

## 2 Variables aléatoires

- Définition et exemples
- Distribution de probabilité
- Opérations, égalité presque sûre et égalité en distribution
- Espérance mathématique
- Variance, risque, inégalité de Tchebychev
- Quelques distributions discrètes classiques
- Quelques distributions continues classiques
- Moments, fonction génératrice des moments

## 2 Variables aléatoires

### ■ Définition et exemples

■ Distribution de probabilité

■ Opérations, égalité presque sûre et égalité en distribution

■ Espérance mathématique

■ Variance, risque, inégalité de Tchebychev

■ Quelques distributions discrètes classiques

■ Quelques distributions continues classiques

■ Moments, fonction génératrice des moments

## Définition et exemples

Soit une expérience aléatoire  $E$ . Soit un espace probabilisé  $(\Omega, \mathcal{A}, P)$  associé.

Il est commun que le résultat de  $E$  soit un nombre ou qu'on puisse naturellement associer un nombre au résultat. Ceci conduit au concept de **variable aléatoire**.

### Définition

Une **variable aléatoire (v.a.)** est une fonction

$$\begin{aligned} X : \Omega &\rightarrow \mathbb{R} \\ \omega &\mapsto X(\omega) \end{aligned}$$

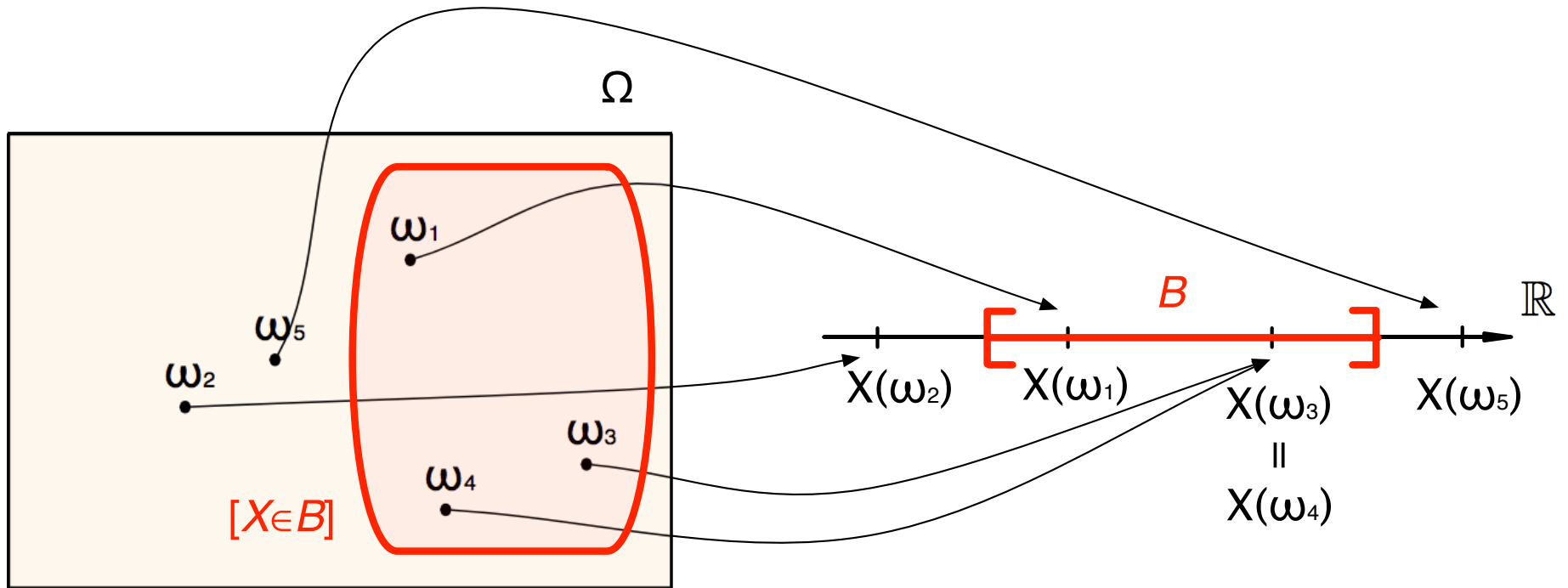
*vérifiant la condition technique que, pour tout borélien  $B \in \mathcal{B}$ ,*

$$[X \in B] \in \mathcal{A} \quad (*)$$

*où  $[X \in B] := X^{-1}(B) := \{\omega \in \Omega : X(\omega) \in B\}$ . La v.a. est dite **discrète** si  $X(\Omega)$ , l'ensemble de ses valeurs possibles  $\{X(\omega) : \omega \in \Omega\}$ , est fini ou infini dénombrable.*

A ce stade,  $P$  ne joue aucun rôle.

# Définition et exemples



# Définition et exemples

Exemple :

$E$  = lancer de deux dés (distinguables)

$\Omega = \{(1, 1), (1, 2), \dots, (1, 6), (2, 1), \dots, (6, 6)\} (\rightsquigarrow \mathcal{A} = \mathcal{P}(\Omega))$

$X$  = la somme des résultats de chaque dé

Autrement dit,

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

$$(i, j) \mapsto X(i, j) = i + j.$$

- Puisque  $\mathcal{A} = \mathcal{P}(\Omega)$ , (\*) est automatiquement satisfaite. Donc  $X$  est une v.a.
- Puisque  $X(\Omega) = \{2, 3, \dots, 12\}$  est fini,  $X$  est une v.a. discrète.

---

Si  $\Omega$  est fini ou infini dénombrable, on peut toujours prendre  $\mathcal{A} = \mathcal{P}(\Omega)$ .

Comme ci-dessus, toute fonction  $X : \Omega \rightarrow \mathbb{R}$  vérifie alors (\*) et est donc une v.a.

# Définition et exemples

Si  $\Omega$  est fini ou infini dénombrable, toutes les v.a. sont discrètes.

Si  $\Omega$  est infini non dénombrable, on peut avoir des v.a. discrètes ou non discrètes.



$E$  = lancer une fléchette.  $\Omega$  = l'ensemble de tous les points d'impact possibles (infini non dénombrable).  $\mathcal{A} = \mathcal{B}^2 := \sigma(\{(a_1, b_1] \times (a_2, b_2] : a_1 < b_1, a_2 < b_2\})$ .

La distance  $Y$  du point d'impact au centre de la cible est une v.a. non discrète.

Le nombre  $X$  de points marqués est une v.a. discrète

$E$  = lancer une pièce jusqu'à obtenir pile pour la première fois

$\Omega = \{P, (F, P), (F, F, P), (F, F, F, P), \dots\}$  ( $\rightsquigarrow \mathcal{A} = \mathcal{P}(\Omega)$ )

$X$  = le nombre de lancers nécessaires

$X$  est une v.a. discrète (avec une infinité de valeurs possibles)

$E$  = mesurer la durée de vie d'une ampoule (en heures)

$\Omega = (0, \infty)$

$\mathcal{A} = \{B \cap (0, \infty) : B \in \mathcal{B}\}$

$X$  = la durée de vie elle-même

$X$  est une v.a. non discrète

$E$  = mesurer (en heures) la durée d'une enchère eBay de 7 jours avec l'option "achat immédiat"

$\Omega = (0, 7 \times 24] = (0, 168]$

$\mathcal{A} = \{B \cap (0, 168] : B \in \mathcal{B}\}$

$X$  = la durée elle-même

$X$  est une v.a. non discrète

## 2 Variables aléatoires

- Définition et exemples
- Distribution de probabilité
- Opérations, égalité presque sûre et égalité en distribution
- Espérance mathématique
- Variance, risque, inégalité de Tchebychev
- Quelques distributions discrètes classiques
- Quelques distributions continues classiques
- Moments, fonction génératrice des moments



## Distribution : cas général

Considérons deux jeux pour lesquels il faut payer 10€ pour jouer.

- Dans le 1er, on gagne 30€ si on obtient un as en tirant une carte dans un jeu de 52 cartes, et 0€ sinon.
- Dans le 2nd, on gagne 30€ si on obtient face en lançant une pièce de monnaie, et 0€ sinon.

Les v.a.  $X_1$  et  $X_2$  correspondant au gain net dans ces jeux ont le même ensemble de valeurs possibles ( $X_1(\Omega) = X_2(\Omega) = \{-10, 20\}$ ), mais sont pourtant très différentes.

---

Ce qui caractérise complètement une v.a.  $X$  est **la distribution (de probabilité) de  $X$** , qui est la collection de toutes les probabilités de la forme  $P[X \in B]$ —ou, de manière équivalente, la loi de probabilité  $P^X$  sur  $(\mathbb{R}, \mathcal{B})$  définie par  $P^X[B] = P[X \in B]$ .

(Le rôle de  $(*)$  est d'assurer qu'on puisse considérer  $P[X \in B] \forall B \in \mathcal{B}$ )

---

Comment décrire efficacement la distribution de  $X$  ?

Si  $X$  est une v.a. discrète, admettant les valeurs possibles  $x_i, i \in \mathcal{I}$ , on a que

$$P[X \in B] = \sum_{i: x_i \in B} P[X = x_i] \quad \forall B \in \mathcal{B}, \quad (1)$$

où  $[X = x_i] = X^{-1}(\{x_i\})$ . Il en découle que la distribution de  $X$  est complètement caractérisée par la collection  $\{(x_i, p_i := P[X = x_i]) : i \in \mathcal{I}\}$  des valeurs possibles et des probabilités correspondantes.

Distribution de $X$				
valeurs possibles	$x_1$	$x_2$	...	$(x_k)$
probabilités	$p_1$	$p_2$	...	$(p_k)$

Par un abus de langage, on appelle aussi ce tableau "distribution de  $X$ " (ce qui se justifie par le fait qu'il permet de calculer, via (1), la distribution au sens strict)

# Distribution : cas discret

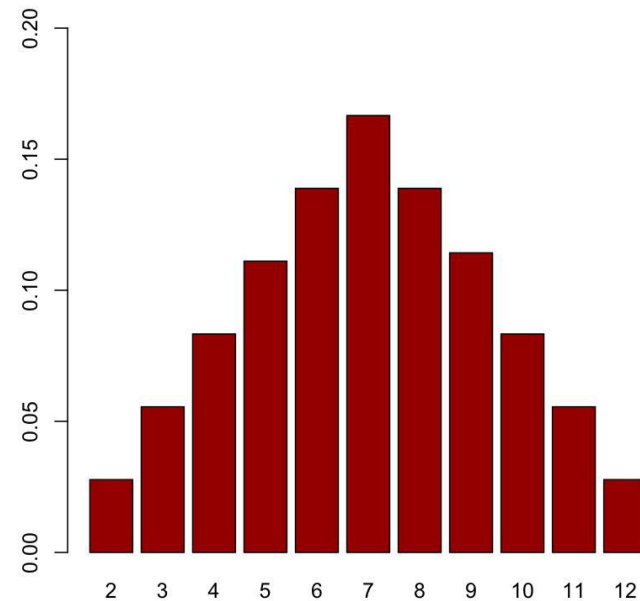
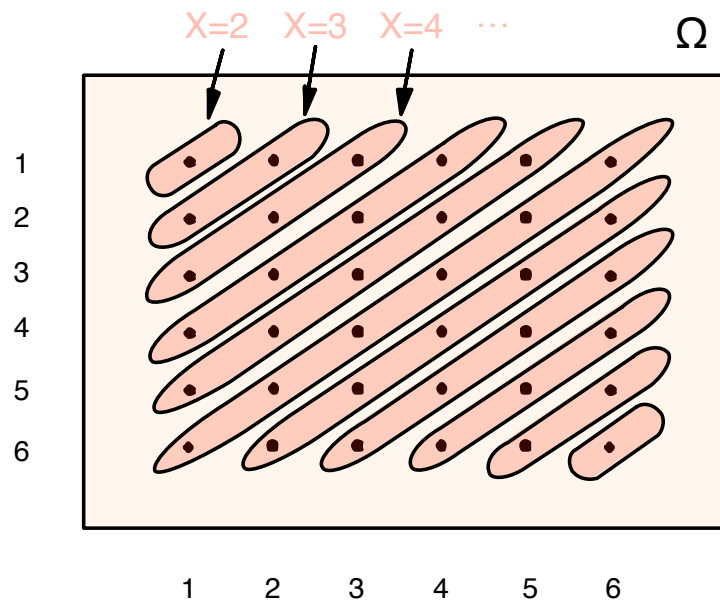
$E$  = lancer de deux dés (distinguables)

$\Omega = \{(1, 1), (1, 2), \dots, (1, 6), (2, 1), \dots, (6, 6)\}$  ( $\rightsquigarrow \mathcal{A} = \mathcal{P}(\Omega)$ )

$P$  : équiprobabilité

$X$  = la somme des résultats de chaque dé

Distribution de $X$											
valeurs possibles	2	3	4	5	6	7	8	9	10	11	12
probabilités	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$



## Distribution : cas discret

Comme annoncé, ceci définit une loi de probabilité  $P^X$  sur  $(\mathbb{R}, \mathcal{B})$ .

Par exemple, on a

$$P^X[\{3\}] = P[X = 3] = \frac{2}{36}$$

$$P^X[\{3, 4\}] = P[X \in \{3, 4\}] = \frac{2}{36} + \frac{3}{36} = \frac{5}{36}$$

$$P^X[[3, 4]] = P[X \in [3, 4]] = P[X \in \{3, 4\}] = \frac{2}{36} + \frac{3}{36} = \frac{5}{36}$$

$$P^X[[3.5, 5 + \sqrt{2})] = P[X \in [3.5, 5 + \sqrt{2})] = P[X \in \{4, 5, 6\}] = \frac{3}{36} + \frac{4}{36} + \frac{5}{36} = \frac{12}{36}$$

⋮

# Distribution : cas discret

Exemple avec une infinité de valeurs possibles :

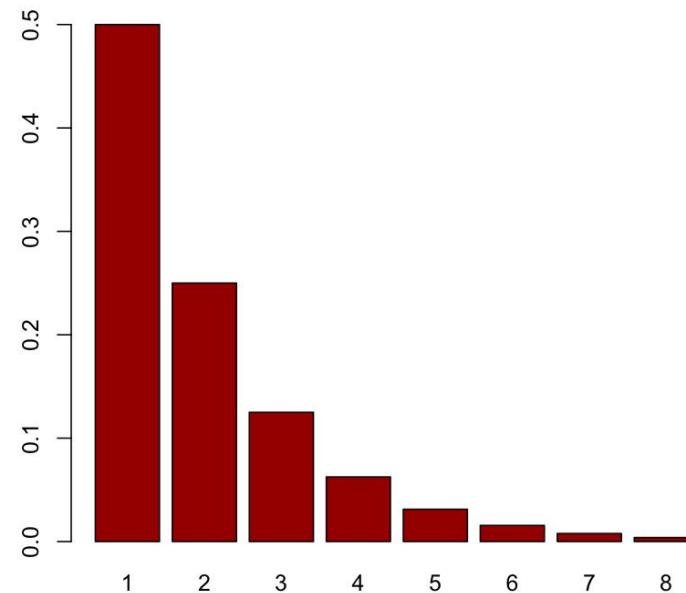
$E$  = on lance une pièce jusqu'à obtenir pile pour la première fois

$\Omega = \{P, (F, P), (F, F, P), (F, F, F, P), \dots\}$  ( $\rightsquigarrow \mathcal{A} = \mathcal{P}(\Omega)$ )

$P$  donnée par  $(p_1, p_2, p_3, p_4, \dots) = (\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \dots)$

$X$  = le nombre de lancers nécessaires

Distribution de $X$					
valeurs possibles	1	2	3	4	...
probabilités	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	...



## Distribution : \* retour au cas général \*

Soit  $X$  une v.a. (pas nécessairement discrète).

### Définition

La *fonction de répartition* de  $X$  est la fonction

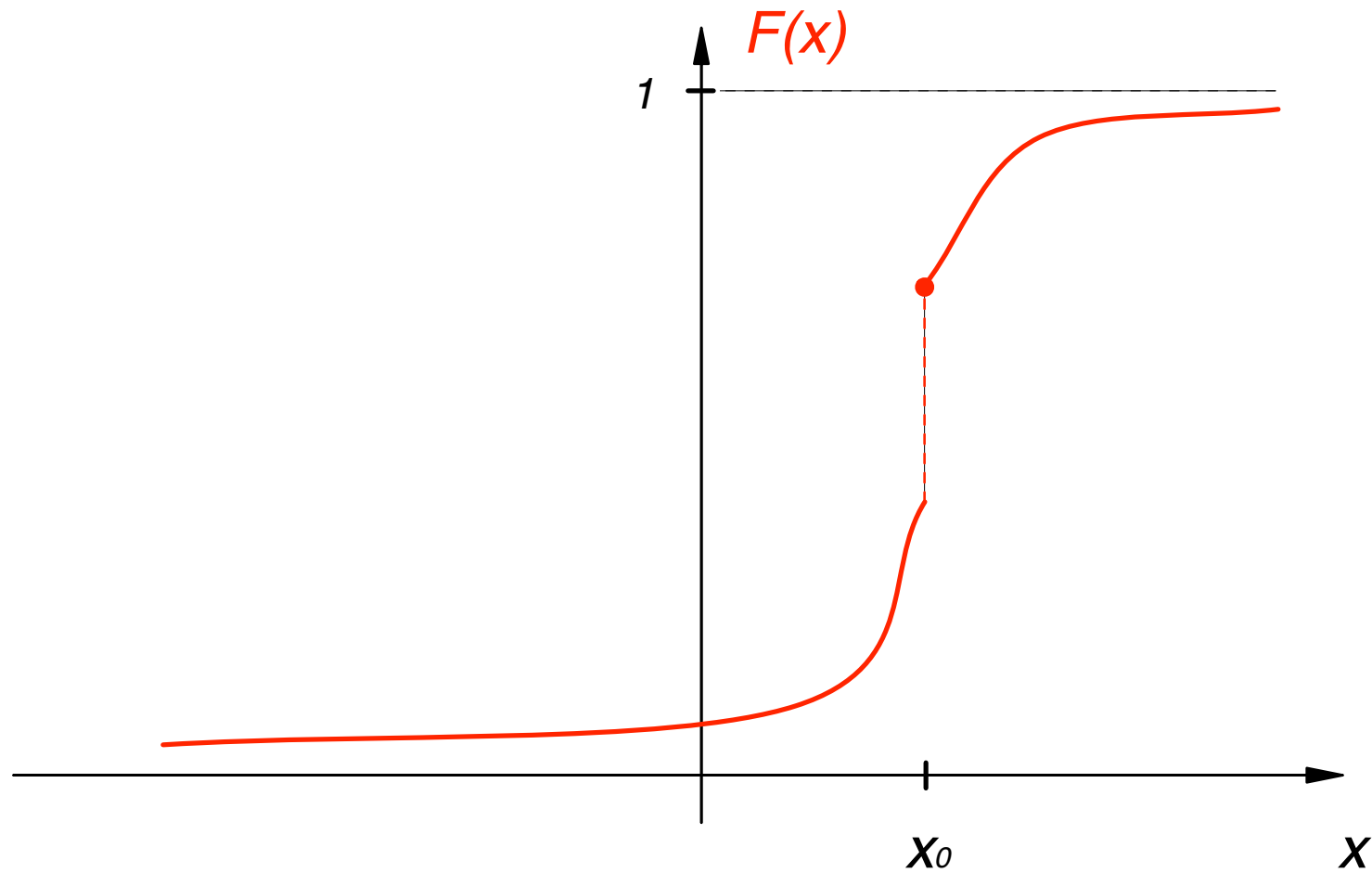
$$\begin{aligned} F : \mathbb{R} &\rightarrow [0, 1] \\ x &\mapsto P[X \leq x]. \end{aligned}$$

Propriétés caractéristiques :

- $\lim_{x \rightarrow -\infty} F(x) = 0$  et  $\lim_{x \rightarrow \infty} F(x) = 1$
- $F$  est non décroissante ( $x < x' \Rightarrow F(x) \leq F(x')$ )
- En notant  $F(x_0^+) := \lim_{x \searrow x_0} F(x)$ , on a  $F(x_0^+) = F(x_0)$  (continuité à droite)

Toute fonction  $F$  satisfaisant ces propriétés est la fonction de répartition d'une v.a.

# Distribution : \* retour au cas général \*



Fonction de répartition d'une v.a.  $X$

Propriétés supplémentaires :

- En notant  $F(x_0^-) := \lim_{x \nearrow x_0} F(x)$ , on a

$$F(x_0) - F(x_0^-) = P[X = x_0].$$

Une fonction de répartition n'est donc pas toujours continue à gauche (et n'est donc pas toujours continue).

- Pour tout  $a < b$ , on a

$$F(b) - F(a) = P[a < X \leq b]. \quad (2)$$

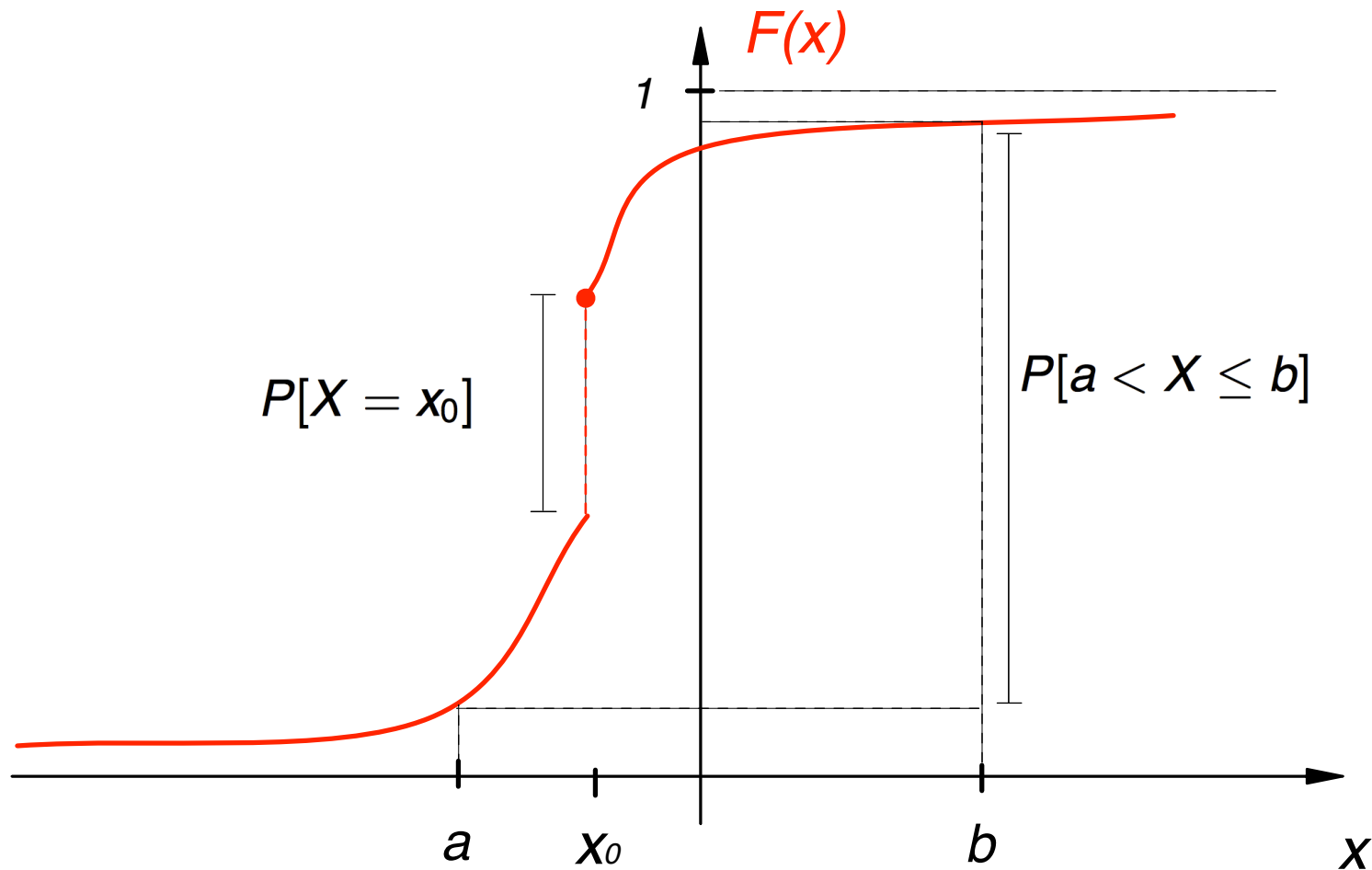
En prenant la limite pour  $b \rightarrow \infty$ , on obtient  $1 - F(a) = P[X > a]$ .

---

Il découle de (2) que  $F$  caractérise complètement la distribution de  $X$  ( $P^X \Leftrightarrow F$ ).



# Distribution : \* retour au cas général \*



Fonction de répartition d'une v.a.  $X$  et lecture des probabilités

Un exemple :

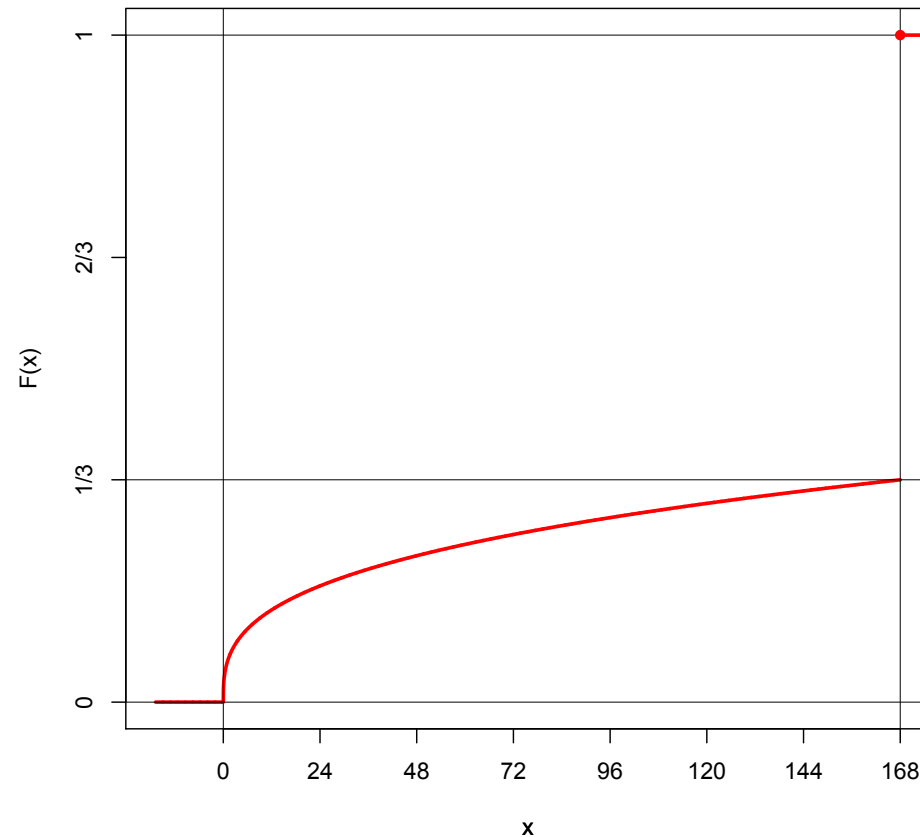
$E$  = mesurer (en heures) la durée d'une enchère eBay de 7 jours avec l'option "achat immédiat"

$\Omega = (0, 7 \times 24] = (0, 168]$

$\mathcal{A} = \{B \cap (0, 168] : B \in \mathcal{B}\}$

$X$  = la durée elle-même

$X$  est une v.a. non discrète.



Fonction de répartition ( $F$ ) de la durée de l'enchère (en heures).  $P[\text{achat immédiat}] = P[X < 168] = F(168^-) = 1/3$  et  $P[\text{l'enchère va à son terme}] = P[X = 168] = F(168) - F(168^-) = 2/3$ . On voit aussi que la probabilité d'avoir un achat immédiat dans l'intervalle  $[x, x + \delta]$  est décroissante en  $x$  (ce qui est raisonnable)

## Distribution : \* retour au cas général \*

La fonction de répartition permet de définir une notion qui sera cruciale dans la partie "inférence statistique" du cours.

### Définition

Soit  $\alpha \in (0, 1)$  fixé. Le **quantile d'ordre  $\alpha$**  de la distribution de  $X$  est le nombre

$$x_\alpha = \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}.$$

S'il existe un unique nombre  $c$  tel que  $F(c) = \alpha$ , alors  $x_\alpha = c$ . Mais la définition ci-dessus permet de définir  $x_\alpha$  même dans les cas où il n'y a pas de tel  $c$ .

Terminologie :

- $x_{\frac{1}{2}}$  est la médiane
- $x_{\frac{1}{4}}$  et  $x_{\frac{3}{4}}$  sont les 1er et 3ème quartiles
- $x_{\frac{i}{10}}$ ,  $i = 1, 2, \dots, 9$  sont les déciles
- $x_{\frac{i}{100}}$ ,  $i = 1, 2, \dots, 99$  sont les percentiles

# Distribution : \* retour au cas général \*

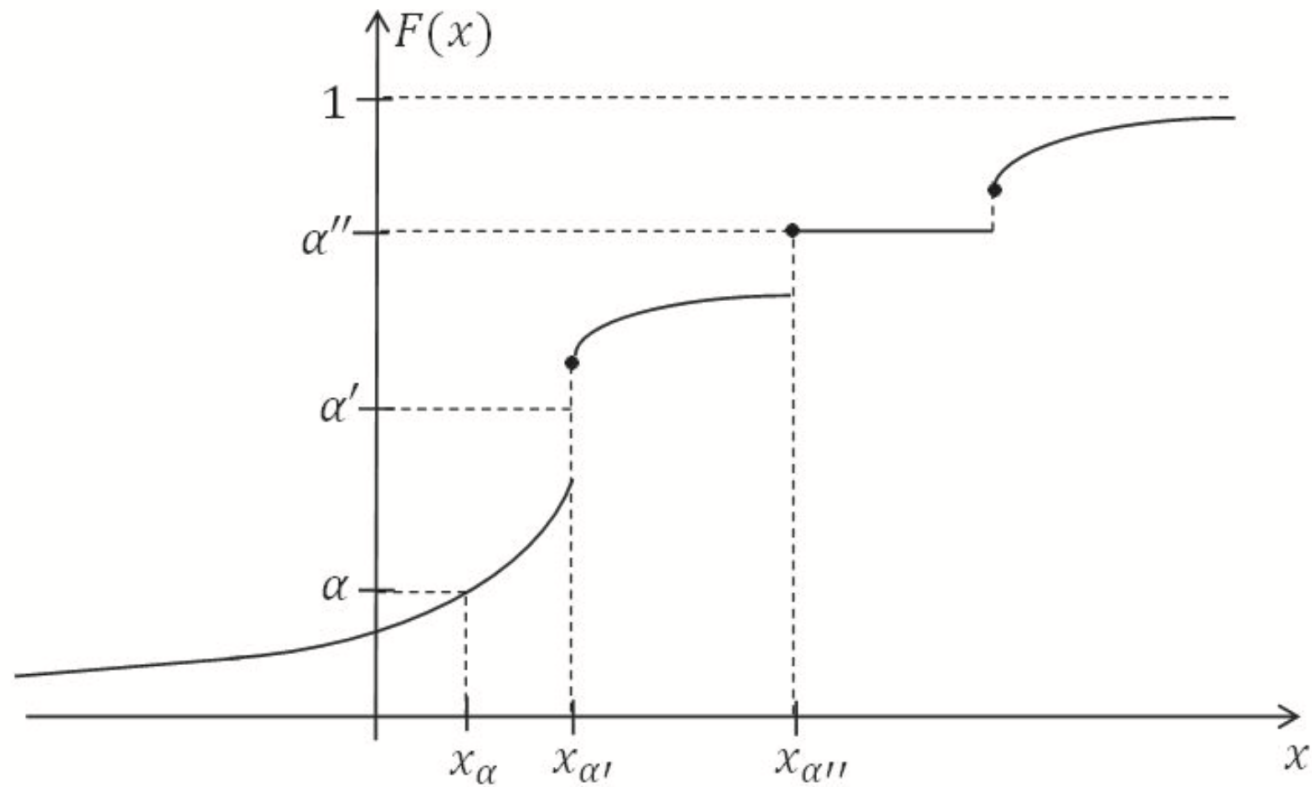
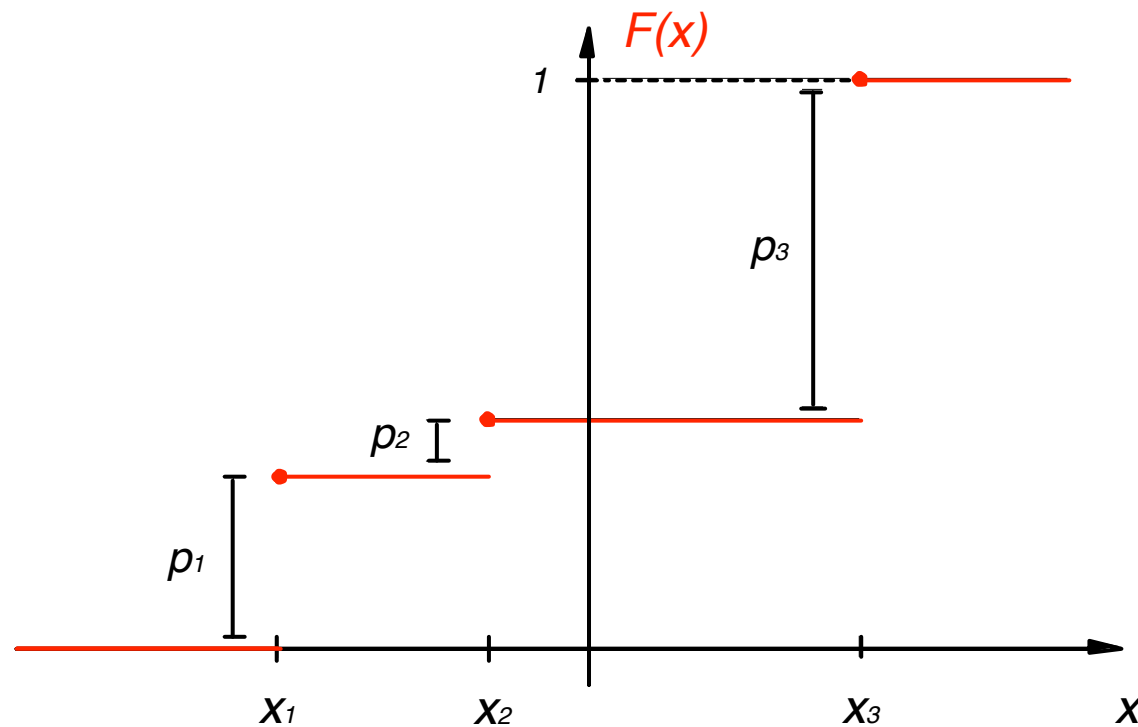


Illustration graphique des quantiles

## Distribution : cas discret

Soit  $X$  une v.a. discrète, de distribution  $(x_i, p_i = P[X = x_i])$ ,  $i \in \mathcal{I}$ , sa distribution.

Alors on a  $F(x) = P[X \leq x] = \sum_{i: x_i \leq x} p_i$ .



Une classe importante de v.a. non discrètes présente des fonctions  $F$  continues.

Pour de telles v.a., on aura donc

$$P[X = x] = F(x) - F(x^-) = 0 \quad \forall x \in \mathbb{R}.$$

Ces v.a. sont d'une nature fondamentalement différente de celle des v.a. discrètes.

Exemples :

- la durée de vie d'une ampoule
- le temps d'attente du 71
- la quantité de pluie en août prochain à Bruxelles
- ...

Puisque de telles v.a. satisfont  $P[X = x] = 0 \forall x$ , on aura alors

$$\begin{aligned}1 - F(a) &= P[X > a] \\ &= P[X \geq a]\end{aligned}$$

et

$$\begin{aligned}F(b) - F(a) &= P[a < X \leq b] \\ &= P[a \leq X \leq b] \\ &= P[a \leq X < b] \\ &= P[a < X < b].\end{aligned}$$

Ceci montre comment déduire ces probabilités de la fonction de répartition.



## Distribution : cas continu

$E$  = mesurer la durée de vie d'une ampoule (en heures)

$\Omega = (0, \infty)$

$\mathcal{A} = \{B \cap (0, \infty) : B \in \mathcal{B}\}$

$X$  = la durée de vie elle-même

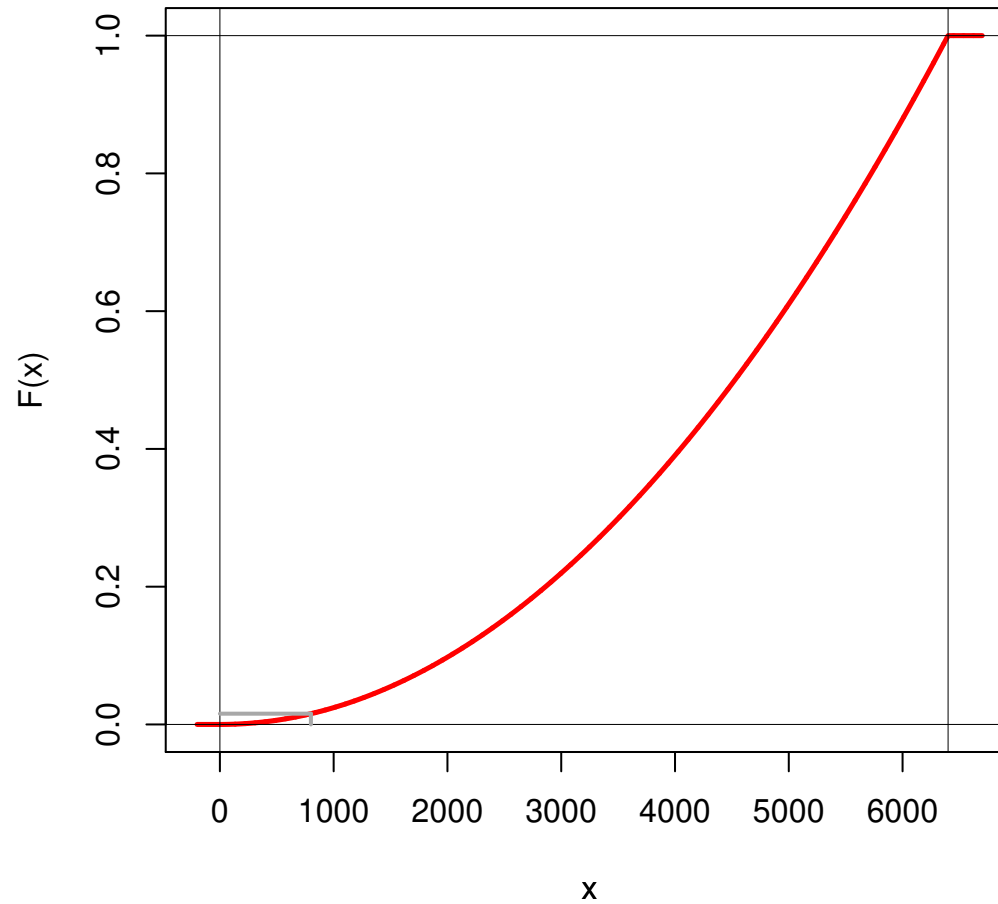
Supposons que la fonction de répartition de  $X$  soit donnée par

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{x^2}{6400^2} & \text{si } 0 \leq x < 6400 \\ 1 & \text{si } x \geq 6400. \end{cases}$$

Si le fabricant s'engage à rembourser l'ampoule dans le cas où celle-ci ne survit pas 800 heures au moins, la probabilité que le fabricant doive faire ce geste vaut

$$P[X < 800] = P[X \leq 800] = F(800) = \frac{800^2}{6400^2} = \frac{1}{64} \approx 0.016.$$

# Distribution : cas continu



Fonction de répartition  $F$  de la durée de vie de l'ampoule

La v.a. modélisant la durée de vie d'une ampoule ci-dessus est une v.a. **continue**.

### Définition

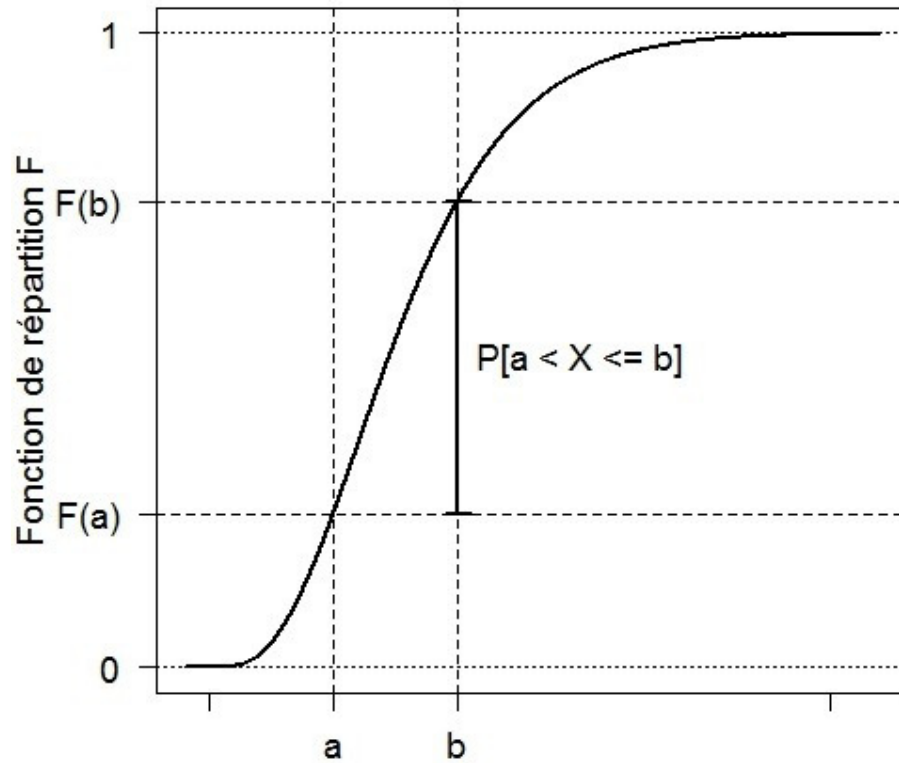
La v.a.  $X$  est **continue**  $\Leftrightarrow$  il existe une fonction  $f : \mathbb{R} \rightarrow \mathbb{R}^+$  telle que

$$P[a < X \leq b] = \int_a^b f(x) dx \quad (= F(b) - F(a)),$$

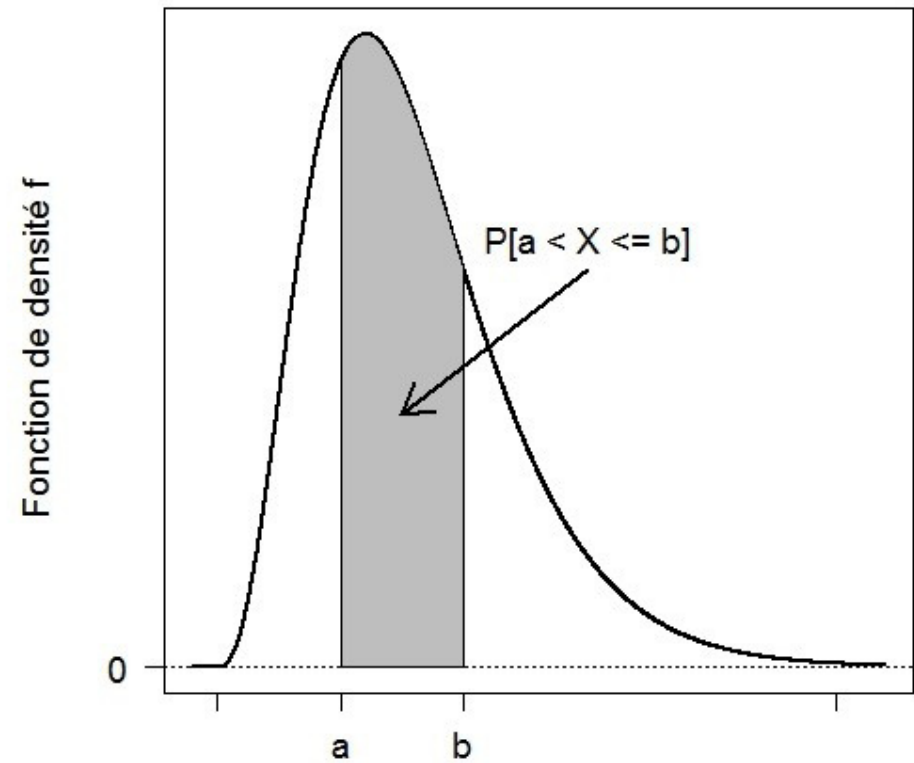
pour tout  $a < b$ . La fonction  $f$  est appelée **fonction de densité (de probabilité)** de  $X$ .

Si  $X$  est une v.a. continue, alors  $P[a < X \leq b]$  correspond à **l'aire sous le graphe de  $f$  au dessus de l'intervalle  $[a, b]$** .

# Distribution : cas continu



Fonction de répartition ( $F$ )



Fonction de densité ( $f$ )

En prenant  $a \rightarrow -\infty$  dans la définition ci-dessus, on obtient

$$F(x) = \int_{-\infty}^x f(z) dz,$$

ce qui montre qu'on peut calculer  $F$  à partir de  $f$ .

Si  $f$  est continue<sup>\*</sup>, le théorème fondamental du calcul différentiel et intégral fournit

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

et on peut donc aussi calculer  $f$  à partir de  $F$ .

~>  $f$  fournit donc une autre caractérisation de la distribution de  $X$  ( $P^X \Leftrightarrow F \Leftrightarrow f$ )!

---

\*. Si  $f$  n'est pas continue, on peut montrer que  $f(x) = F'(x)$  presque partout

Propriétés caractéristiques de  $f$  :

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

Toute fonction  $f$  satisfaisant ces deux propriétés est la fonction de densité d'une v.a. continue.

---

Interprétation de  $f(x)$  :

Pour  $\Delta x$  petit, on a

$$P[X \in (x, x + \Delta x]] = F(x + \Delta x) - F(x) \approx f(x) \Delta x,$$

de sorte que  $f(x)$  est proportionnelle à la probabilité que  $X$  se réalise dans un (petit) voisinage fixé de  $x$  ; au plus grande la valeur de  $f(x)$ , au plus il est probable que  $X$  se réalisera "autour de  $x$ ".

## Distribution : cas continu

$E$  = mesurer la durée de vie d'une ampoule (en heures)

$\Omega = (0, \infty)$

$\mathcal{A} = \{B \cap (0, \infty) : B \in \mathcal{B}\}$

$X$  = la durée de vie elle-même

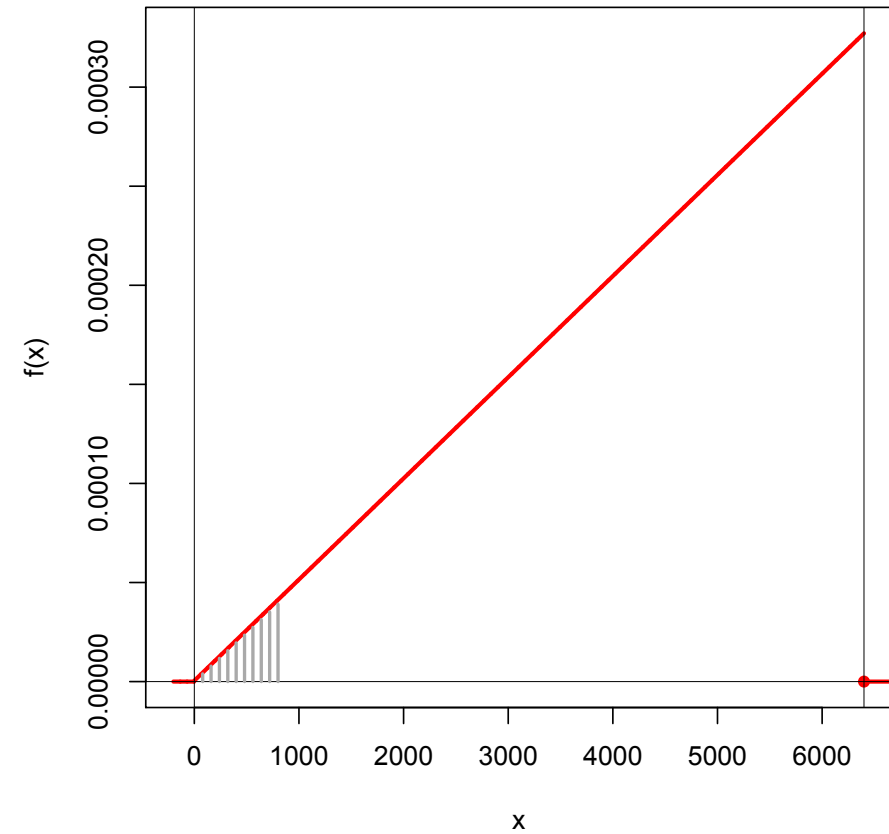
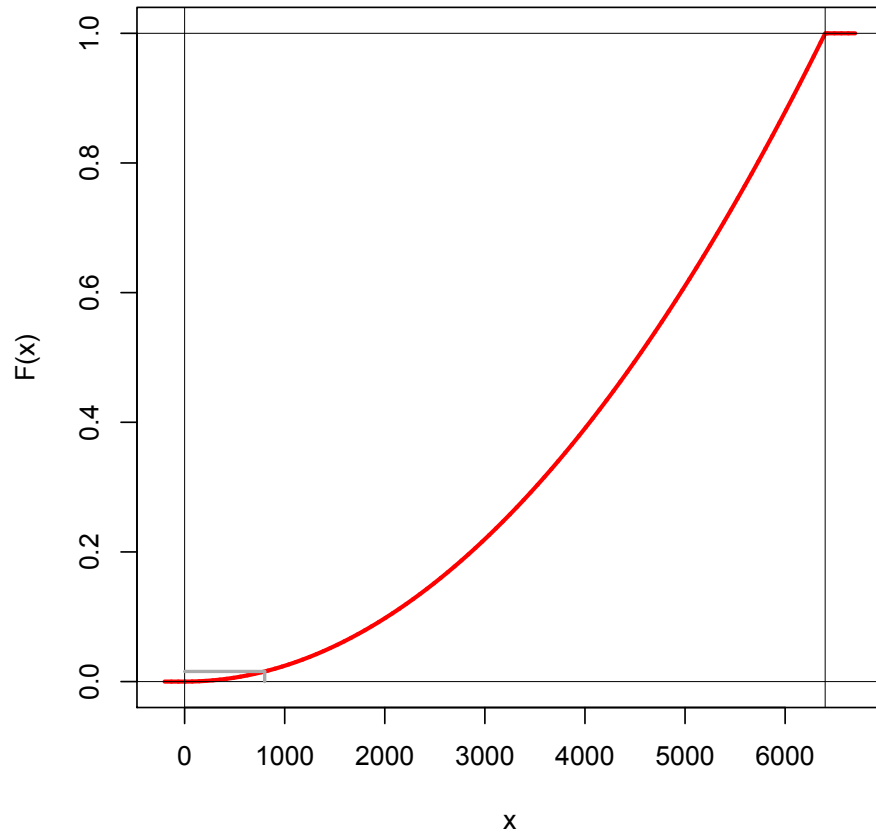
Si la fonction de répartition de  $X$  est donnée par

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{x^2}{6400^2} & \text{si } 0 \leq x < 6400 \\ 1 & \text{si } x \geq 6400, \end{cases}$$

on obtient la fonction de densité

$$f(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{2x}{6400^2} & \text{si } 0 \leq x < 6400 \\ 0 & \text{si } x \geq 6400. \end{cases}$$

# Distribution : cas continu



Fonction de répartition ( $F$ ) et fonction de densité ( $f$ ) de la durée de vie de l'ampoule



Sur base de la densité  $f$ , on peut recalculer que la probabilité que l'ampoule cède en moins de 800 heures vaut

$$P[X < 800] = \int_{-\infty}^{800} f(x) dx = \int_0^{800} \frac{2x}{6400^2} dx = \left[ \frac{x^2}{6400^2} \right]_0^{800} = \frac{1}{64}.$$

---

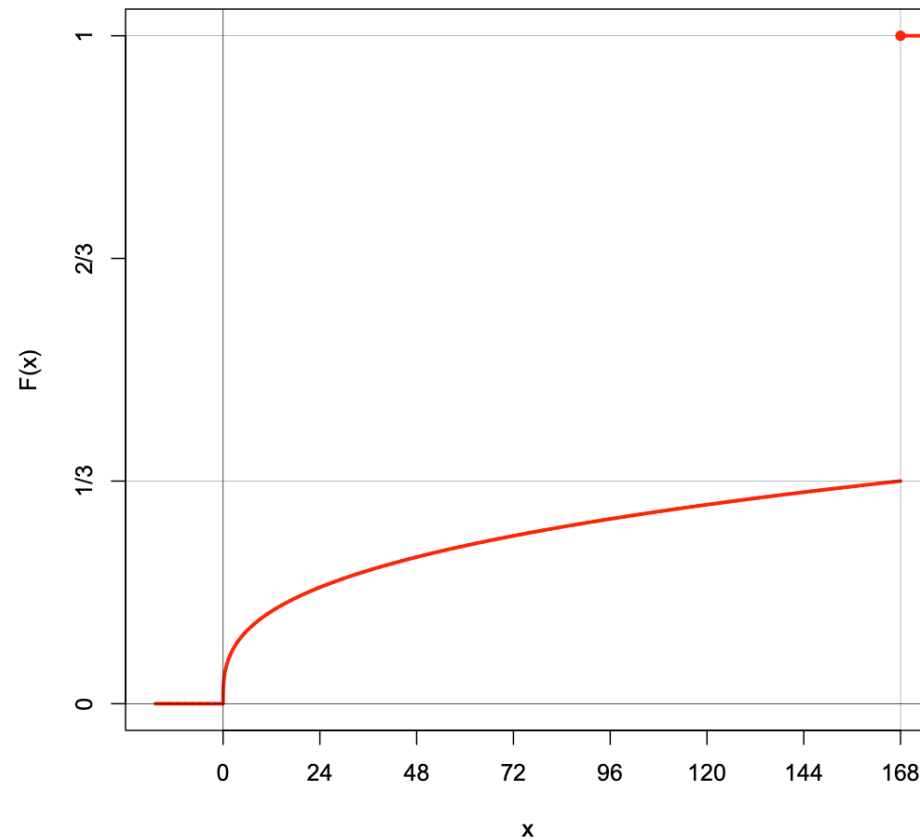
Clairement, 800 est l'unique valeur  $c$  telle que  $F(c) = \frac{1}{64}$ .

↪ 800 est le quantile d'ordre  $\alpha = \frac{1}{64}$ .

Graphiquement (voir la page précédente),

- $x = 800$  est l'abscisse pour laquelle l'ordonnée sur le graphe de  $F$  vaut  $\frac{1}{64}$ .
- L'aire sous le graphe de  $f$  à gauche de  $x = 800$  vaut  $\frac{1}{64}$ .

Il existe des v.a. qui ne sont ni discrètes ni continues...



Fonction de répartition ( $F$ ) de la durée de l'enchère eBay (en heures)

## 2 Variables aléatoires

- Définition et exemples
- Distribution de probabilité
- Opérations, égalité presque sûre et égalité en distribution
- Espérance mathématique
- Variance, risque, inégalité de Tchebychev
- Quelques distributions discrètes classiques
- Quelques distributions continues classiques
- Moments, fonction génératrice des moments

# Opérations sur les variables aléatoires

Soient  $X_1, X_2$  des v.a. définies sur un même espace  $(\Omega, \mathcal{A}, P)$ .

---

Puisque des v.a. sont des fonctions, on peut les combiner.

## Définition

La v.a.  $X_1 + X_2$  est définie par la relation

$$X_1 + X_2 : \Omega \mapsto \mathbb{R}$$

$$\omega \mapsto (X_1 + X_2)(\omega) = X_1(\omega) + X_2(\omega).$$

On définit de façon analogue les v.a.  $X_1 X_2$  et (si  $X_2(\omega) \neq 0$  pour tout  $\omega$  :)  $X_1 / X_2$ .

Enfin, si le domaine de  $g : \mathbb{C} \rightarrow \mathbb{R}$  contient l'ensemble des valeurs possibles de  $X$ , on peut considérer la v.a.  $g(X)$ , définie à travers  $(g(X))(\omega) = g(X(\omega))$  pour tout  $\omega$ .

Les définitions suivantes procèdent également comme pour les fonctions.

## Définition

- (i)  $X_1 = X_2 \Leftrightarrow$  pour tout  $\omega \in \Omega$ , on a  $X_1(\omega) = X_2(\omega)$
- (ii)  $X_1 \leq X_2 \Leftrightarrow$  pour tout  $\omega \in \Omega$ , on a  $X_1(\omega) \leq X_2(\omega)$
- (iii)  $X_1 \geq X_2 \Leftrightarrow$  pour tout  $\omega \in \Omega$ , on a  $X_1(\omega) \geq X_2(\omega)$

Mais les concepts plus faibles suivants sont plus utiles en probabilité.

## Définition

- (i)  $X_1 = X_2$  p.s. ("*presque sûrement*")  $\Leftrightarrow P[X_1 = X_2] = 1$
- (ii)  $X_1 \leq X_2$  p.s.  $\Leftrightarrow P[X_1 \leq X_2] = 1$
- (iii)  $X_1 \geq X_2$  p.s.  $\Leftrightarrow P[X_1 \geq X_2] = 1$

Ici, il faut lire  $P[X_1 = X_2]$  comme  $P[\{\omega \in \Omega : X_1(\omega) = X_2(\omega)\}]$ , etc.

# Egalité en distribution

Soient  $X_1$  une v.a. définie sur un espace probabilisé  $(\Omega_1, \mathcal{A}_1, P_1)$  et  $X_2$  une v.a. définie sur un espace probabilisé  $(\Omega_2, \mathcal{A}_2, P_2)$ .

## Définition

$X_1$  et  $X_2$  sont égales en distribution (notation :  $X_1 \stackrel{\mathcal{D}}{=} X_2$ )  
 $\Leftrightarrow$  Pour tout  $B \in \mathcal{B}$ , on a  $P_1[X_1 \in B] = P_2[X_2 \in B]$ .

Deux v.a. sont donc égales en distribution... si elles partagent la même distribution.

---

Remarque :  $X_1 = X_2$  p.s.  $\Rightarrow X_1 \stackrel{\mathcal{D}}{=} X_2$ , mais pas l'inverse. Par exemple, lors du lancer d'une pièce de monnaie équilibrée, les v.a.  $X_1$  et  $X_2$  définies par

$$X_1(\omega) = \begin{cases} 1 & \text{si } \omega = P \\ 0 & \text{si } \omega = F \end{cases} \quad \text{et} \quad X_2(\omega) = 1 - X_1(\omega) = \begin{cases} 0 & \text{si } \omega = P \\ 1 & \text{si } \omega = F \end{cases}$$

vérifient  $X_1 \stackrel{\mathcal{D}}{=} X_2$ , mais pas  $X_1 = X_2$  p.s. (exercice).

## 2 Variables aléatoires

- Définition et exemples
- Distribution de probabilité
- Opérations, égalité presque sûre et égalité en distribution
- **Espérance mathématique**
- Variance, risque, inégalité de Tchebychev
- Quelques distributions discrètes classiques
- Quelques distributions continues classiques
- Moments, fonction génératrice des moments

On **résume** souvent l'information contenue dans la distribution d'une v.a.  $X$  au moyen d'un petit nombre d'indicateurs (les **moments**), qui requerront le concept suivant.

## Définition

(i) Pour  $X$  **discrète**, de distribution  $(x_i, p_i = P[X = x_i])$ ,  $i \in \mathcal{I}$ , l'espérance mathématique  $E[X]$  de  $X$  existe **si**  $E[|X|] = \sum_{i \in \mathcal{I}} |x_i| p_i$  est bien définie dans  $\mathbb{R}$ .

Dans ce cas,

$$E[X] = \sum_{i \in \mathcal{I}} x_i p_i.$$

(ii) Pour  $X$  **continue**, de fonction de densité  $f$ , l'espérance mathématique  $E[X]$  de  $X$  existe **si**  $E[|X|] = \int_{-\infty}^{\infty} |x| f(x) dx$  est bien définie dans  $\mathbb{R}$  (converge dans  $\mathbb{R}$ ).

Dans ce cas,

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

En (i)–(ii),  $E[X]$  est donc une **moyenne pondérée** des valeurs possibles de  $X$ , où les poids sont déterminés par la vraisemblance de chaque valeur.



# Espérance mathématique

Exemple 1 (cas discret avec  $\#\mathcal{I} < \infty$ ) :

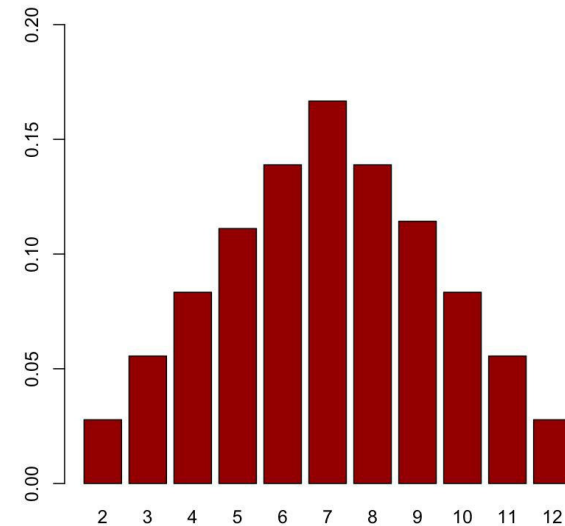
$E$  = lancer de deux dés (distinguables)

$\Omega = \{(1, 1), (1, 2), \dots, (1, 6), (2, 1), \dots, (6, 6)\}$  ( $\rightsquigarrow \mathcal{A} = \mathcal{P}(\Omega)$ )

$P$  : équiprobabilité

Distribution de $X =$ la somme des résultats de chaque dé											
valeurs possibles $x_i$	2	3	4	5	6	7	8	9	10	11	12
probabilités $p_i$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$$\begin{aligned}\rightsquigarrow E[X] &= \sum_{i=1}^{11} x_i p_i = x_1 p_1 + \dots + x_{11} p_{11} \\ &= 2 \times \frac{1}{36} + 3 \times \frac{2}{36} + \dots + 12 \times \frac{1}{36} \\ &= 7\end{aligned}$$



# Espérance mathématique

Exemple 2 (cas discret avec  $\#\mathcal{I} = \infty$ ) :

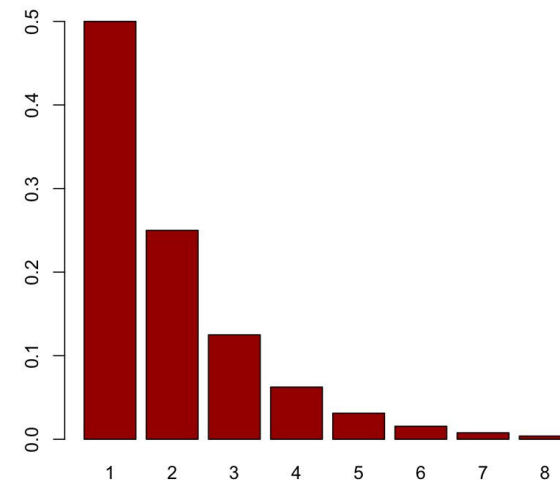
$E$  = lancer une pièce jusqu'à obtenir pile pour la première fois

$\Omega = \{P, (F, P), (F, F, P), (F, F, F, P), \dots\}$  ( $\rightsquigarrow \mathcal{A} = \mathcal{P}(\Omega)$ )

$P$  donné par  $(p_1, p_2, p_3, p_4, \dots) = (\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \dots)$

Distribution de $X$ = le nombre de lancers nécessaires					
valeurs possibles $x_i$	1	2	3	4	...
probabilités $p_i$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	...

$$\begin{aligned}\rightsquigarrow E[X] &= \sum_{i=1}^{\infty} x_i p_i \\ &= 1 \times \frac{1}{2} + 2 \times \frac{1}{4} + 3 \times \frac{1}{8} + 4 \times \frac{1}{16} + \dots \\ &= 2\end{aligned}$$



Exemple 3 (cas continu) :

$E$  = mesurer la durée de vie d'une ampoule (en heures)

$\Omega = (0, \infty)$

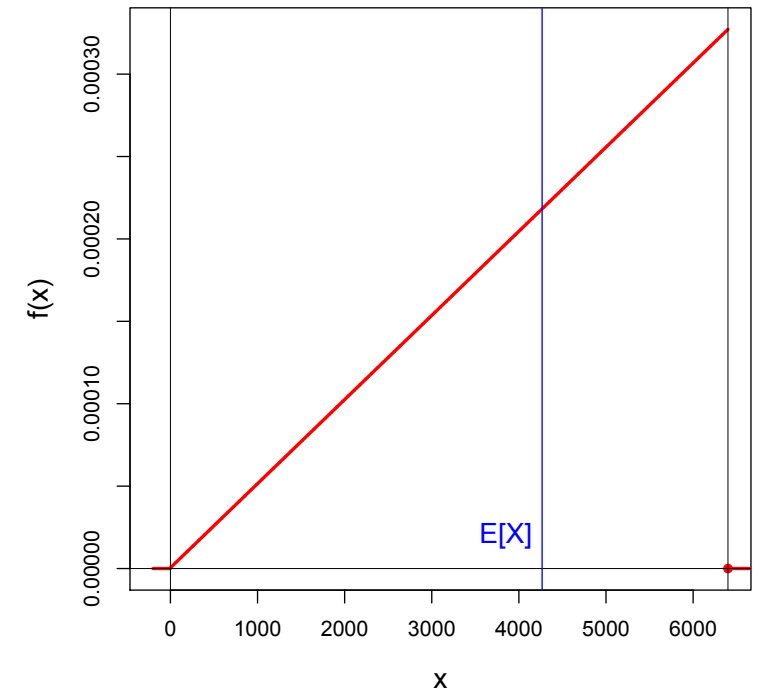
$\mathcal{A} = \{B \cap (0, \infty) : B \in \mathcal{B}\}$

Puisque la fonction de densité de la durée de vie  $X$  est

$$f(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{2x}{6400^2} & \text{si } 0 \leq x < 6400 \\ 0 & \text{si } x \geq 6400, \end{cases}$$

on obtient

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f(x) dx = \int_0^{6400} \frac{2x^2}{6400^2} dx \\ &= \frac{2}{6400^2} \left[ \frac{x^3}{3} \right]_0^{6400} = \frac{2 \times 6400}{3} \approx 4266.67 \dots \end{aligned}$$



La condition qui veut que  $E[|X|]$  est bien définie dans  $\mathbb{R}$  assure que  $E[X]$  est elle-même bien définie et finie (car on a toujours  $|E[X]| \leq E[|X|]$ ).

---

Cette condition n'est pas toujours satisfaite :

Si  $X$  prend les valeurs possibles  $x_i = (-2)^i$  avec probabilité  $p_i = (\frac{1}{2})^i$  ( $i = 1, 2, \dots$ ),

$$E[|X|] = \sum_{i=1}^{\infty} |x_i| p_i = \sum_{i=1}^{\infty} 2^i \times \frac{1}{2^i} = \sum_{i=1}^{\infty} 1 = \infty.$$

Dans ce cas, la série

$$E[X] = \sum_{i=1}^{\infty} x_i p_i = (-1) + 1 + (-1) + 1 + \dots$$

ne converge pas, de sorte que  $E[X]$  n'est pas bien définie.

Dans la suite, on devra considérer l'espérance de certaines **transformées**  $g(X)$  de  $X$ , où le domaine de  $g : C \rightarrow \mathbb{R}$  contient l'ensemble des valeurs possibles de  $X$ .

## Théorème

Soit  $g$  comme ci-dessus et tel que  $E[|g(X)|]$  est bien définie dans  $\mathbb{R}$ . Alors

(i) Pour  $X$  **discrète**, de distribution  $(x_i, p_i = P[X = x_i]), i \in \mathcal{I}$ ,

$$E[g(X)] = \sum_{i \in \mathcal{I}} g(x_i) p_i.$$

(ii) Pour  $X$  **continue**, de fonction de densité  $f$ ,

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

On a donc **deux méthodes** pour calculer  $E[g(X)]$  :

1. Utiliser le théorème ci-dessus
2. Poser  $Y = g(X)$  et calculer  $E[g(X)] = E[Y]$  selon la définition de ch.1-p.35.

Exemple :

Si la v.a. discrète  $X$  a la distribution ci-dessous, que vaut  $E[g(X)] = E[X^2]$  ?

Distribution de $X$					
valeurs possibles $x_i$	-2	-1	0	1	2
probabilités $p_i$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

La [méthode 1](#) livre directement

$$\begin{aligned}\rightsquigarrow E[X^2] &= \sum_{i=1}^5 (x_i)^2 p_i \\ &= (-2)^2 \times \frac{1}{5} + (-1)^2 \times \frac{1}{5} + (0)^2 \times \frac{1}{5} + (1)^2 \times \frac{1}{5} + (2)^2 \times \frac{1}{5} \\ &= \frac{10}{5} = 2\end{aligned}$$

Pour la [méthode 2](#), on obtient d'abord la distribution de  $Y = X^2$  :

Distribution de $Y = X^2$			
valeurs possibles $y_i$	0	1	4
probabilités $p_i^{(y)}$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$

( $P[Y = 4] = P[[X = -2] \cup [X = 2]] = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}$ , mais  $P[Y = 0] = P[X = 0] = \frac{1}{5}$ )

Ceci donne alors

$$\begin{aligned} E[X^2] = E[Y] &= \sum_{i=1}^3 y_i p_i^{(y)} \\ &= 0 \times \frac{1}{5} + 1 \times \frac{2}{5} + 4 \times \frac{2}{5} \\ &= \frac{10}{5} = 2 \end{aligned}$$

Propriétés principales de l'espérance mathématique :

## Théorème

Soient  $X_1, X_2, X$  des v.a. et  $a \in \mathbb{R}$ . Alors

(i)  $E[X_1 + X_2] = E[X_1] + E[X_2]$

(ii)  $E[aX_1] = aE[X_1]$

(iii)  $E[a] = a$

(iv) Si  $X_1 \leq X_2$  p.s., alors  $E[X_1] \leq E[X_2]$

(v) Si  $E[X] = 0$  et  $X \geq 0$  p.s., alors  $X = 0$  p.s.

(vi)  $|E[X_1 X_2]| \leq \sqrt{E[X_1^2]} \sqrt{E[X_2^2]}$  (inégalité de Cauchy-Schwarz), et l'égalité a lieu si et seulement si  $X_2 = cX_1$  p.s. ou  $X_1 = cX_2$  p.s. pour un certain  $c \in \mathbb{R}$ .

De (i)–(ii), il découle que  $E[\cdot]$  est un **opérateur linéaire** :

$$E[a_1 X_1 + a_2 X_2] = a_1 E[X_1] + a_2 E[X_2] \text{ pour toutes les v.a. } X_1, X_2 \text{ et pour tout } a_1, a_2 \in \mathbb{R}.$$

Dans l'exemple précédent, on a  $E[X^2] = 2 \neq 0^2 = (E[X])^2$ , ce qui montre qu'on n'a pas toujours  $E[X_1 X_2] = E[X_1]E[X_2]$ .



## 2 Variables aléatoires

- Définition et exemples
- Distribution de probabilité
- Opérations, égalité presque sûre et égalité en distribution
- Espérance mathématique
- **Variance, risque, inégalité de Tchebychev**
- Quelques distributions discrètes classiques
- Quelques distributions continues classiques
- Moments, fonction génératrice des moments

L'espérance mathématique  $E[X]$  de  $X$  est une quantité qui donne une information sur la grandeur (ou la position) de  $X$ , à travers sa valeur moyenne (ou "attendue").

Mais  $E[X]$  est loin de décrire de façon complète la distribution de  $X$ .

---

Exemple :

Considérons deux actions  $A_1$  et  $A_2$ , et leur évolution en une unité de temps fixée.

- Si l'économie reste telle qu'elle est, les deux actions vont augmenter de 5%.
- Si l'économie s'améliore,  $A_1$  va prendre 10% et  $A_2$  50%.
- Si l'économie se détériore,  $A_1$  va perdre 10% et  $A_2$  50%.

En outre, supposons que l'économie s'améliorera avec proba .2, se détériorera avec proba .2, donc restera inchangée avec proba .6.

Si on note respectivement  $X_1$  et  $X_2$  les pourcentages gagnés par  $A_1$  et  $A_2$ , on a

$$\begin{cases} E[X_1] = (-10\%) \times .2 + 5\% \times .6 + 10\% \times .2 = 3\% \\ E[X_2] = (-50\%) \times .2 + 5\% \times .6 + 50\% \times .2 = 3\% \end{cases}$$

Les bénéfices attendus pour  $A_1$  et  $A_2$  coïncident...

Mais les deux actions sont réellement différentes :

*Le risque associé à  $A_2$  est beaucoup plus élevé que celui associé à  $A_1$ .*

Cette caractéristique, qui est bien présente dans la distribution de  $X_1$  et  $X_2$ , ne l'est pas dans le résumé de ces distributions que constituent  $E[X_1]$  et  $E[X_2]$ .

Le concept de **variance** permet de mesurer le "**risque**" ou la **variabilité** d'une v.a.

## Définition

Supposons que  $E[X^2]$  soit bien définie dans  $\mathbb{R}$ . Alors la variance de  $X$  est

$$\sigma_X^2 = \text{Var}[X] = E[(X - \mu_X)^2] = \begin{cases} \sum_{i \in \mathcal{I}} (x_i - \mu_X)^2 p_i & \text{si } X \text{ est discrète} \\ \int_{-\infty}^{\infty} (x - \mu_X)^2 f(x) dx & \text{si } X \text{ est continue,} \end{cases}$$

où  $\mu_X = E[X]$ .

Interprétation :  $\sigma_X^2$  est la valeur attendue du carré de l'écart  $|X - \mu_X|$  entre  $X$  et  $\mu_X$ .

---

Notons que la variance est bien définie dans  $\mathbb{R}$  parce que

- $E[X^2] < \infty \Rightarrow E[|X|] < \infty$  (voir ch.2-p.90), ce qui fournit  $\mu_X = E[X] \in \mathbb{R}$ .
- $\text{Var}[X] = E[(X - \mu_X)^2] \leq E[2X^2 + 2\mu_X^2] = 2E[X^2] + 2\mu_X^2 < \infty$

## Théorème

Soient  $X$  une v.a. telle que  $E[X^2]$  soit bien définie dans  $\mathbb{R}$  et  $a \in \mathbb{R}$ . Alors

(i)  $\text{Var}[X + a] = \text{Var}[X]$

(ii)  $\text{Var}[aX] = a^2 \text{Var}[X]$

(iii)  $\text{Var}[a] = 0$

(iv)  $\text{Var}[X] = E[X^2] - (\mu_X)^2$

(v) Si  $\text{Var}[X] = 0$ , alors  $X = c$  p.s. pour un certain  $c \in \mathbb{R}$ .

Le point (iv), qui facilite le calcul de la variance en pratique, découle du fait que

$$\begin{aligned}\text{Var}[X] &= E[(X - \mu_X)^2] \\ &= E[X^2 - 2\mu_X X + (\mu_X)^2] \\ &= E[X^2] - 2\mu_X E[X] + E[(\mu_X)^2] \\ &= E[X^2] - 2(\mu_X)^2 + (\mu_X)^2 \\ &= E[X^2] - (\mu_X)^2.\end{aligned}$$

Le point (v) est une conséquence directe du point (v) du théorème en page 43.

En pratique, on utilise ainsi

$$\text{Var}[X] = E[X^2] - (\mu_X)^2 = \begin{cases} \sum_{i \in \mathcal{I}} (x_i)^2 p_i - (\mu_X)^2 & \text{pour } X \text{ discrète} \\ \int_{-\infty}^{\infty} x^2 f(x) dx - (\mu_X)^2 & \text{pour } X \text{ continue.} \end{cases}$$

---

L'exemple des deux actions :

Puisque

$$\begin{cases} E[X_1^2] = (-10\%)^2 \times .2 + (5\%)^2 \times .6 + (10\%)^2 \times .2 = 55(\%)^2 \\ E[X_2^2] = (-50\%)^2 \times .2 + (5\%)^2 \times .6 + (50\%)^2 \times .2 = 1015(\%)^2, \end{cases}$$

on obtient

$$\begin{cases} \text{Var}[X_1] = E[X_1^2] - (\mu_{X_1})^2 = 55(\%)^2 - (3\%)^2 = 46(\%)^2 \\ \text{Var}[X_2] = E[X_2^2] - (\mu_{X_2})^2 = 1015(\%)^2 - (3\%)^2 = 1006(\%)^2, \end{cases}$$

ce qui traduit bien le côté **plus "volatile"** de la seconde action.

Dans cet exemple, les variances se mesurent en  $(\%)^2$ ...

↪ Parfois, on mesure plutôt la dispersion par l'écart-type, qui a l'avantage de s'exprimer **dans les mêmes unités que  $X$** .

## Définition

L'écart-type de  $X$  est  $\sigma_X = \sqrt{\text{Var}[X]}$ .

Dans l'exemple précédent, on a  $\sigma_{X_1} \approx 6.78\%$  et  $\sigma_{X_2} \approx 31.72\%$ .

---

A  $X$ , on associera souvent sa version "**centrée-réduite**"

$$Z = \frac{X - \mu_X}{\sigma_X},$$

qui satisfait toujours  $\mathbf{E}[Z] = 0$  et  $\mathbf{Var}[Z] = 1$  (exercice).

# Inégalité de Tchebychev

## Théorème (inégalité de Tchebychev)

Soit une v.a.  $X$  telle que  $E[X^2]$  soit bien définie dans  $\mathbb{R}$  et  $a > 0$ .

Alors  $P[|X - \mu_X| \geq a\sigma_X] \leq 1/a^2$

"il y a au plus une chance sur  $a^2$  que  $X$  s'écarte de  $\mu_X$  de plus de  $a$  fois  $\sigma_X$ ".

Preuve : Posons

$$Y = \begin{cases} 1 & \text{si } (X - \mu_X)^2 \geq a^2 \sigma_X^2 \\ 0 & \text{sinon.} \end{cases}$$

On a toujours  $(X - \mu_X)^2 \geq a^2 \sigma_X^2 Y$  (pourquoi ?) Donc on a

$$\begin{aligned} \sigma_X^2 &= E[(X - \mu_X)^2] \geq E[a^2 \sigma_X^2 Y] = a^2 \sigma_X^2 E[Y] \\ &= a^2 \sigma_X^2 (1 \times P[Y = 1] + 0 \times P[Y = 0]) \\ &= a^2 \sigma_X^2 P[(X - \mu_X)^2 \geq a^2 \sigma_X^2], \\ &= a^2 \sigma_X^2 P[|X - \mu_X| \geq a\sigma_X], \end{aligned}$$

ce qui fournit le résultat.

□



## 2 Variables aléatoires

- Définition et exemples
- Distribution de probabilité
- Opérations, égalité presque sûre et égalité en distribution
- Espérance mathématique
- Variance, risque, inégalité de Tchebychev
- Quelques distributions discrètes classiques
- Quelques distributions continues classiques
- Moments, fonction génératrice des moments

# Quelques distributions discrètes classiques

(i)  $X$  est de **distribution de Bernoulli de paramètre  $p$**  ( $\in [0, 1]$ )  
(notation :  $X \sim \text{Bern}(p)$ ) si sa distribution est donnée par

valeurs possibles	0	1
probabilités	$1 - p$	$p$

Ceci modélise des expériences de type "succès (1) ou échec (0)".

On vérifie directement qu'on a  $E[X] = p$  et  $\text{Var}[X] = p(1 - p)$ .

---

Si  $A$  est un événement relatif à l'espace probabilisé  $(\Omega, \mathcal{A}, P)$ , la v.a.

$$X = I_A := \begin{cases} 1 & \text{si } A \text{ se produit} \\ 0 & \text{sinon} \end{cases}$$

est de loi de Bernoulli de paramètre  $p = P[A]$  (on utilisera indistinctement les termes **distribution** et **loi**).

# Quelques distributions discrètes classiques

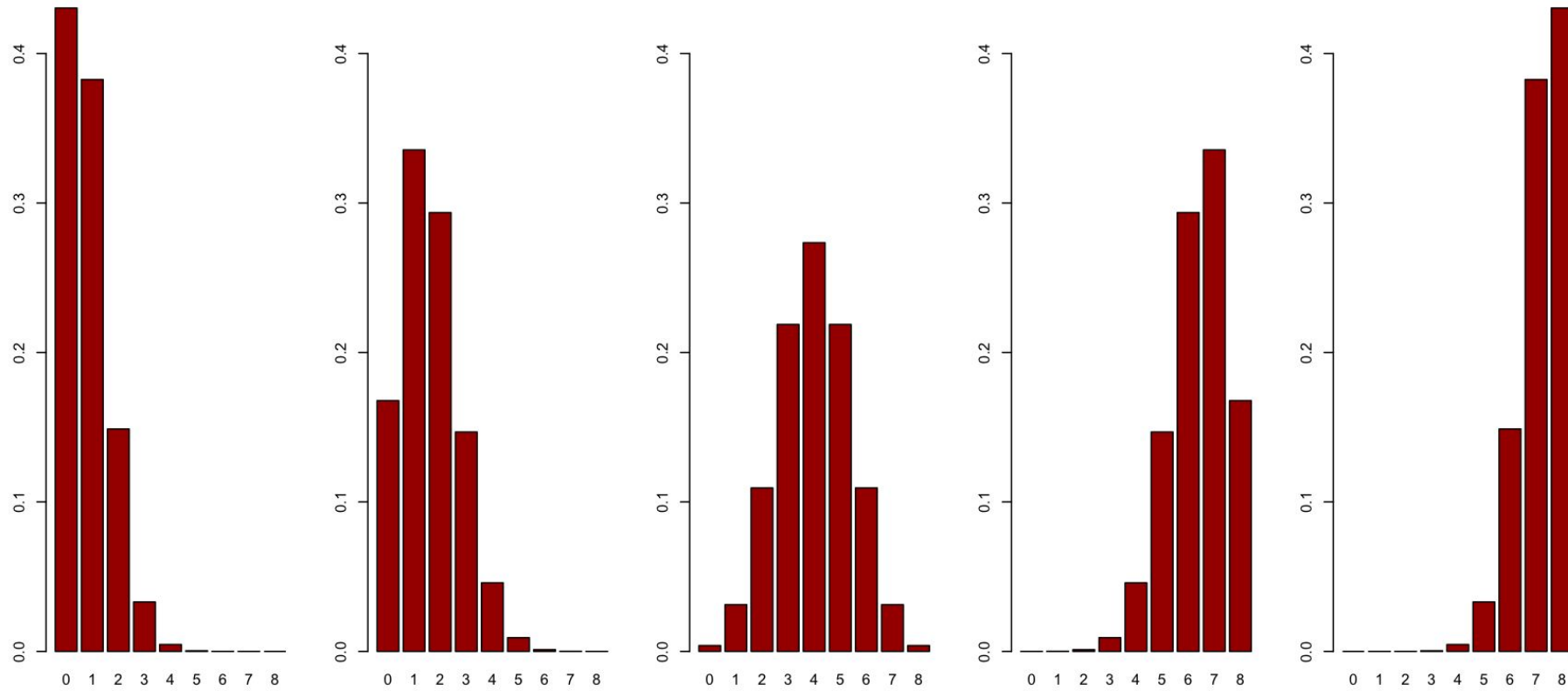
(ii)  $X$  est de **distribution binomiale de paramètres**  $n$  ( $\in \mathbb{N}_0$ ) **et**  $p$  ( $\in [0, 1]$ )  
(notation :  $X \sim \text{Bin}(n, p)$ ) si sa distribution est donnée par

Distribution de $X$						
valeurs possibles	0	1	...	$k$	...	$n$
probabilités	$(1 - p)^n$	$np(1 - p)^{n-1}$	...	$\binom{n}{k} p^k (1 - p)^{n-k}$	...	$p^n$

La v.a.  $X$  compte le nombre de succès dans une suite de  $n$  expériences de type "succès-échec" (où un succès se produit chaque fois avec probabilité  $p$ ) répétées de façon indépendante.

On vérifie que  $E[X] = np$  et  $\text{Var}[X] = np(1 - p)$ .

# Quelques distributions discrètes classiques



Répartition des probabilités de  $\text{Bin}(n, p)$ , pour  $n = 8$  et  $p = .1, .2, .5, .8, .9$

# Quelques distributions discrètes classiques

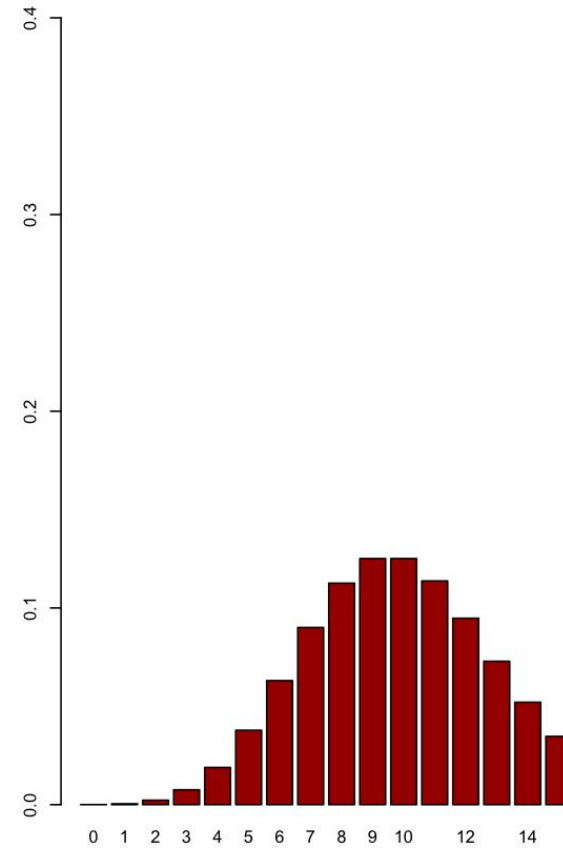
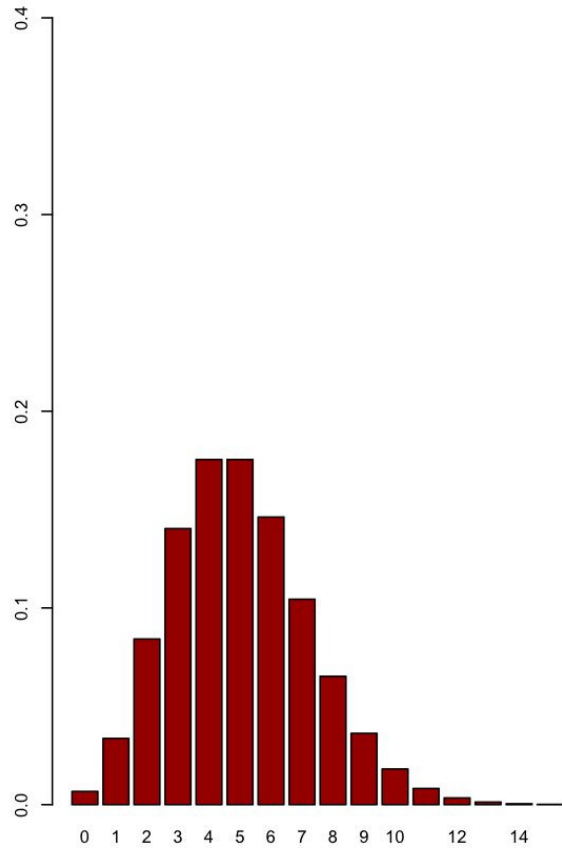
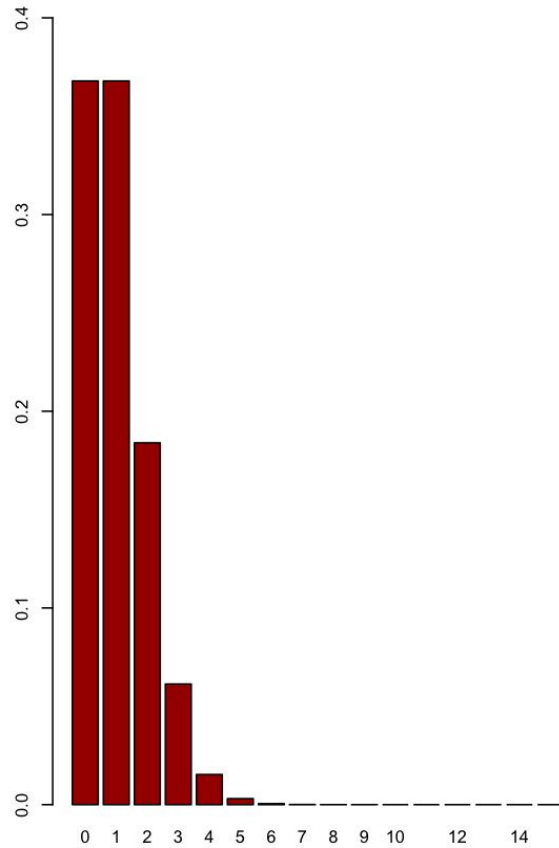
- (iii)  $X$  est de **distribution de Poisson de paramètre  $\lambda$**  ( $\in \mathbb{R}_0^+$ )  
(notation :  $X \sim \text{Poi}(\lambda)$  ou  $X \sim \mathcal{P}(\lambda)$ ) si sa distribution est donnée par

Distribution de $X$	
valeurs possibles	0      1      ... $k$ ...
probabilités	$e^{-\lambda}$ $e^{-\lambda}\lambda$ ... $e^{-\lambda}\lambda^k/(k!)$ ...

Ci-dessous, on verra qu'une v.a.  $X$  de ce type compte le nombre de réalisations d'un événement rare au cours d'une certaine période de temps (nombre de décès suite à une maladie rare au cours d'une année, nombre de Ferrari passant par le boulevard Général Jacques au cours d'une journée, etc.)

On vérifie que  $E[X] = \lambda$  et  $\text{Var}[X] = \lambda$ .

# Quelques distributions discrètes classiques



Répartition des probabilités de  $\text{Poi}(\lambda)$ , pour  $\lambda = 1, 5, 10$

# Quelques distributions discrètes classiques

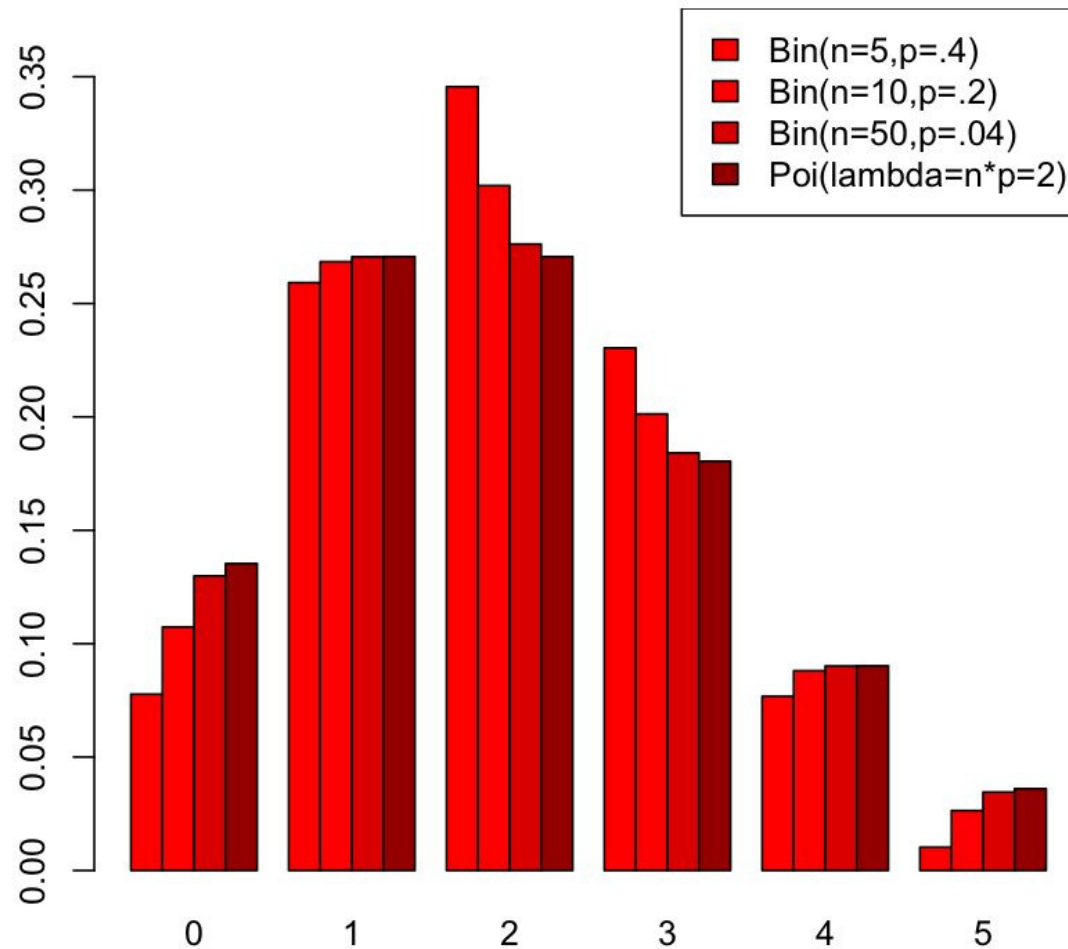
Si  $n \rightarrow \infty$  et  $p \rightarrow 0$  de telle manière que  $np \rightarrow \lambda (> 0)$ ,

$$P[\text{Bin}(n, p) = k] \rightarrow P[\text{Poi}(\lambda) = k] \quad \forall k \in \mathbb{N}.$$

Ce résultat

- justifie l'interprétation de la distribution de Poisson (=le comptage du nombre de réalisations d'un événement rare au cours d'une certaine période de temps)
- permet d'approximer une  $\text{Bin}(n, p)$  par une  $\text{Poi}(np)$  (pour  $n$  grand,  $p$  petit)
- est compatible avec le fait que  $E[X] = \lambda$  et  $\text{Var}[X] = \lambda$ .

# Quelques distributions discrètes classiques



Approximation d'une binomiale par une Poisson



## Quelques distributions discrètes classiques

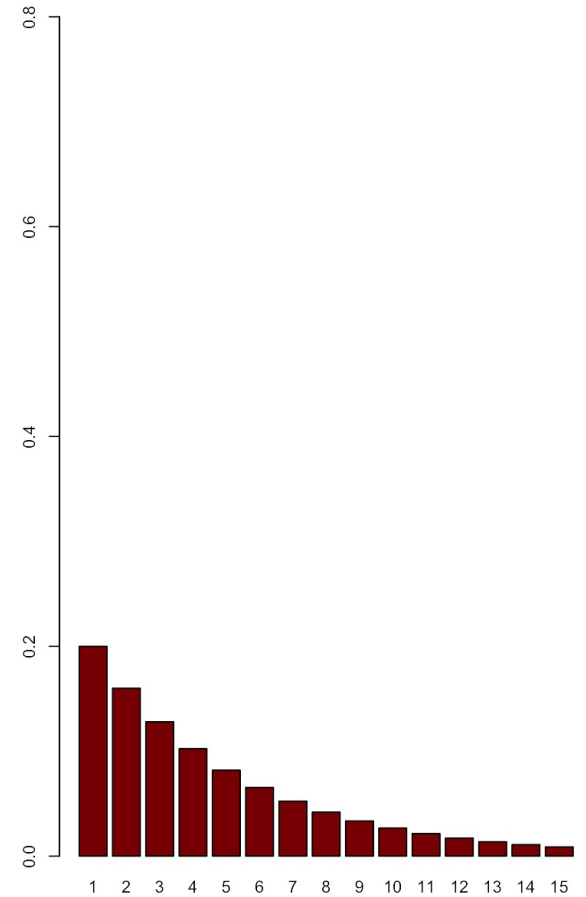
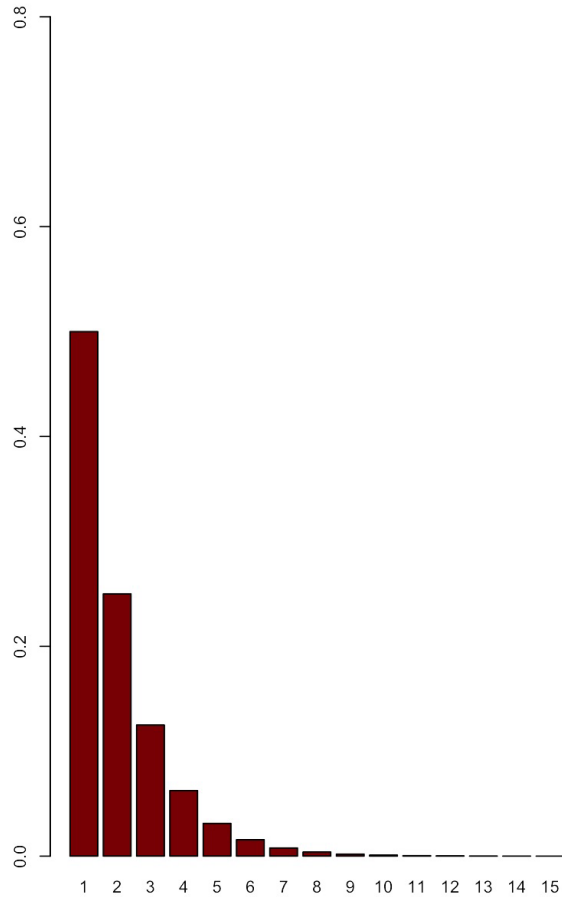
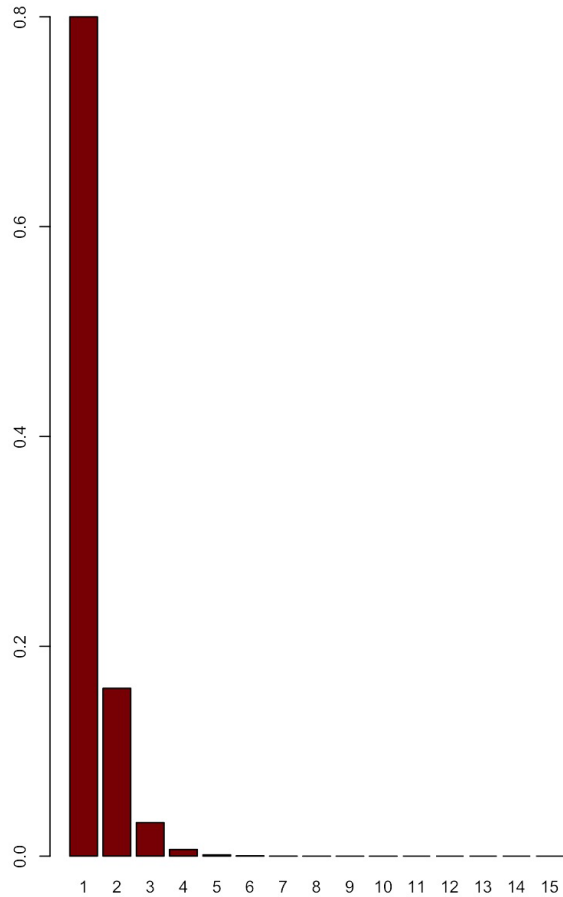
(iv)  $X$  est de **distribution géométrique de paramètre  $p$**  ( $\in (0, 1)$ )  
(notation :  $X \sim \text{Geom}(p)$ ) si sa distribution est donnée par

Distribution de $X$					
valeurs possibles	1	2	...	$k$	...
probabilités	$p$	$(1 - p)p$	...	$(1 - p)^{k-1}p$	...

La v.a.  $X$  compte le nombre de répétitions (indépendantes) d'une expérience de type succès-échec nécessaires pour avoir le premier succès.

On vérifie que  $E[X] = 1/p$  et  $\text{Var}[X] = (1 - p)/p^2$ .

# Quelques distributions discrètes classiques



Répartition des probabilités de  $\text{Geom}(p)$ , pour  $p = .8, .5, .2$

# Quelques distributions discrètes classiques

On a donc considéré les distributions

- de Bernoulli,
- binomiales,
- de Poisson, et
- géométriques.

On pourrait encore présenter les distributions

- binomiales négatives,
- hypergéométriques,
- ...

Nous renvoyons cependant aux monographies donnés en référence pour l'étude de ces distributions supplémentaires

## **2** Variables aléatoires

- Définition et exemples
- Distribution de probabilité
- Opérations, égalité presque sûre et égalité en distribution
- Espérance mathématique
- Variance, risque, inégalité de Tchebychev
- Quelques distributions discrètes classiques
- **Quelques distributions continues classiques**
- Moments, fonction génératrice des moments

# Quelques distributions continues classiques

(i)  $X$  est de **distribution uniforme sur  $(a, b)$**  ( $-\infty < a < b < \infty$ )  
(notation :  $X \sim \text{Unif}(a, b)$ ) si  $X$  admet la fonction de densité définie par

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

Ceci fournit

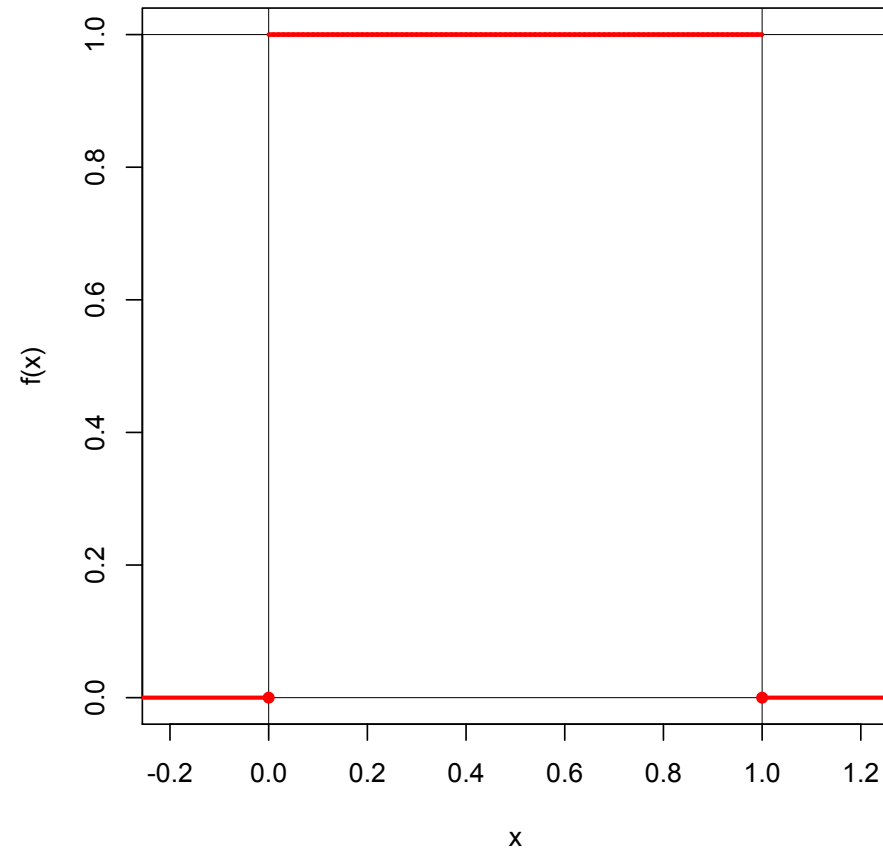
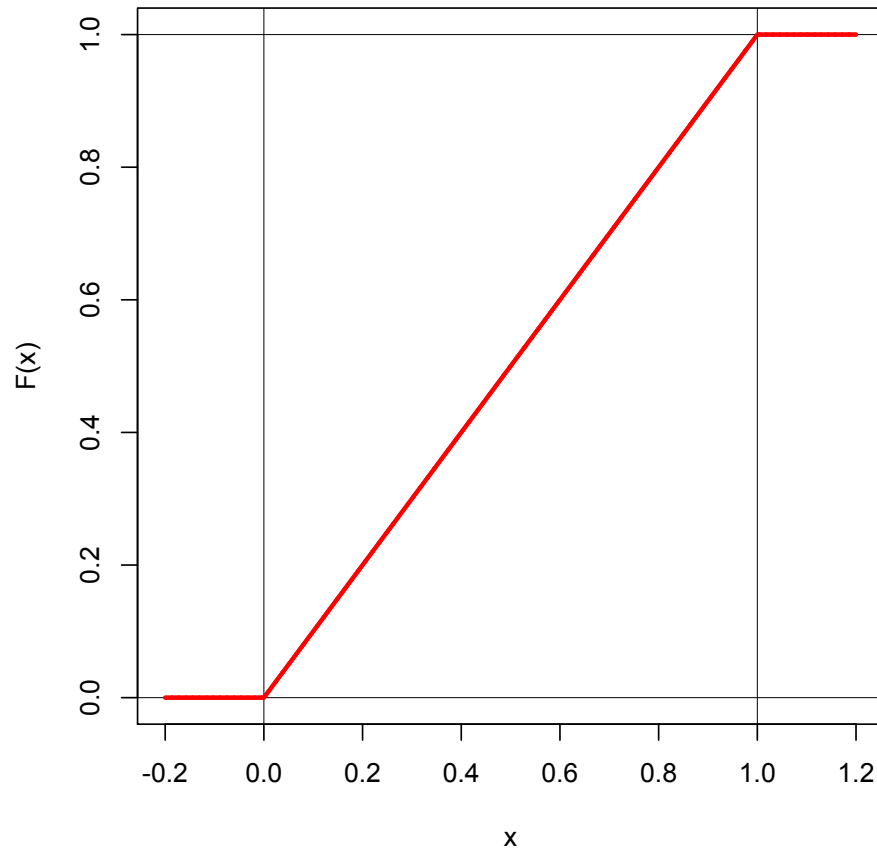
$$F(x) = \begin{cases} 0 & \text{si } x \leq a \\ \frac{x-a}{b-a} & \text{si } x \in (a, b) \\ 1 & \text{si } x \geq b. \end{cases}$$

Pour  $[c, c+h] \subset (a, b)$ , on a  $P[c \leq X \leq c+h] = h/(b-a)$ , qui ne dépend pas de la position de l'intervalle  $[c, c+h]$  dans  $(a, b)$   $\rightsquigarrow$  **Distribution uniforme**

---

Le temps d'attente du bus 71 est une v.a. de loi  $\text{Unif}(0, d)$ , où  $d$  est l'écart (supposé constant) entre deux bus consécutifs !

# Quelques distributions continues classiques



Fonction de répartition ( $F$ ) et fonction de densité ( $f$ ) de  $X \sim \text{Unif}(0, 1)$

# Quelques distributions continues classiques

On a

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_a^b x \times \frac{1}{b-a} dx = \left[ \frac{x^2}{2(b-a)} \right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}.$$

D'autre part, comme

$$\begin{aligned} E[X^2] &= \int_{-\infty}^{\infty} x^2 f(x) dx = \int_a^b x^2 \times \frac{1}{b-a} dx \\ &= \left[ \frac{x^3}{3(b-a)} \right]_a^b = \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3}, \end{aligned}$$

on obtient

$$\text{Var}[X] = E[X^2] - (E[X])^2 = \dots = \frac{(b-a)^2}{12}.$$

# Quelques distributions continues classiques

(ii)  $X$  est de **distribution exponentielle de paramètre  $\lambda$**  ( $\in \mathbb{R}_0^+$ )  
(notation :  $X \sim \text{Exp}(\lambda)$ ) si  $X$  admet la fonction de densité définie par

$$f(x) = \begin{cases} \frac{1}{\lambda} e^{-x/\lambda} & \text{si } x \geq 0 \\ 0 & \text{sinon.} \end{cases}$$

Ceci fournit

$$F(x) = \begin{cases} 1 - e^{-x/\lambda} & \text{si } x \geq 0 \\ 0 & \text{sinon.} \end{cases}$$

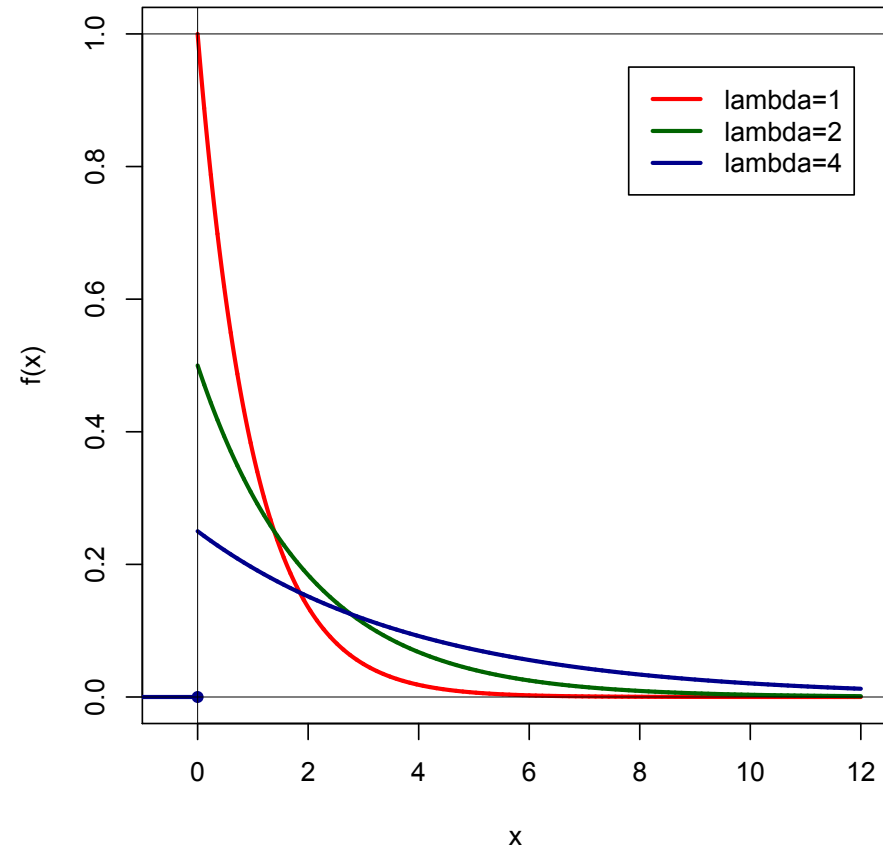
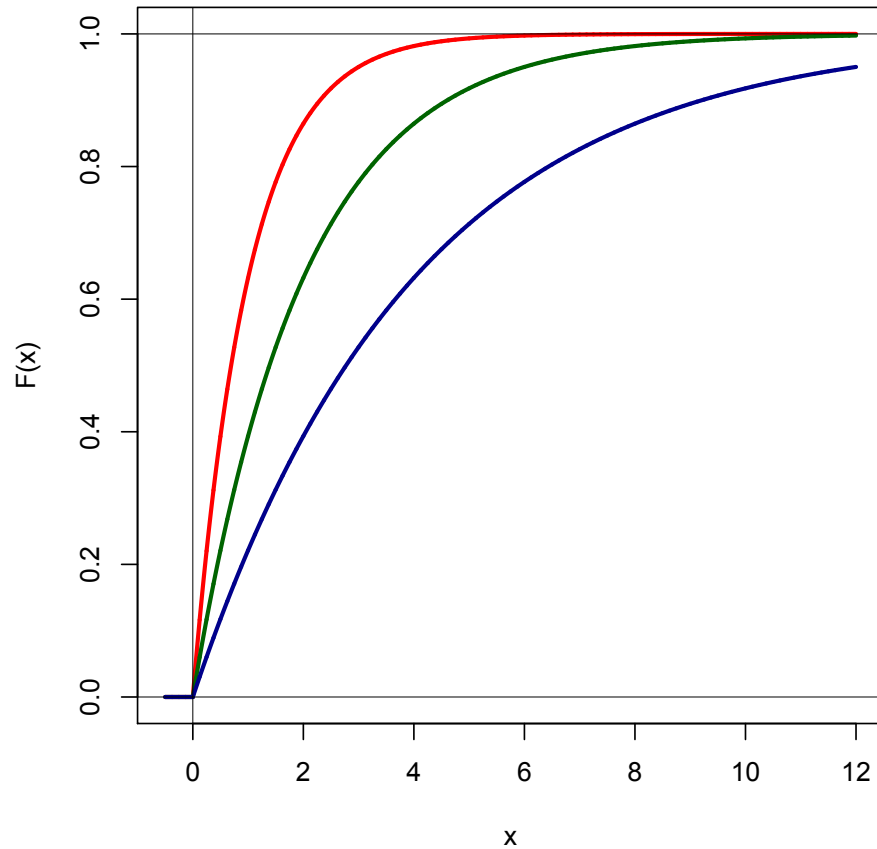
On vérifiera facilement que

$$E[X] = \lambda \quad \text{et} \quad \text{Var}[X] = \lambda^2$$

(exercice).



# Quelques distributions continues classiques



Fonction de répartition ( $F$ ) et fonction de densité ( $f$ ) de  $X \sim \text{Exp}(\lambda)$ ,  $\lambda = 1, 2$  et  $4$

# Quelques distributions continues classiques

Cette distribution est souvent utilisée pour modéliser des durées de vie (ou des temps d'attente)

---

La distribution exponentielle est "**sans mémoire**" : pour tout  $x, h > 0$ ,

$$\begin{aligned} P[X > x + h | X > x] &= \frac{P[[X > x + h] \cap [X > x]]}{P[X > x]} = \frac{P[X > x + h]}{P[X > x]} \\ &= \frac{1 - P[X \leq x + h]}{1 - P[X \leq x]} = \frac{e^{-(x+h)/\lambda}}{e^{-x/\lambda}} = e^{-h/\lambda} = 1 - P[X \leq h] = P[X > h]. \end{aligned}$$

*"La probabilité de survivre  $h$  unités de temps supplémentaires ne dépend pas de l'âge  $x$  atteint".*

Exercice : montrer que la distribution géométrique satisfait une propriété similaire

# Quelques distributions continues classiques

(iii)  $X$  est de **distribution normale centrée réduite** (notation :  $X \sim \mathcal{N}(0, 1)$ ) si  $X$  admet la fonction de densité définie par

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad (\stackrel{\text{not}}{=} \varphi(x)).$$

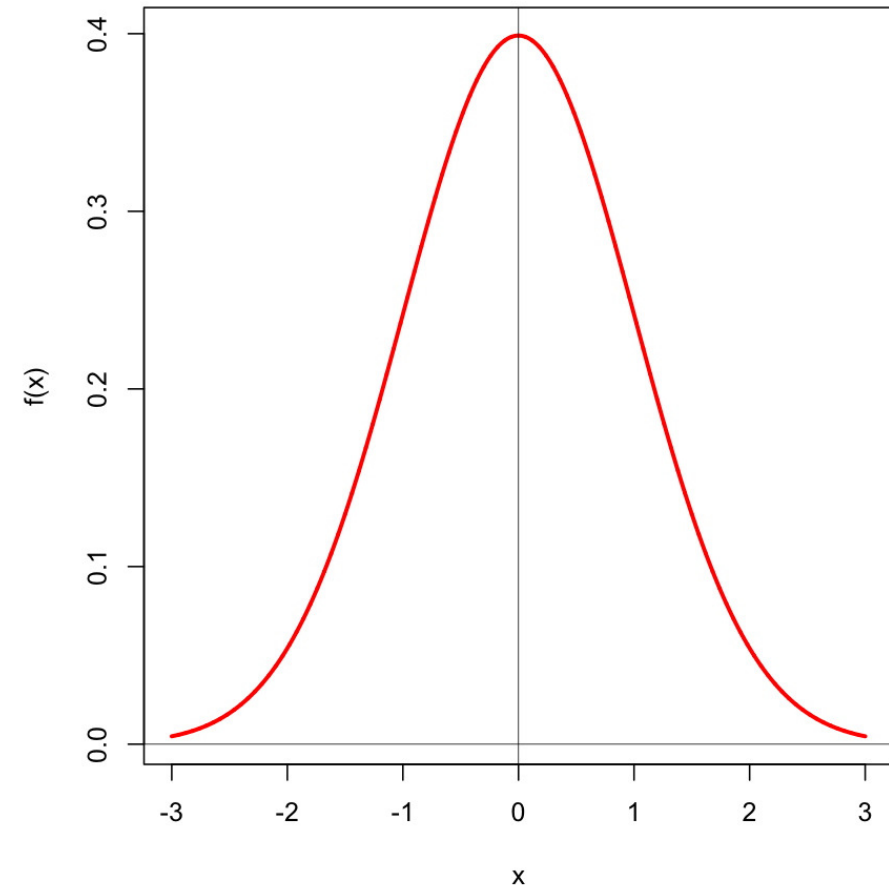
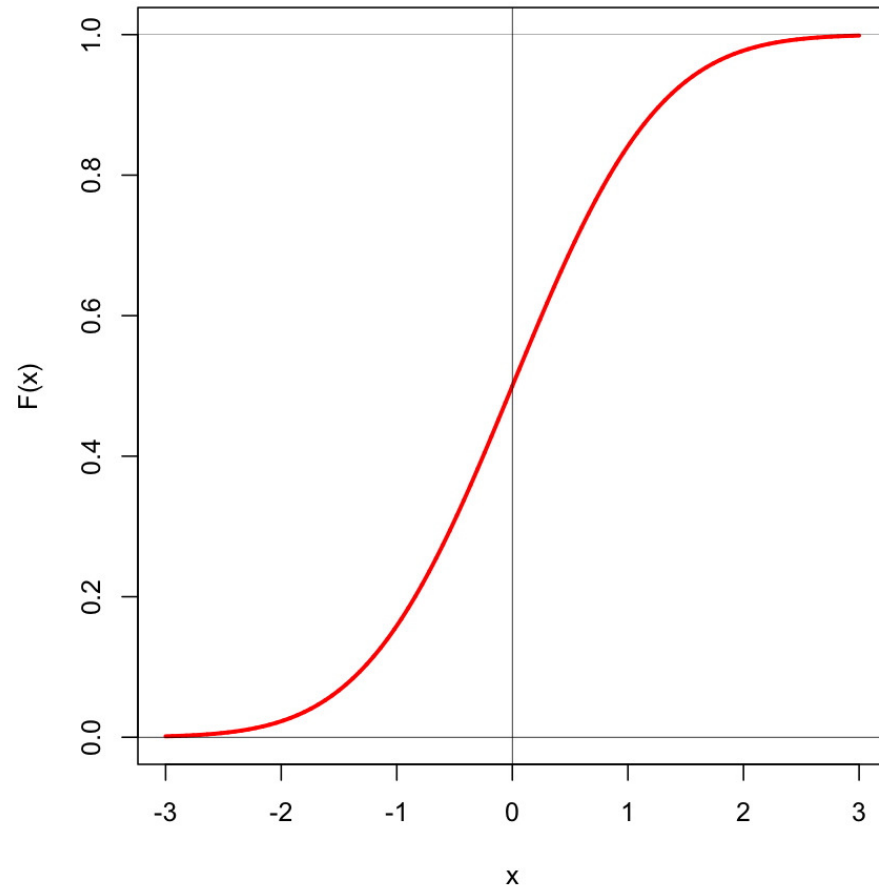
La fonction de répartition associée

$$F(x) = \int_{-\infty}^x \varphi(y) dy \quad (\stackrel{\text{not}}{=} \Phi(x))$$

n'admet pas de forme explicite.

On vérifie que  $E[X] = 0$  et  $\text{Var}[X] = 1$ .

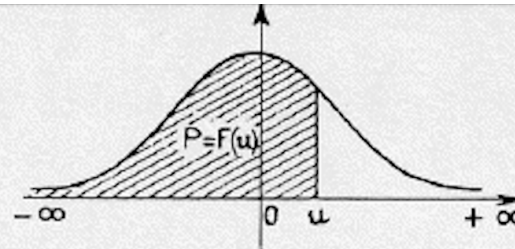
# Quelques distributions continues classiques



Fonction de répartition ( $F = \Phi$ ) et fonction de densité ( $f = \varphi$ ) de  $X \sim \mathcal{N}(0, 1)$

Le graphe de  $\varphi$  est souvent appelé **cloche de Gauss**

# Quelques distributions continues classiques

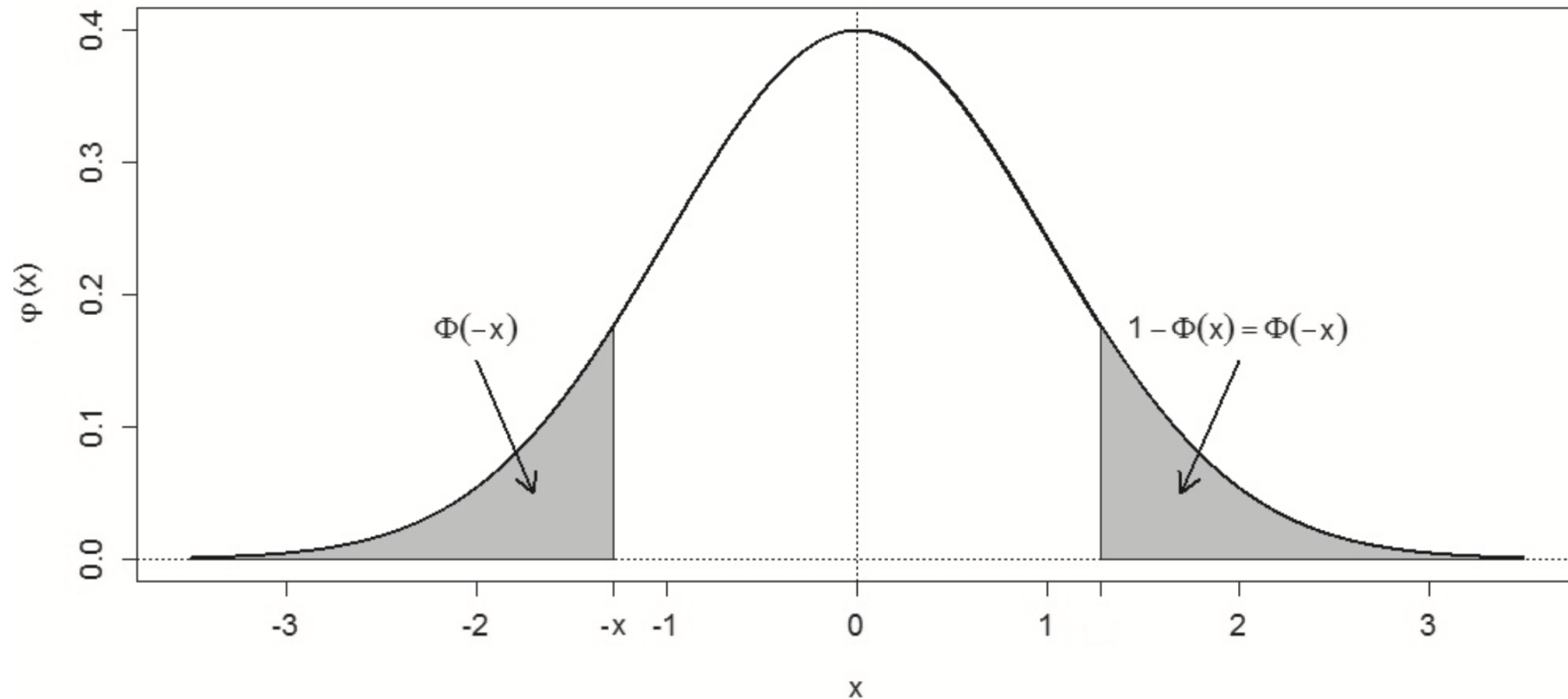


u	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7290	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767

Table de  $\Phi(z)$  pour  $z \geq 0$

# Quelques distributions continues classiques

## Propriétés de symétrie de la loi $\mathcal{N}(0, 1)$



Si  $X \sim \mathcal{N}(0, 1)$ , on a  $P[X < -x] = P[X > x]$  pour tout  $x \geq 0$   
(c'est pour cela qu'on ne tabule  $\Phi(x)$  que pour  $x \geq 0$ )

# Quelques distributions continues classiques

(iii)  $X$  est de **distribution normale de paramètres**  $\mu$  ( $\in \mathbb{R}$ ) et  $\sigma^2$  ( $\in \mathbb{R}_0^+$ )  
(notation :  $X \sim \mathcal{N}(\mu, \sigma^2)$ ) si  $X \stackrel{D}{=} \sigma Z + \mu$ , où  $Z \sim \mathcal{N}(0, 1)$ .

La fonction de répartition de  $X$  est  $F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$  et la densité de  $X$  est

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

(exercice). De la définition ci-dessus, il découle directement que

$$\mathbf{E}[X] = \mathbf{E}[\sigma Z + \mu] = \sigma \mathbf{E}[Z] + \mu = \mu$$

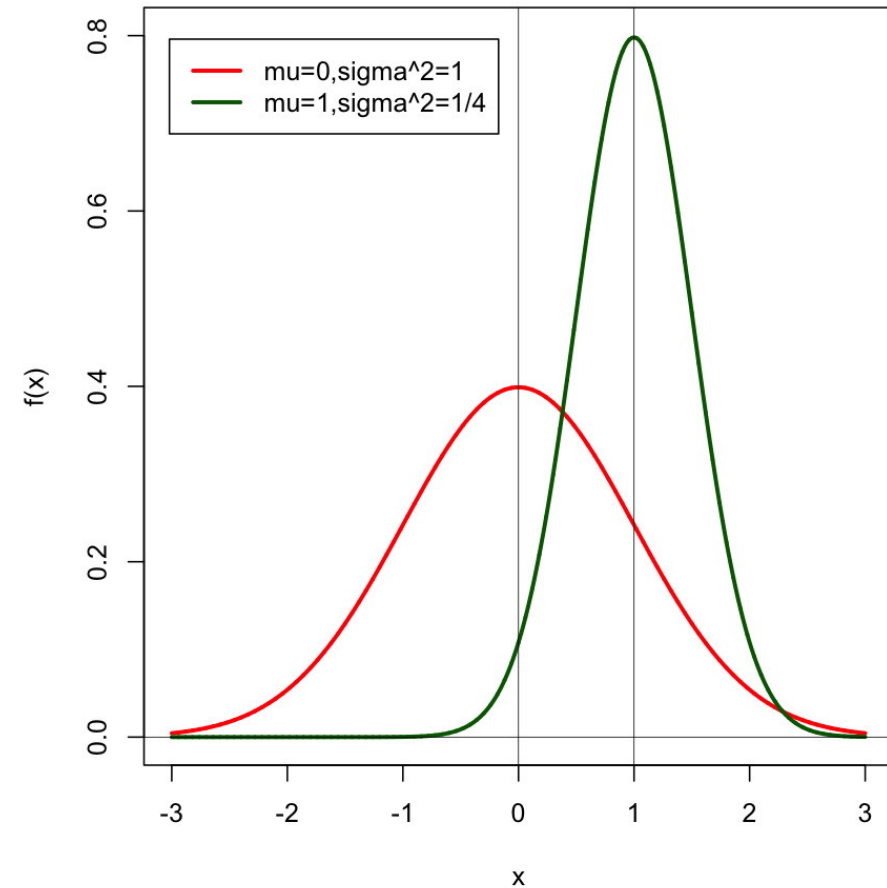
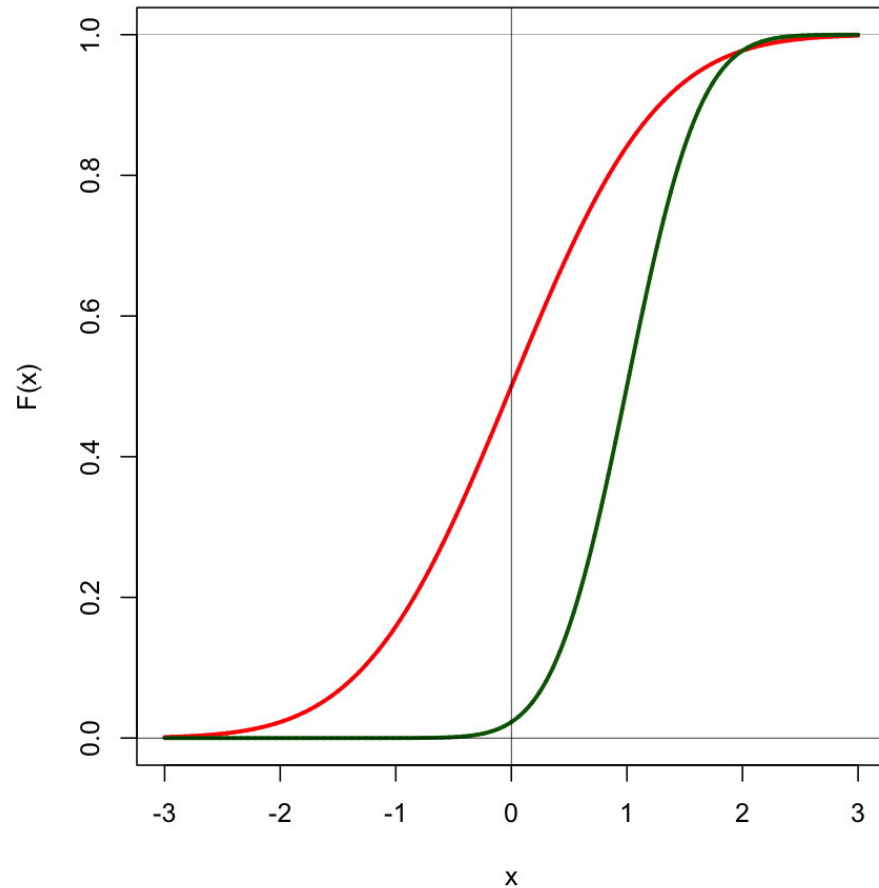
$$\mathbf{Var}[X] = \mathbf{Var}[\sigma Z + \mu] = \mathbf{Var}[\sigma Z] = \sigma^2 \mathbf{Var}[Z] = \sigma^2,$$

ce qui donne une interprétation claire aux paramètres de la loi  $\mathcal{N}(\mu, \sigma^2)$ .

---

Si  $X \sim \mathcal{N}(\mu, \sigma^2)$ , alors sa version centrée-réduite  $Z = (X - \mu)/\sigma$  est de loi  $\mathcal{N}(0, 1)$

# Quelques distributions continues classiques



Fonction de répartition ( $F$ ) et fonction de densité ( $f$ ) de  $X \sim \mathcal{N}(0, 1)$  et de  $X \sim \mathcal{N}(1, \frac{1}{4})$



# Quelques distributions continues classiques



L'ancien billet de 10 Deutsche Marks

## Quelques distributions continues classiques

(iv)  $X$  est de **distribution lognormale de paramètres**  $\mu \in \mathbb{R}$  et  $\sigma^2 \in \mathbb{R}_0^+$   
(notation :  $X \sim \mathcal{LN}(\mu, \sigma^2)$ ) si  $X \stackrel{\mathcal{D}}{=} e^{\mu + \sigma Z}$ , où  $Z \sim \mathcal{N}(0, 1)$ .

On vérifiera (exercice) que la fonction de répartition de  $X$  est

$$F(x) = \begin{cases} \Phi\left(\frac{(\ln x) - \mu}{\sigma}\right) & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$$

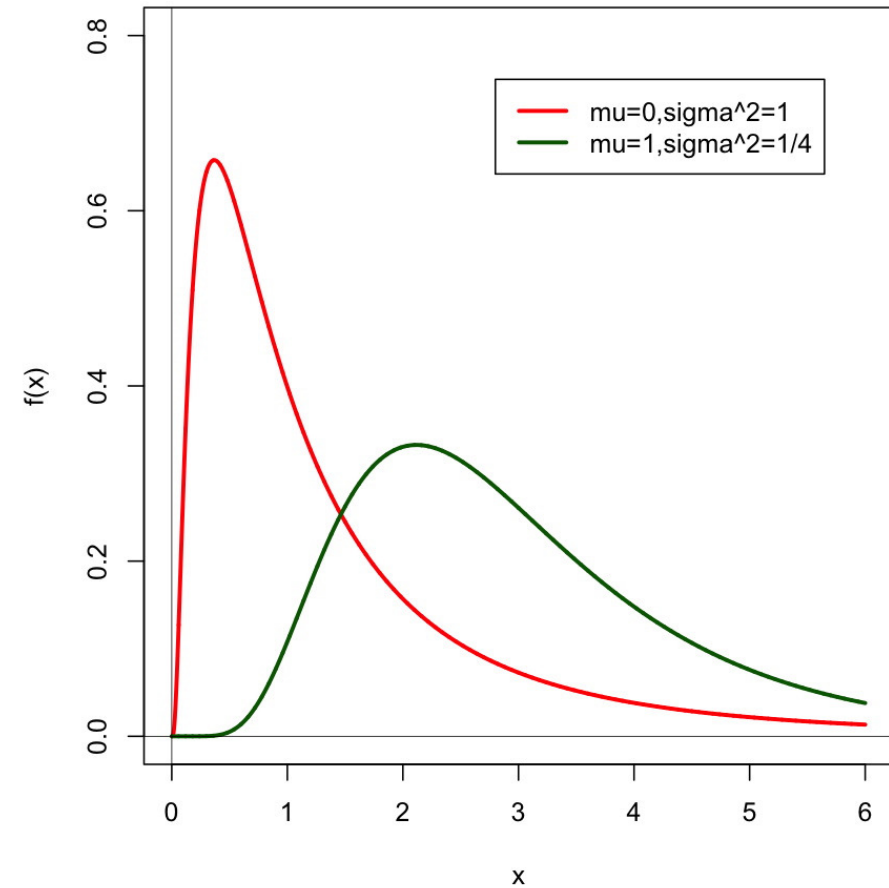
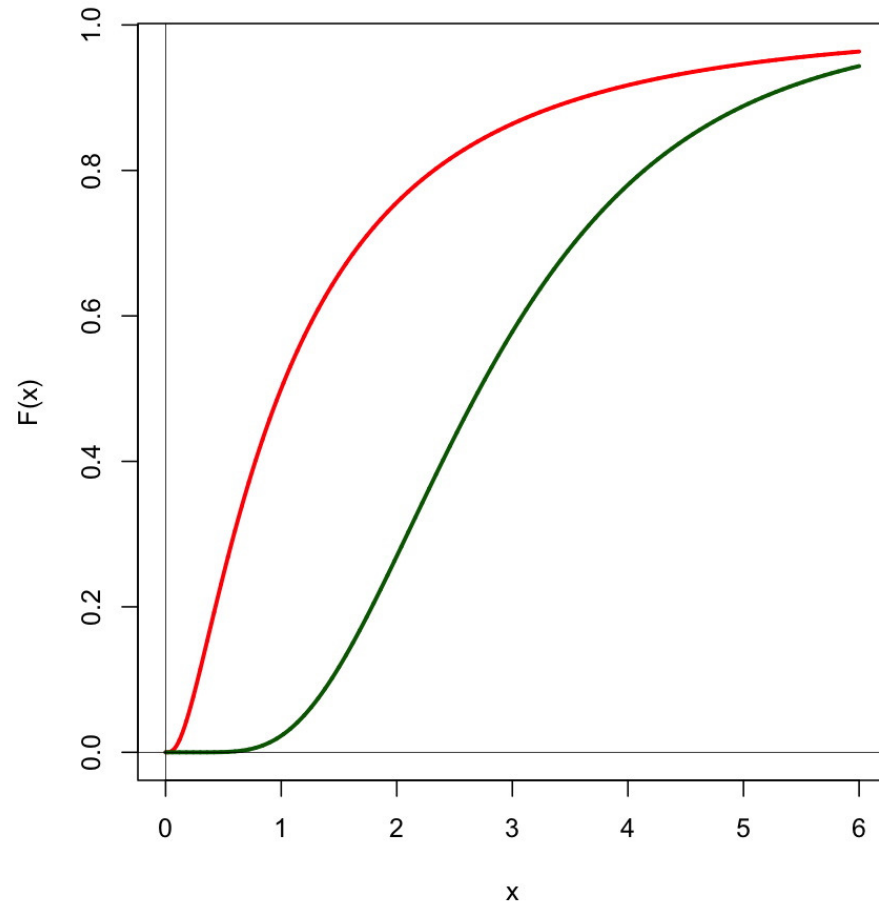
et que sa fonction de densité est

$$f(x) = \begin{cases} \frac{1}{x\sqrt{2\pi\sigma^2}} e^{-\frac{((\ln x) - \mu)^2}{2\sigma^2}} & \text{si } x > 0 \\ 0 & \text{sinon.} \end{cases}$$

On en déduit que la médiane de cette distribution vaut  $F^{-1}\left(\frac{1}{2}\right) = e^{\mu}$ .

Enfin, on peut montrer que  $E[X] = e^{\mu + (\sigma^2/2)}$  et  $\text{Var}[X] = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}$ .

# Quelques distributions continues classiques



Fonction de répartition ( $F$ ) et fonction de densité ( $f$ ) de  $X \sim \mathcal{LN}(0, 1)$  et de  $X \sim \mathcal{LN}(1, \frac{1}{4})$

## Quelques distributions continues classiques

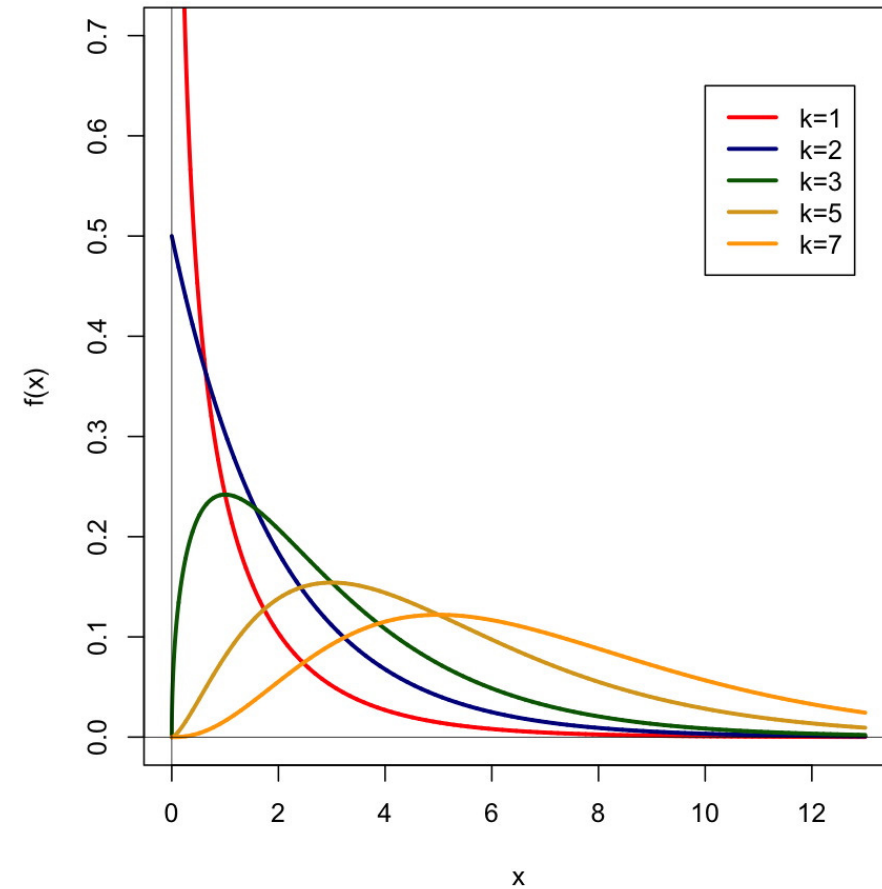
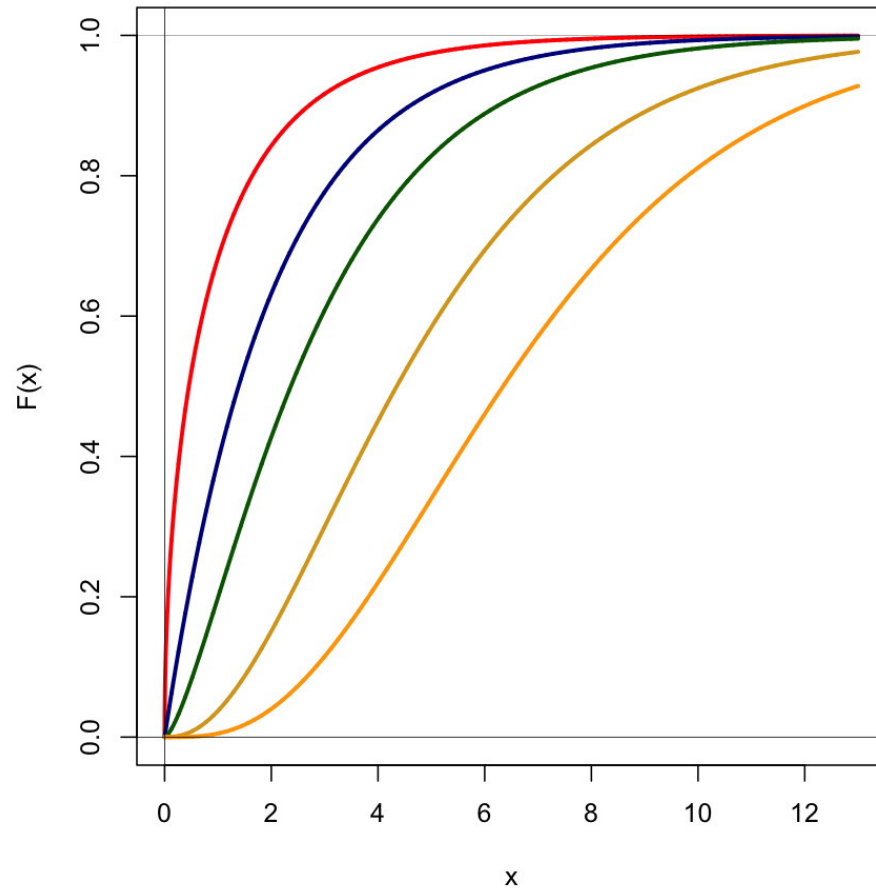
(v)  $X$  est de **distribution chi-carré à  $k$  degrés de liberté** ( $k \in \mathbb{N}_0$ ) (notation :  $X \sim \chi_k^2$ ) si  $X$  admet la fonction de densité définie par

$$f(x) = \begin{cases} \frac{1}{2^{\frac{k}{2}} \Gamma(\frac{k}{2})} x^{\frac{k}{2}-1} e^{-x/2} & \text{si } x \geq 0 \\ 0 & \text{sinon,} \end{cases}$$

où  $x \mapsto \Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$  est la fonction Gamma d'Euler.

On peut montrer que  $E[X] = k$  et  $\text{Var}[X] = 2k$ .

# Quelques distributions continues classiques



Fonction de répartition ( $F$ ) et fonction de densité ( $f$ ) de  $X \sim \chi_k^2$ ,  $k = 1, 2, 3, 5$ , et  $7$

## Quelques distributions continues classiques

(vi)  $X$  est de **distribution de Student à  $\nu$  degrés de liberté** ( $\nu \in \mathbb{N}_0$ )  
(notation :  $X \sim t_\nu$ ) si  $X$  admet la fonction de densité définie par

$$f(x) = \frac{\Gamma((\nu + 1)/2)}{\Gamma(\nu/2)\sqrt{\nu\pi}} (1 + x^2/\nu)^{-(\nu+1)/2}.$$

On peut montrer que

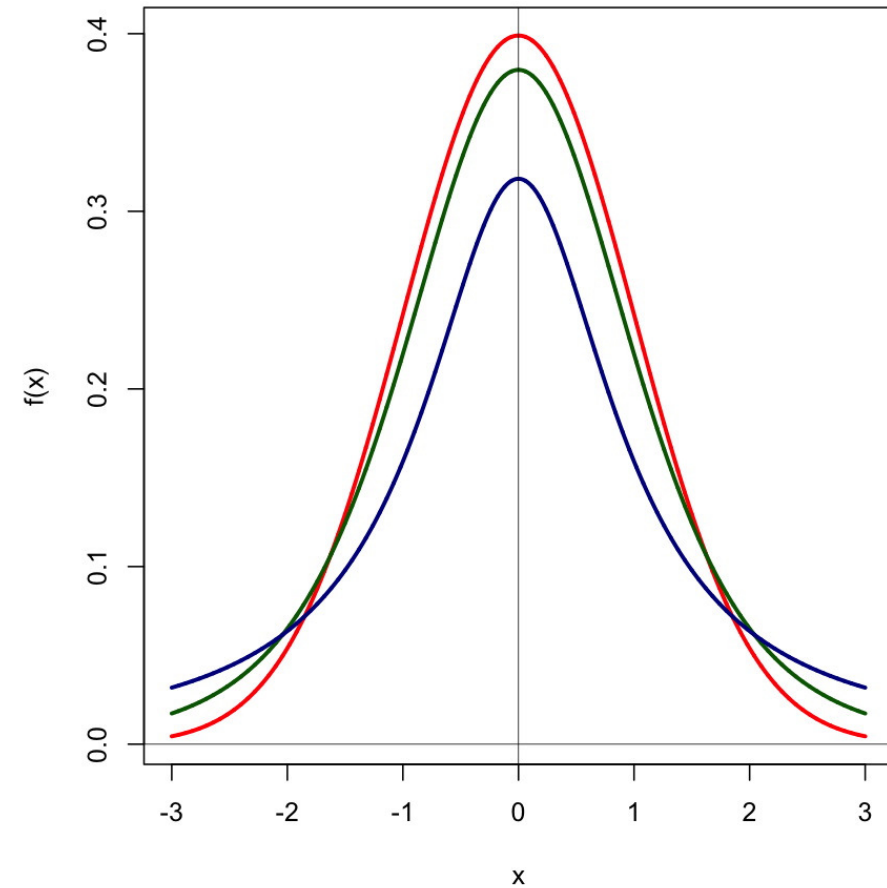
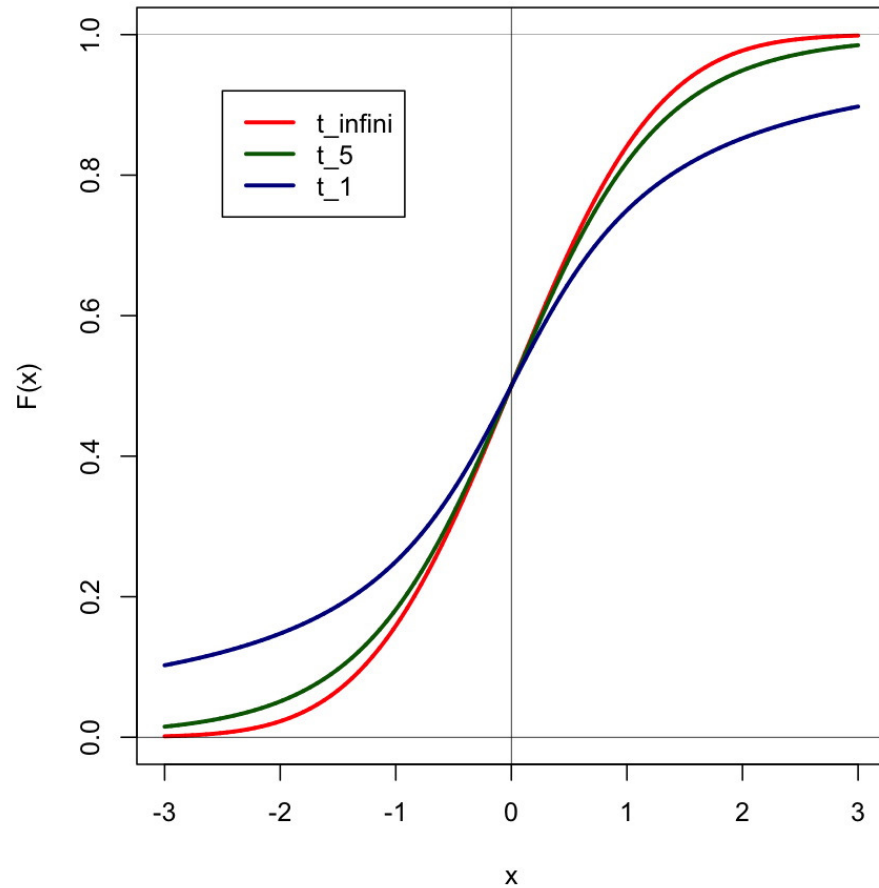
$$E[X] = 0 \quad \text{pour } \nu > 1 \quad \text{et} \quad \text{Var}[X] = \frac{\nu}{\nu - 2} \quad \text{pour } \nu > 2.$$

Ces distributions modélisent des phénomènes où des événements extrêmes se produisent plus souvent que pour la loi normale ("queues lourdes"). Le poids des queues augmente quand  $\nu$  diminue.

Pour  $\nu \leq 1$ , les queues sont si lourdes que  $E[X]$  n'est pas bien défini.

Pour  $\nu \leq 2$ ,  $E[X^2]$  (donc aussi  $\text{Var}[X]$ ) n'est pas bien défini.

# Quelques distributions continues classiques



Fonction de répartition ( $F$ ) et fonction de densité ( $f$ ) de  $X \sim t_{\nu}$ ,  $\nu = \infty, 5, 1$  ( $t_{\infty} = \mathcal{N}(0, 1)$ )

## Quelques distributions continues classiques

(vii)  $X$  est de **distribution de Fisher-Snedecor (F) à  $k_1$  et  $k_2$  degrés de liberté** ( $k_1, k_2 \in \mathbb{N}_0$ ) (notation :  $X \sim F_{k_1, k_2}$ ) si  $X$  admet la fonction de densité définie par

$$f(x) = \begin{cases} \frac{1}{B(\frac{k_1}{2}, \frac{k_2}{2})} x^{\frac{k_1}{2}-1} \sqrt{\frac{(k_1 x)^{k_1} k_2^{k_2}}{(k_1 x + k_2)^{k_1+k_2}}} & \text{si } x \geq 0 \\ 0 & \text{sinon,} \end{cases}$$

où  $(x, y) \mapsto B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt$  est la fonction Beta.

On peut montrer que

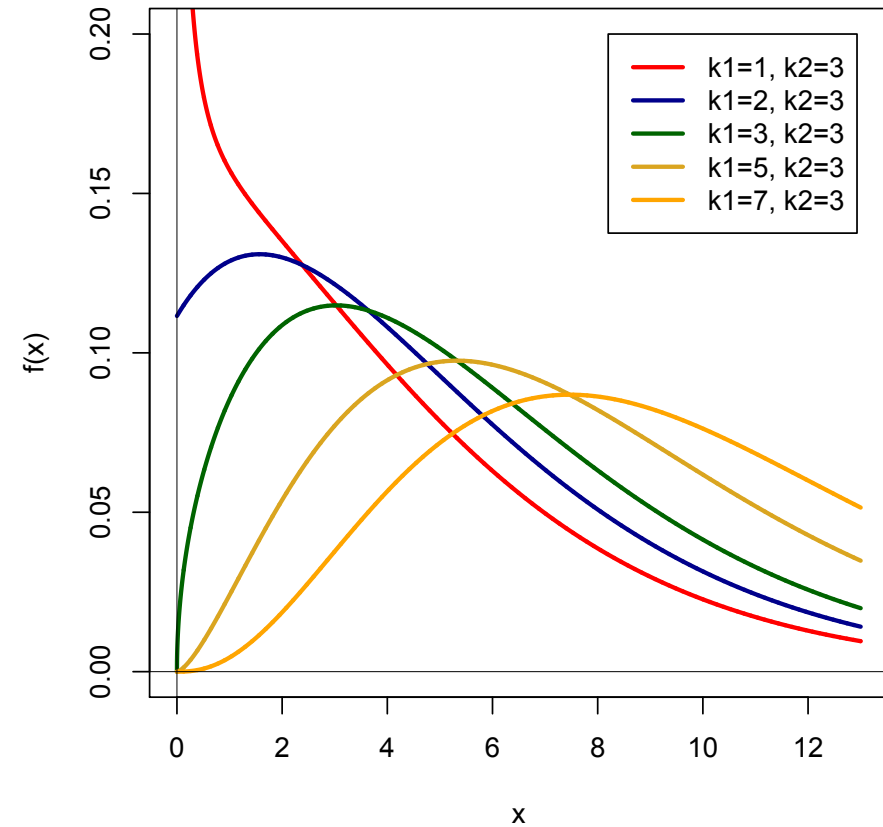
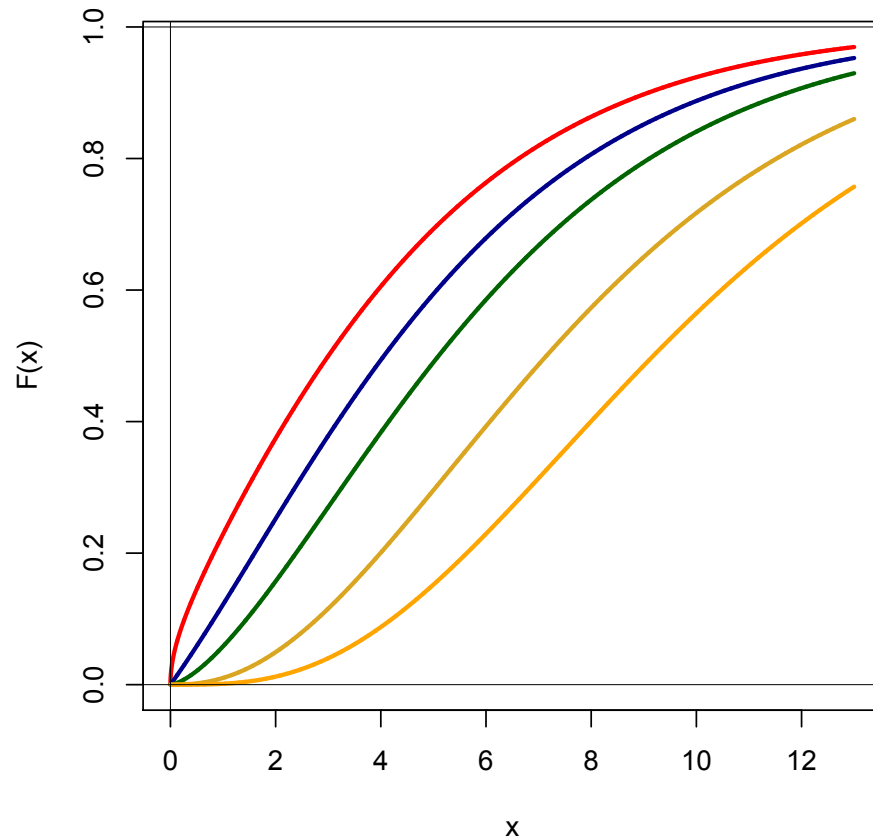
$$E[X] = \frac{k_2}{k_2 - 2} \quad \text{pour } k_2 > 2$$

et

$$\text{Var}[X] = \frac{2k_2^2(k_1 + k_2 - 2)}{k_1(k_2 - 2)^2(k_2 - 4)} \quad \text{pour } k_2 > 4.$$



# Quelques distributions continues classiques



Fonction de répartition ( $F$ ) et fonction de densité ( $f$ ) de  $X \sim F_{k_1, k_2}$ ,  
avec  $k_1 = 1, 2, 3, 5, \text{ et } 7$ , et  $k_2 = 3$  dans chaque cas

# Quelques distributions continues classiques

Les lois  $\chi_k^2$ ,  $t_\nu$ ,  $F_{k_1, k_2}$  admettent les **définitions équivalentes** suivantes, qui apparaissent ici comme des propriétés de ces lois :

- $X \sim \chi_k^2 \Leftrightarrow X \stackrel{\mathcal{D}}{=} \sum_{i=1}^k Z_i^2$ , où les  $Z_i \sim \mathcal{N}(0, 1)$  sont mutuellement indépendantes.
- $X \sim t_\nu \Leftrightarrow X \stackrel{\mathcal{D}}{=} \frac{Z}{\sqrt{Y/\nu}}$ , où  $Z \sim \mathcal{N}(0, 1)$  et  $Y \sim \chi_\nu^2$  sont mutuellement indépendantes.
- $X \sim F_{k_1, k_2} \Leftrightarrow X \stackrel{\mathcal{D}}{=} \frac{Y_1/k_1}{Y_2/k_2}$ , où  $Y_1 \sim \chi_{k_1}^2$  et  $Y_2 \sim \chi_{k_2}^2$  sont mutuellement indépendantes.

La raison pour laquelle nous n'avons pas adopté ces propriétés comme définitions est que la notion d'indépendance de v.a. ne sera définie qu'au ch.3.

Au terme du Q1, ce sont néanmoins les propriétés ci-dessus qui devraient permettre à chacun d'identifier les lois  $\chi_k^2$ ,  $t_\nu$ , ou  $F_{k_1, k_2}$ .

# Quelques distributions continues classiques

Dans le cas continu aussi, nous renvoyons aux monographies donnés en référence pour la présentation d'autres distributions continues classiques, parmi lesquelles

- les lois gamma,
- les lois beta,
- ...

## 2 Variables aléatoires

- Définition et exemples
- Distribution de probabilité
- Opérations, égalité presque sûre et égalité en distribution
- Espérance mathématique
- Variance, risque, inégalité de Tchebychev
- Quelques distributions discrètes classiques
- Quelques distributions continues classiques
- Moments, fonction génératrice des moments

# Moments

Pour mesurer **la position** d'une v.a.  $X$ , on utilise **son espérance**  $\mu_X = E[X]$ .

Pour mesurer **la dispersion** d'une v.a.  $X$ , on utilise **sa variance**  $\sigma_X^2 = E[(X - \mu_X)^2]$ .

Pour mesurer d'**autres caractéristiques**, on peut utiliser d'**autres moments**.

## Définition

Soit  $k \in \{1, 2, \dots\}$  tel que  $E[|X|^k]$  est bien définie dans  $\mathbb{R}$ .

Le **moment non centré d'ordre  $k$**  de  $X$  est  $\mu'_{k,X} = E[X^k]$ .

Le **moment centré d'ordre  $k$**  de  $X$  est  $\mu_{k,X} = E[(X - \mu_X)^k]$ .

On a donc  $\mu'_{1,X} = \mu_X$  et  $\mu_{2,X} = \sigma_X^2$ .

En pratique,

$$\mu'_{k,X} = E[X^k] = \begin{cases} \sum_{i \in \mathcal{I}} (x_i)^k p_i & \text{pour } X \text{ discrète} \\ \int_{-\infty}^{\infty} x^k f(x) dx & \text{pour } X \text{ continue} \end{cases}$$

et

$$\mu_{k,X} = E[(X - \mu_X)^k] = \begin{cases} \sum_{i \in \mathcal{I}} (x_i - \mu_X)^k p_i & \text{pour } X \text{ discrète} \\ \int_{-\infty}^{\infty} (x - \mu_X)^k f(x) dx & \text{pour } X \text{ continue.} \end{cases}$$

Sur la base du moment d'ordre 3, on peut définir une **mesure d'asymétrie**

## Définition

Le **coefficient d'asymétrie** (skewness) de  $X$  est

$$\gamma_{1,X} := \frac{\mu_{3,X}}{(\mu_{2,X})^{3/2}} \quad \left( = \frac{\mu_{3,X}}{\sigma_X^3} \right).$$

Interprétation :

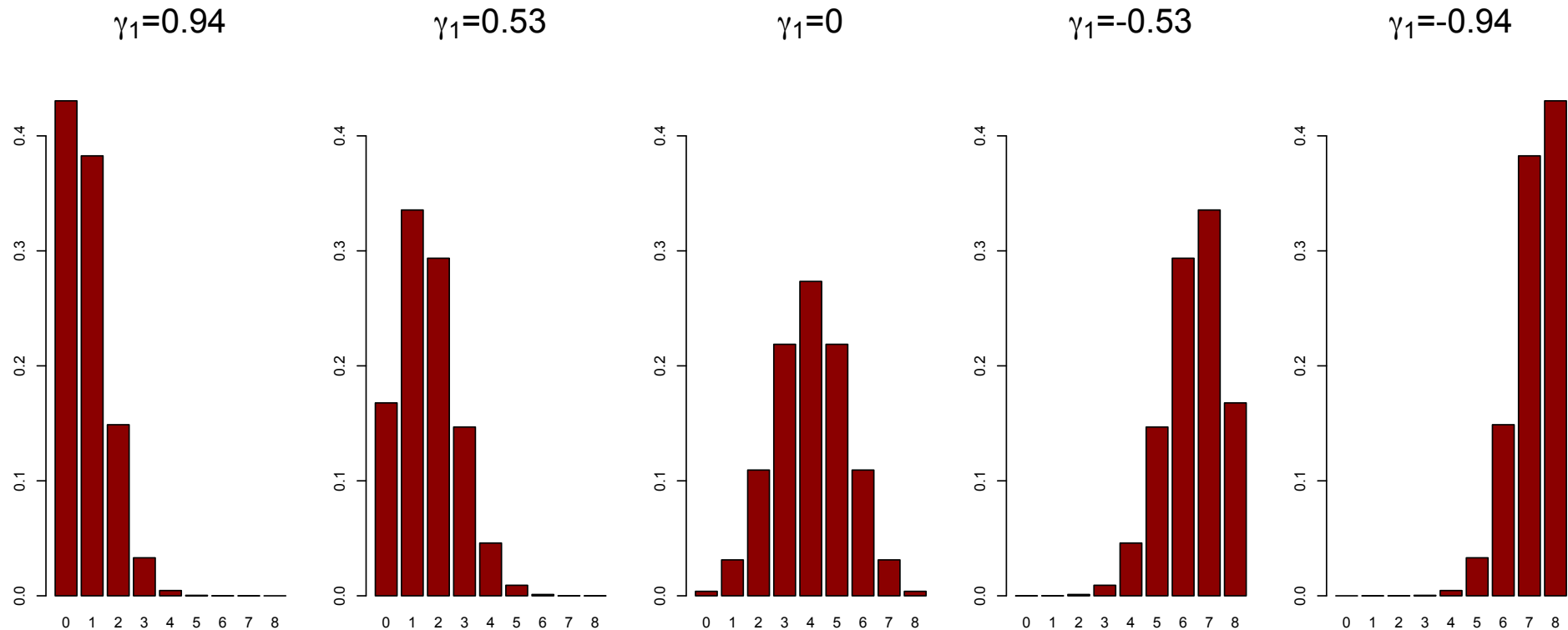
- Si la distribution est symétrique par rapport à une certaine valeur, alors  $\gamma_1 = 0$
- Si la distribution est "asymétrique à droite", alors  $\gamma_1 < 0$
- Si la distribution est "asymétrique à gauche", alors  $\gamma_1 > 0$

(clarification sur la base de la figure suivante)

Exemples classiques : salaires, dépenses,...

Cette mesure d'asymétrie est invariante sous changement de position et d'échelle, dans le sens où  $\gamma_{1,aX+b} = \gamma_{1,X}$  pour tout  $a > 0$  et tout  $b$ .

# Moments



Répartition des probabilités de  $\text{Bin}(n, p)$ , pour  $n = 8$  et  $p = .1, .2, .5, .8, .9$ , avec, dans chaque cas, la valeur numérique de  $\gamma_1 = (1 - 2p) / \sqrt{np(1 - p)}$

Sur la base du moment d'ordre 4, on peut définir une **mesure du poids des queues**

## Définition

Le **coefficient d'aplatissement** (kurtosis) de  $X$  est

$$\gamma_{2,X} := \frac{\mu_{4,X}}{(\mu_{2,X})^2} - 3 \quad \left( = \frac{\mu_{4,X}}{\sigma_X^4} - 3 \right).$$

Interprétation :

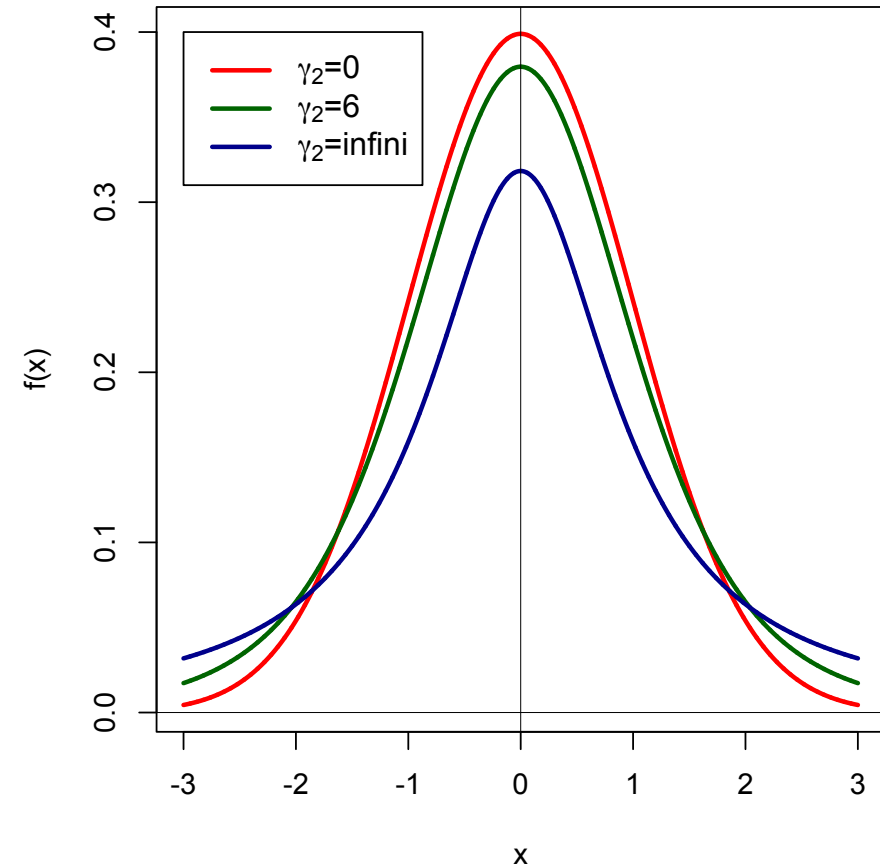
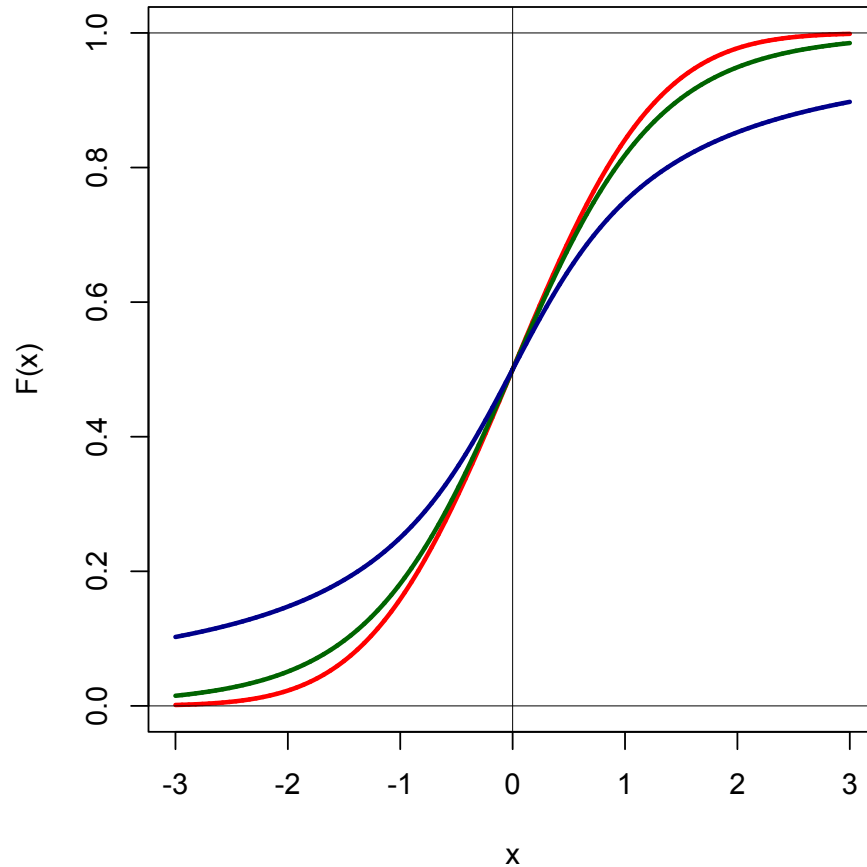
- Si les queues sont "aussi lourdes" que dans le cas normal, alors  $\gamma_2 = 0$
- Si les queues sont "plus lourdes" que dans le cas normal, alors  $\gamma_2 > 0$
- Si les queues sont "plus légères" que dans le cas normal, alors  $\gamma_2 < 0$

Exemple classique : les returns d'actions

Comme le coefficient d'asymétrie, le coefficient d'aplatissement est invariant sous changement de position et d'échelle :  $\gamma_{2,aX+b} = \gamma_{2,X}$  pour tout  $a \neq 0$  et tout  $b$ .



# Moments



Fonction de répartition ( $F$ ) et fonction de densité ( $f$ )  
de  $X \sim t_\infty (= \mathcal{N}(0, 1))$  (rouge), de  $X \sim t_5$  (vert), et de  $X \sim t_1$  (bleu)

Pour pouvoir parler de  $\mu_{k,X}$  et  $\mu'_{k,X}$ , il faut que "X admette des moments d'ordre k".

## Définition

*X admet des moments d'ordre k*  $\Leftrightarrow E[|X|^k]$  est bien définie dans  $\mathbb{R}$

En particulier, si  $X \sim t_1$ , c'est-à-dire si X admet la densité

$$f(x) = \frac{1}{\pi(1+x^2)},$$

alors

$$E[|X|] = \int_{-\infty}^{\infty} |x| f(x) dx = \int_{-\infty}^{\infty} \frac{|x|}{\pi(1+x^2)} dx$$

ne converge pas dans  $\mathbb{R}$ , de sorte qu'on ne peut pas parler de  $\mu_X$ . Plus généralement, si  $X \sim t_\nu$ , alors X n'admet pas de moments d'ordre  $k \geq \nu$ , ce qui est à l'origine des restrictions sur  $\nu$  à la page ch.2-p.78.

---

C'est la présence de queues lourdes qui empêche de considérer certains moments

## Théorème

Supposons que  $X$  admet des moments d'ordre  $k$  (pour un certain  $k \in \mathbb{R}_0^+$ ).

Alors  $X$  admet des moments d'ordre  $\ell$  pour tout  $\ell \in [0, k)$ .

Preuve : fixons  $\ell \in [0, k)$ . Puisque

$$g : \mathbb{R}^+ \rightarrow \mathbb{R}^+ : t \mapsto \frac{t^\ell}{1 + t^k}.$$

est continue sur  $\mathbb{R}^+$  et satisfait  $\lim_{t \rightarrow \infty} g(t) = 0$ ,  $g$  est bornée sur  $\mathbb{R}^+$ . Donc

$$\mathbb{E}[|X|^\ell] = \mathbb{E}\left[\underbrace{\frac{|X|^\ell}{1 + |X|^k}}_{\leq C} (1 + |X|^k)\right] \leq \mathbb{E}[C(1 + |X|^k)] = C + C\mathbb{E}[|X|^k] < \infty,$$

ce qui montre que  $\mathbb{E}[|X|^\ell]$  est bien définie dans  $\mathbb{R}$ . □

---

↪ "Si on peut parler de variance, on peut parler de moyenne" (voir ch.2-p.46)

# Fonction génératrice des moments

## Définition

La *fonction génératrice des moments* d'une v.a.  $X$  est

$$M : \mathbb{R} \mapsto \mathbb{R}$$
$$t \mapsto M(t) := \mathbb{E}[e^{tX}] = \begin{cases} \sum_{i \in \mathcal{I}} e^{tx_i} p_i & \text{si } X \text{ est discrète} \\ \int_{-\infty}^{\infty} e^{tx} f(x) dx & \text{si } X \text{ est continue.} \end{cases}$$

Exemples pour quelques lois classiques :

Loi discrète	$M(t)$	Loi continue	$M(t)$
Bern( $p$ )	$1 - p + pe^t$	Unif( $a, b$ )	$\frac{e^{tb} - e^{ta}}{t(b - a)}$
Bin( $n, p$ )	$(1 - p + pe^t)^n$	Exp( $\lambda$ )	$\frac{1}{1 - \lambda t}$
Poi( $\lambda$ )	$e^{\lambda(e^t - 1)}$	$\mathcal{N}(\mu, \sigma^2)$	$e^{\mu t + \frac{1}{2} \sigma^2 t^2}$
Geom( $p$ )	$\frac{pe^t}{1 - (1 - p)e^t}$	$\chi_k^2$	$(1 - 2t)^{-k/2}, t < \frac{1}{2}$

# Fonction génératrice des moments

Exemple de calcul de  $M(t)$  : si  $X \sim \mathcal{N}(0, 1)$ ,

$$\begin{aligned} M(t) &= \mathbb{E}[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x) dx = \int_{-\infty}^{\infty} e^{tx} \left( \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right) dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx - \frac{x^2}{2} - \frac{t^2}{2} + \frac{t^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-t)^2}{2}} e^{\frac{t^2}{2}} dx = e^{\frac{t^2}{2}}, \end{aligned}$$

où la dernière égalité découle du fait que

$$x \mapsto \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-t)^2}{2}}$$

est la fonction de densité de la loi  $\mathcal{N}(t, 1)$ , et s'intègre donc à 1.

Il arrive que cette fonction ne soit pas définie.

Exemple : si  $X \sim t_\nu$ ,  $M(t)$  n'est bien définie qu'en  $t = 0$ .

---

Comme son nom l'indique, la fonction génératrice des moments permet entre autres de **calculer les moments** de  $X$ .

## Théorème

*Pour tout entier positif  $k$ ,*

$$\frac{d^k M(t)}{dt^k} \Big|_{t=0} = \mu'_{k,X} (= E[X^k])$$

# Fonction génératrice des moments

Soit  $X \sim \text{Bin}(n, p)$ . Comme vu plus haut, on a alors  $M(t) = (1 - p + pe^t)^n$ .

On obtient donc

$$\mathbf{E}[X] = \mu'_1 = \left. \frac{dM(t)}{dt} \right|_{t=0} = [n(1 - p + pe^t)^{n-1} pe^t] \Big|_{t=0} = np.$$

De même,

$$\begin{aligned} \mu'_2 &= \left. \frac{d^2 M(t)}{dt^2} \right|_{t=0} \\ &= [n(n-1)(1 - p + pe^t)^{n-2} p^2 e^{2t} + n(1 - p + pe^t)^{n-1} pe^t] \Big|_{t=0} \\ &= n(n-1)p^2 + np \end{aligned}$$

livre  $\mathbf{Var}[X] = \mu'_2 - (\mu'_1)^2 = n(n-1)p^2 + np - (np)^2 = np(1 - p)$ .

En continuant, on pourrait obtenir l'expression  $\gamma_1 = (1 - 2p)/\sqrt{np(1 - p)}$  donnée en ch.2-p.86.

Nous concluons ce chapitre avec le résultat suivant.

## Théorème

*Soient  $X$  et  $Y$  des v.a. admettant respectivement les fonctions génératrices des moments  $M_X(t)$  et  $M_Y(t)$  dans un voisinage de  $t = 0$ . Alors  **$X$  et  $Y$  sont égales en distribution**  $\Leftrightarrow$  il existe  $\varepsilon > 0$  tel que  $M_X(t) = M_Y(t) \forall t \in (-\varepsilon, \varepsilon)$ .*

Par conséquent, pour établir que deux v.a. ont la même distribution, il suffit de montrer qu'elles partagent la même fonction génératrice des moments (au moins dans un voisinage de 0).

Ceci sera utilisé au chapitre suivant.