,1)$

Lois de student

Loi de student

Soit U une va suivant une loi normale centrée réduite et V suivant une loi du χ^2 à n degrés de liberté, U et V étant indépendantes.

On dit alors que

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

suit une loi de Student à n degrés de liberté.