

# Modélisation: sujets traités

- Généralités: abstraction, simplification, validation, recommandations.
- Choix des lois de probabilité.
- Différents types de distributions standard univariées et multivariées. Lois empiriques. Dépendence.
- Processus stochastiques: processus de Poisson, Brownien, ...
- Estimation.
- Mesures de performance sur horizon fini et infini. Actualisation.

## Modélisation: principes de base

**Validation:** Le modèle est-t-il assez précis pour l'**application visée**?

**Vérification:** Le programme implante-t-il le modèle correctement?

**Crédibilité:** Les utilisateurs potentiels ont-ils suffisamment confiance dans le modèle pour l'utiliser?

## Modélisation: principes de base

**Validation:** Le modèle est-t-il assez précis pour l'**application visée**?

**Vérification:** Le programme implante-t-il le modèle correctement?

**Crédibilité:** Les utilisateurs potentiels ont-ils suffisamment confiance dans le modèle pour l'utiliser?

La validité est **relative**, jamais absolue.

Elle dépend de l'**objectif** du modèle.

Question numéro 1: À qui servira ce modèle?

Le niveau de détail approprié en dépend.

**Principe:** **minimiser** la complexité du modèle sous la contrainte qu'il donne une assez bonne approximation pour nos besoins.

Règle générale: regarder la forêt avant de regarder les arbres.

Règle générale: regarder la forêt avant de regarder les arbres.

Commencer par un modèle **plus simple**, approximatif, moins coûteux.

Un modèle simple et réaliste lorsqu'il existe) apporte beaucoup: il nous fait comprendre ce qui est le plus important.

Règle générale: regarder la forêt avant de regarder les arbres.

Commencer par un modèle **plus simple**, approximatif, moins coûteux.

Un modèle simple et réaliste lorsqu'il existe) apporte beaucoup: il nous fait comprendre ce qui est le plus important.

Soyons **parcimonieux** dans les hypothèses et les paramètres.

Règle générale: regarder la forêt avant de regarder les arbres.

Commencer par un modèle **plus simple**, approximatif, moins coûteux.

Un modèle simple et réaliste lorsqu'il existe) apporte beaucoup: il nous fait comprendre ce qui est le plus important.

Soyons **parcimonieux** dans les hypothèses et les paramètres.

On raffinerait le modèle seulement lorsque nécessaire (et profitable).

Ne pas construire un modèle trop raffiné compte tenu des données disponibles.

Règle générale: regarder la forêt avant de regarder les arbres.

Commencer par un modèle **plus simple**, approximatif, moins coûteux.

Un modèle simple et réaliste lorsqu'il existe) apporte beaucoup: il nous fait comprendre ce qui est le plus important.

Soyons **parcimonieux** dans les hypothèses et les paramètres.

On raffinerait le modèle seulement lorsque nécessaire (et profitable).

Ne pas construire un modèle trop raffiné compte tenu des données disponibles.

Important: une **bonne connaissance** du système à modéliser et du problème à résoudre.



Règle générale: regarder la forêt avant de regarder les arbres.

Commencer par un modèle **plus simple**, approximatif, moins coûteux.

Un modèle simple et réaliste lorsqu'il existe) apporte beaucoup: il nous fait comprendre ce qui est le plus important.

Soyons **parcimonieux** dans les hypothèses et les paramètres.

On raffinerait le modèle seulement lorsque nécessaire (et profitable).

Ne pas construire un modèle trop raffiné compte tenu des données disponibles.

Important: une **bonne connaissance** du système à modéliser et du problème à résoudre.

Le modèle et le programme doivent être **flexibles**, faciles à **modifier**. Modularité.

La modélisation/simulation est un processus **itératif**.

Règle générale: regarder la forêt avant de regarder les arbres.

Commencer par un modèle **plus simple**, approximatif, moins coûteux.

Un modèle simple et réaliste lorsqu'il existe) apporte beaucoup: il nous fait comprendre ce qui est le plus important.

Soyons **parcimonieux** dans les hypothèses et les paramètres.

On raffinerait le modèle seulement lorsque nécessaire (et profitable).

Ne pas construire un modèle trop raffiné compte tenu des données disponibles.

Important: une **bonne connaissance** du système à modéliser et du problème à résoudre.

Le modèle et le programme doivent être **flexibles**, faciles à **modifier**. Modularité.

La modélisation/simulation est un processus **itératif**.

Et cela demeure en grande partie **un art**: pas de "recette" universelle.

Si le modèle sert à répondre à plusieurs questions, la **validité** se pose pour chacune de ces questions.

Si le modèle sert à répondre à plusieurs questions, la **validité** se pose pour chacune de ces questions.

La validation coûte cher. Plus on veut valider, plus ça coûte cher.  
Aussi, plus on veut un modèle réaliste, plus ça coûte cher.

Si le modèle sert à répondre à plusieurs questions, la **validité** se pose pour chacune de ces questions.

La validation coûte cher. Plus on veut valider, plus ça coûte cher.

Aussi, plus on veut un modèle réaliste, plus ça coûte cher.

On doit faire un compromis entre d'une part les coûts de modélisation et de validation, et d'autre part les coûts associés aux conséquences d'un modèle non valide (ou pas suffisamment réaliste).

# Techniques de validation

## Validation à vue.

- Est-ce que tout semble raisonnable?
- Montrer aux experts ou aux clients. Expérience, intuition.
- Animation graphique.
- Essayez des situations extrêmes.

# Techniques de validation

## Validation à vue.

- Est-ce que tout semble raisonnable?
- Montrer aux experts ou aux clients. Expérience, intuition.
- Animation graphique.
- Essayez des situations extrêmes.

## Tests statistiques des hypothèses.

- Testez les approximations, les lois de probabilité, les hypothèses, etc.
- Comparaisons graphiques, intervalles de confiance, tests d'hypothèses.
- Attention: les tests formels ne sont pas nécessairement appropriés.
- Analyse de sensibilité. Exemple: centre d'appels.

# Techniques de validation

## Validation à vue.

- Est-ce que tout semble raisonnable?
- Montrer aux experts ou aux clients. Expérience, intuition.
- Animation graphique.
- Essayez des situations extrêmes.

## Tests statistiques des hypothèses.

- Testez les approximations, les lois de probabilité, les hypothèses, etc.
- Comparaisons graphiques, intervalles de confiance, tests d'hypothèses.
- Attention: les tests formels ne sont pas nécessairement appropriés.
- Analyse de sensibilité. Exemple: centre d'appels.

## Validation Opérationnelle (comportement du modèle).

- Comparer avec un système existant. Calibration.
- Valider ensuite avec un ensemble indépendant de données.
- Hypothèse à tester: le modèle est valide pour le niveau de précision acceptable, sous les conditions expérimentales.
- Est-ce que le modèle prédit bien l'avenir ? Validation prospective.
- Tests de Turing. Crédibilité.



## Qualité des données

Il faut de l'information **claire** et **fiable** sur le système.

Des données **pertinentes**, **correctes**, et en **quantité** suffisante.

## Qualité des données

Il faut de l'information **claire** et **fiable** sur le système.

Des données **pertinentes**, **correctes**, et en **quantité** suffisante.

On a rarement les données que l'on veut!

**Peu** de données:

- Échantillon trop petit.
- Seulement des **résumés** (moyenne, variance, etc.).
- Pas assez **précis** (e.g., durées des appels téléphoniques...).
- Information **indirecte**, par ex. des entrevues.

## Qualité des données

Il faut de l'information **claire** et **fiable** sur le système.

Des données **pertinentes**, **correctes**, et en **quantité** suffisante.

On a rarement les données que l'on veut!

**Peu** de données:

- Échantillon trop petit.
- Seulement des **résumés** (moyenne, variance, etc.).
- Pas assez **précis** (e.g., durées des appels téléphoniques...).
- Information **indirecte**, par ex. des entrevues.

Des données venant d'une **loi "reliée"** mais pas la bonne.

- Trop d'**aggrégation** (e.g., mensuel au lieu de quotidien).
- Mauvaise **période** (e.g., on a il y a 3 ans, mais on voudrait l'an prochain).
- Mauvais **endroit** (e.g., Toronto au lieu de Montréal).
- Distributions **censurées** (e.g., on veut les demandes, on a les ventes).

## Qualité des données

Il faut de l'information **claire** et **fiable** sur le système.

Des données **pertinentes**, **correctes**, et en **quantité** suffisante.

On a rarement les données que l'on veut!

**Peu** de données:

- Échantillon trop petit.
- Seulement des **résumés** (moyenne, variance, etc.).
- Pas assez **précis** (e.g., durées des appels téléphoniques...).
- Information **indirecte**, par ex. des entrevues.

Des données venant d'une **loi "reliée"** mais pas la bonne.

- Trop d'**aggrégation** (e.g., mensuel au lieu de quotidien).
- Mauvaise **période** (e.g., on a il y a 3 ans, mais on voudrait l'an prochain).
- Mauvais **endroit** (e.g., Toronto au lieu de Montréal).
- Distributions **censurées** (e.g., on veut les demandes, on a les ventes).

Le processus de validation doit inclure la validation des données.

Si les données sont incertaines, faire de l'analyse de sensibilité,  
pas seulement p.r. aux paramètres, mais aussi p.r. aux types de lois.

# Hypothèses simplificatrices

- Logique du modèle.

# Hypothèses simplificatrices

- Logique du modèle.
- Lois de probabilité simplifiées: exponentielle, normale, etc.

# Hypothèses simplificatrices

- Logique du modèle.
- Lois de probabilité simplifiées: exponentielle, normale, etc.
- Aggrégation du temps, des objets, des ressources, etc.

# Hypothèses simplificatrices

- Logique du modèle.
- Lois de probabilité simplifiées: exponentielle, normale, etc.
- Aggrégation du temps, des objets, des ressources, etc.
- Hypothèses d'indépendance et de stationarité.



**Robustesse.**      Analyse de sensibilité.

**Robustesse.** Analyse de sensibilité.

**Exemple:** File  $M/G/1$  à l'état stationnaire.

$\lambda$  = taux d'arrivée,

$\nu, \sigma^2$  = moyenne et variance des durées de service,

$\rho$  =  $\lambda\nu$  = facteur d'utilisation,

$q$  =  $\rho + \rho^2 \frac{(1 + \sigma^2/\nu^2)}{2(1 - \rho)}$  (Pollaczek-Khintchine)

$w$  =  $\frac{q}{\lambda} = \frac{\lambda(\sigma^2 + \nu^2)}{2(1 - \rho)}$

ne dépendent que de  $\lambda$ ,  $\nu$ , et  $\sigma^2$ .

L'erreur si on approxime la file  $M/G/1$  par  $M/M/1$  ne dépend que de  $\sigma/\nu$ .

**Robustesse.** Analyse de sensibilité.

**Exemple:** File  $M/G/1$  à l'état stationnaire.

$\lambda$  = taux d'arrivée,

$\nu, \sigma^2$  = moyenne et variance des durées de service,

$\rho$  =  $\lambda\nu$  = facteur d'utilisation,

$q$  =  $\rho + \rho^2 \frac{(1 + \sigma^2/\nu^2)}{2(1 - \rho)}$  (Pollaczek-Khintchine)

$w$  =  $\frac{q}{\lambda} = \frac{\lambda(\sigma^2 + \nu^2)}{2(1 - \rho)}$

ne dépendent que de  $\lambda$ ,  $\nu$ , et  $\sigma^2$ .

L'erreur si on approxime la file  $M/G/1$  par  $M/M/1$  ne dépend que de  $\sigma/\nu$ .

Mais si les interarrivées ne sont pas exponentielles, ou pour d'autres mesures de performance, ce n'est plus vrai.

On pourra alors utiliser la simulation pour estimer la sensibilité.

# Types de méthodes pour choisir les lois

## A. Paramétrique.

Choisir une loi standard compacte et estimer les paramètres.

Types de lois: expon., normale, gamma, etc.; estimer; test d'ajustement.

Variables dépendantes: lois multivariées.

# Types de méthodes pour choisir les lois

## A. Paramétrique.

Choisir une loi standard compacte et estimer les paramètres.

Types de lois: expon., normale, gamma, etc.; estimer; test d'ajustement.

Variables dépendantes: lois multivariées.

## B. Loi quasi-empirique ou non paramétrique.

Construire une variante de la fonction de répartition empirique et l'utiliser pour générer les variables. Méthodes d'estimation de densité.

# Types de méthodes pour choisir les lois

## A. Paramétrique.

Choisir une loi standard compacte et estimer les paramètres.

Types de lois: expon., normale, gamma, etc.; estimer; test d'ajustement.

Variables dépendantes: lois multivariées.

## B. Loi quasi-empirique ou non paramétrique.

Construire une variante de la fonction de répartition empirique et l'utiliser pour générer les variables. Méthodes d'estimation de densité.

## C. Simulation par retraçage.

Rejouer simplement l'histoire (le "log") d'un système réel, en changeant seulement les décisions et leurs conséquences. Peut être partiel.

# Types de méthodes pour choisir les lois

## A. Paramétrique.

Choisir une loi standard compacte et estimer les paramètres.

Types de lois: expon., normale, gamma, etc.; estimer; test d'ajustement.

Variables dépendantes: lois multivariées.

## B. Loi quasi-empirique ou non paramétrique.

Construire une variante de la fonction de répartition empirique et l'utiliser pour générer les variables. Méthodes d'estimation de densité.

## C. Simulation par retraçage.

Rejouer simplement l'histoire (le "log") d'un système réel, en changeant seulement les décisions et leurs conséquences. Peut être partiel.

Exemples: patients dans un hôpital, United Airlines.

# Types de méthodes pour choisir les lois

## A. Paramétrique.

Choisir une loi standard compacte et estimer les paramètres.

Types de lois: expon., normale, gamma, etc.; estimer; test d'ajustement.

Variables dépendantes: lois multivariées.

## B. Loi quasi-empirique ou non paramétrique.

Construire une variante de la fonction de répartition empirique et l'utiliser pour générer les variables. Méthodes d'estimation de densité.

## C. Simulation par retraçage.

Rejouer simplement l'histoire (le "log") d'un système réel, en changeant seulement les décisions et leurs conséquences. Peut être partiel.

Exemples: patients dans un hôpital, United Airlines.

Gros défaut: on ne verra jamais d'autres combinaisons que ce qui se trouve dans les données.

Solution: il faut modéliser les dépendences complexes! Difficile...



## Recommendations:

- A s'il y a une certaine justification pour une loi particulière  
(important de comprendre les principes sous-jacents aux différentes lois);
- B si on a vraiment beaucoup de données;
- C peut être utile s'il y a vraiment beaucoup de données et des dépendances trop difficiles à modéliser.

## Recommendations:

- A s'il y a une certaine justification pour une loi particulière  
(important de comprendre les principes sous-jacents aux différentes lois);
- B si on a vraiment beaucoup de données;
- C peut être utile s'il y a vraiment beaucoup de données et des dépendances trop difficiles à modéliser.

Arguments favorisant l'approche paramétrique A:

- Si on a des raisons “physiques” de choisir une loi particulière.

## Recommendations:

- A s'il y a une certaine justification pour une loi particulière  
(important de comprendre les principes sous-jacents aux différentes lois);
- B si on a vraiment beaucoup de données;
- C peut être utile s'il y a vraiment beaucoup de données et des dépendances trop difficiles à modéliser.

Arguments favorisant l'approche paramétrique A:

- Si on a des raisons “physiques” de choisir une loi particulière.
- C'est une façon compacte de représenter les données.

## Recommendations:

- A s'il y a une certaine justification pour une loi particulière  
(important de comprendre les principes sous-jacents aux différentes lois);
- B si on a vraiment beaucoup de données;
- C peut être utile s'il y a vraiment beaucoup de données et des dépendances trop difficiles à modéliser.

Arguments favorisant l'approche paramétrique A:

- Si on a des raisons “physiques” de choisir une loi particulière.
- C'est une façon compacte de représenter les données.
- Procédures disponibles pour générer les v.a. (moins de programmation).

## Recommendations:

- A s'il y a une certaine justification pour une loi particulière  
(important de comprendre les principes sous-jacents aux différentes lois);
- B si on a vraiment beaucoup de données;
- C peut être utile s'il y a vraiment beaucoup de données et des dépendances trop difficiles à modéliser.

Arguments **favorisant** l'approche paramétrique **A**:

- Si on a des raisons “physiques” de choisir une loi particulière.
- C'est une façon compacte de représenter les données.
- Procédures disponibles pour générer les v.a. (moins de programmation).
- Les valeurs que l'on peut générer ne sont pas limitées à l'étendue de l'échantillon (important si on veut pouvoir générer des événements “extrêmes”).  
Cela “lisse” (régularise) les données.

## Recommendations:

- A s'il y a une certaine justification pour une loi particulière  
(important de comprendre les principes sous-jacents aux différentes lois);
- B si on a vraiment beaucoup de données;
- C peut être utile s'il y a vraiment beaucoup de données et des dépendances trop difficiles à modéliser.

### Arguments favorisant l'approche paramétrique A:

- Si on a des raisons “physiques” de choisir une loi particulière.
- C'est une façon compacte de représenter les données.
- Procédures disponibles pour générer les v.a. (moins de programmation).
- Les valeurs que l'on peut générer ne sont pas limitées à l'étendue de l'échantillon (important si on veut pouvoir générer des événements “extrêmes”).  
Cela “lisse” (régularise) les données.

### Arguments contre A:

- Parfois très difficile, voire impossible, de connaître le bon type de loi.  
Parfois, les données ne s'ajustent à aucune des lois disponibles.

## Recommendations:

- A s'il y a une certaine justification pour une loi particulière  
(important de comprendre les principes sous-jacents aux différentes lois);
- B si on a vraiment beaucoup de données;
- C peut être utile s'il y a vraiment beaucoup de données et des dépendances trop difficiles à modéliser.

Arguments **favorisant** l'approche paramétrique **A**:

- Si on a des raisons “physiques” de choisir une loi particulière.
- C'est une façon compacte de représenter les données.
- Procédures disponibles pour générer les v.a. (moins de programmation).
- Les valeurs que l'on peut générer ne sont pas limitées à l'étendue de l'échantillon (important si on veut pouvoir générer des événements “extrêmes”).  
Cela “lisse” (régularise) les données.

Arguments **contre** **A**:

- Parfois très difficile, voire impossible, de connaître le bon type de loi.  
Parfois, les données ne s'ajustent à aucune des lois disponibles.
- L'estimation des paramètres n'est pas toujours robuste.

## Recommendations:

- A s'il y a une certaine justification pour une loi particulière  
(important de comprendre les principes sous-jacents aux différentes lois);
- B si on a vraiment beaucoup de données;
- C peut être utile s'il y a vraiment beaucoup de données et des dépendances trop difficiles à modéliser.

### Arguments favorisant l'approche paramétrique A:

- Si on a des raisons “physiques” de choisir une loi particulière.
- C'est une façon compacte de représenter les données.
- Procédures disponibles pour générer les v.a. (moins de programmation).
- Les valeurs que l'on peut générer ne sont pas limitées à l'étendue de l'échantillon (important si on veut pouvoir générer des événements “extrêmes”).  
Cela “lisse” (régularise) les données.

### Arguments contre A:

- Parfois très difficile, voire impossible, de connaître le bon type de loi.  
Parfois, les données ne s'ajustent à aucune des lois disponibles.
- L'estimation des paramètres n'est pas toujours robuste.
- Lors de l'ajustement, il y a perte ou distorsion d'information.



## Recommendations:

- A s'il y a une certaine justification pour une loi particulière  
(important de comprendre les principes sous-jacents aux différentes lois);
- B si on a vraiment beaucoup de données;
- C peut être utile s'il y a vraiment beaucoup de données et des dépendances trop difficiles à modéliser.

Arguments **favorisant** l'approche paramétrique **A**:

- Si on a des raisons “physiques” de choisir une loi particulière.
- C'est une façon compacte de représenter les données.
- Procédures disponibles pour générer les v.a. (moins de programmation).
- Les valeurs que l'on peut générer ne sont pas limitées à l'étendue de l'échantillon (important si on veut pouvoir générer des événements “extrêmes”).  
Cela “lisse” (régularise) les données.

Arguments **contre** **A**:

- Parfois très difficile, voire impossible, de connaître le bon type de loi.  
Parfois, les données ne s'ajustent à aucune des lois disponibles.
- L'estimation des paramètres n'est pas toujours robuste.
- Lors de l'ajustement, il y a perte ou distorsion d'information.
- Pour certaines lois, la génération de v.a. est difficile ou lente.

## **L'approche paramétrique**

Recommandée seulement si l'ajustement est bon, et sur la base de justifications théoriques.

## L'approche paramétrique

Recommandée seulement si l'ajustement est bon, et sur la base de justifications théoriques.

[Encyclopedie](#) des lois de probabilité: Johnson and Kotz (1969–1972), Johnson, Kotz, et Balakrishnan (1995). Aussi très bien sur Wikipedia.

## L'approche paramétrique

Recommandée seulement si l'ajustement est bon, et sur la base de justifications théoriques.

[Encyclopedie](#) des lois de probabilité: Johnson and Kotz (1969–1972), Johnson, Kotz, et Balakrishnan (1995). Aussi très bien sur Wikipedia.

### Types de paramètres:

[Localisation](#): spécifie l'origine sur l'axe horizontal.

[Échelle](#): étire ou contracte l'axe horizontal, détermine l'échelle sans changer la forme de la courbe.

[Forme](#): modifie plus en profondeur la forme de la fonction de densité ou de masse, et les propriétés de la loi.

Paramètres (et coefficients) d'[asymétrie](#) et d'[aplatissement](#).

## L'approche paramétrique

Recommandée seulement si l'ajustement est bon, et sur la base de justifications théoriques.

[Encyclopedie](#) des lois de probabilité: Johnson and Kotz (1969–1972), Johnson, Kotz, et Balakrishnan (1995). Aussi très bien sur Wikipedia.

### Types de paramètres:

[Localisation](#): spécifie l'origine sur l'axe horizontal.

[Échelle](#): étire ou contracte l'axe horizontal, détermine l'échelle sans changer la forme de la courbe.

[Forme](#): modifie plus en profondeur la forme de la fonction de densité ou de masse, et les propriétés de la loi.

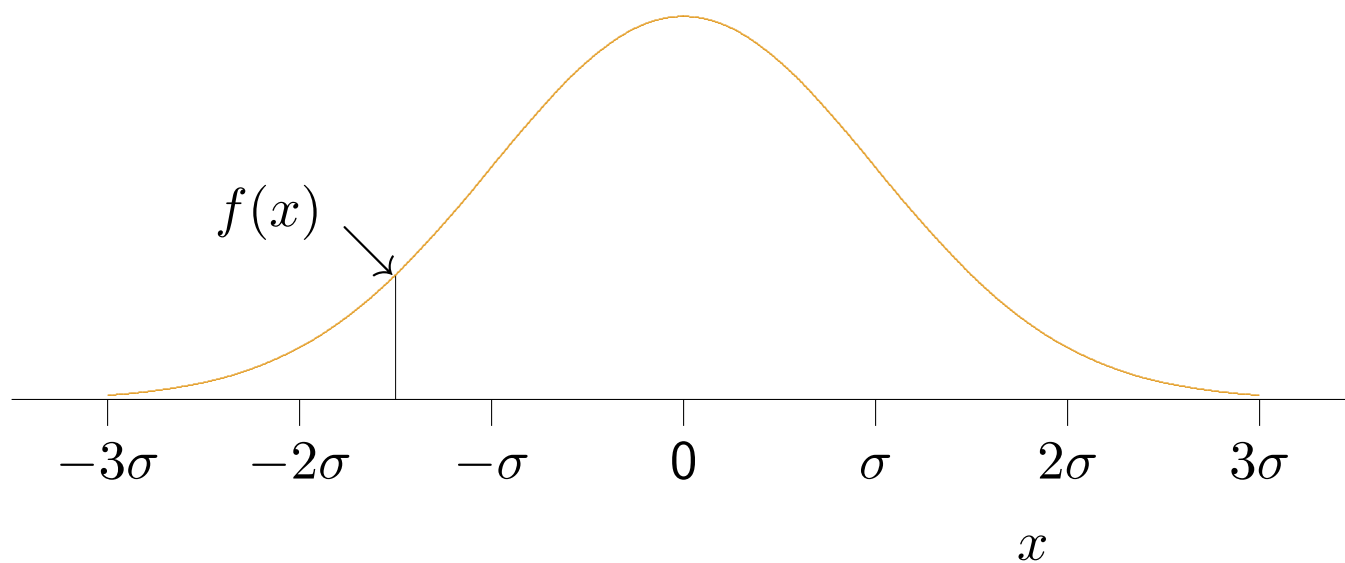
Paramètres (et coefficients) d'[asymétrie](#) et d'[aplatissement](#).

Une [transformation affine](#) ( $Y = aX + b$ ) permet de contrôler la localisation et l'échelle. Il suffit d'étudier une version standardisée.

**Exemple:** La loi normale.

$$X \sim N(\mu, \sigma^2) \text{ ssi } Z = (X - \mu)/\sigma \sim N(0, 1).$$

$$X = \sigma Z + \mu.$$



## Quelques caractérisations des distributions.

Soit  $X$  une v.a. univariée.

Moyenne  $\mu = \mathbb{E}[X]$ ;

Variance  $\sigma^2 = \mathbb{E}[(X - \mu)^2]$ ;      Écart-type  $\sigma$ ;

## Quelques caractérisations des distributions.

Soit  $X$  une v.a. univariée.

Moyenne  $\mu = \mathbb{E}[X]$ ;

Variance  $\sigma^2 = \mathbb{E}[(X - \mu)^2]$ ;      Écart-type  $\sigma$ ;

Coefficient de variation  $\sigma/\mu$ ;



## Quelques caractérisations des distributions.

Soit  $X$  une v.a. univariée.

Moyenne  $\mu = \mathbb{E}[X]$ ;

Variance  $\sigma^2 = \mathbb{E}[(X - \mu)^2]$ ; Écart-type  $\sigma$ ;

Coefficient de variation  $\sigma/\mu$ ;

Coefficient d'asymétrie  $\nu = \mathbb{E}[(X - \mu)^3]/\sigma^3$

## Quelques caractérisations des distributions.

Soit  $X$  une v.a. univariée.

Moyenne  $\mu = \mathbb{E}[X]$ ;

Variance  $\sigma^2 = \mathbb{E}[(X - \mu)^2]$ ;      Écart-type  $\sigma$ ;

Coefficient de variation  $\sigma/\mu$ ;

Coefficient d'asymétrie  $\nu = \mathbb{E}[(X - \mu)^3]/\sigma^3$

Coefficient d'aplatissement  $\kappa = \mathbb{E}[(X - \mu)^4]/\sigma^4$ .

## Quelques caractérisations des distributions.

Soit  $X$  une v.a. univariée.

Moyenne  $\mu = \mathbb{E}[X]$ ;

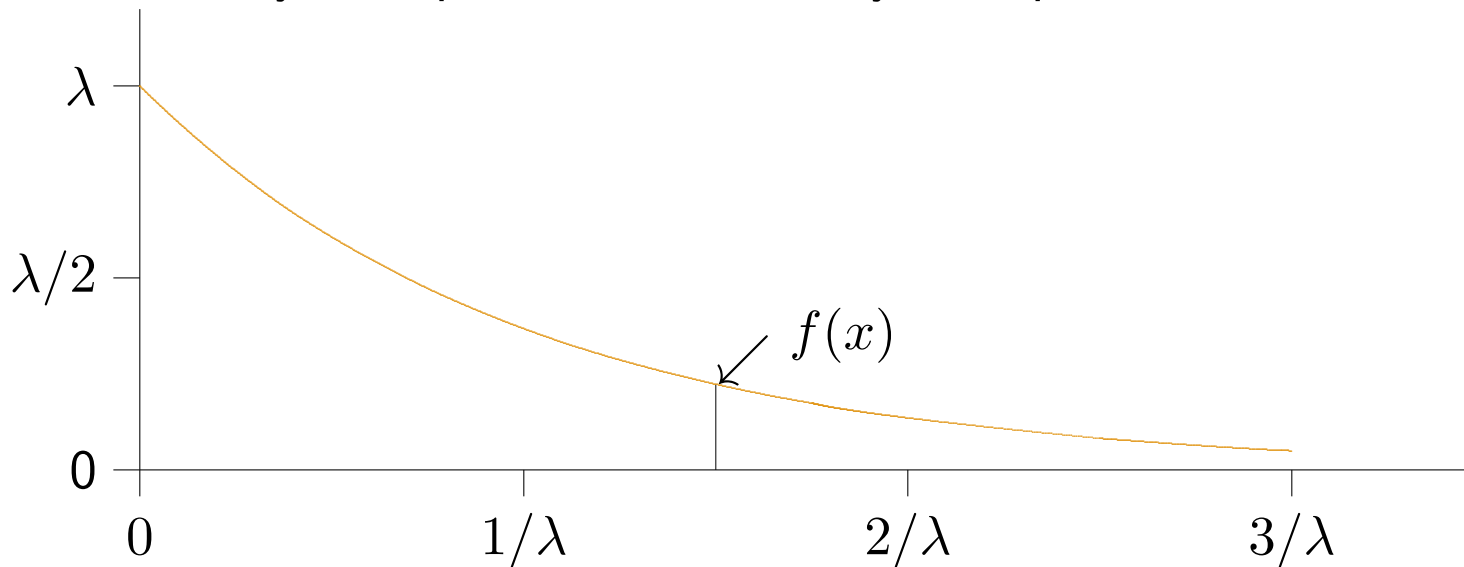
Variance  $\sigma^2 = \mathbb{E}[(X - \mu)^2]$ ; Écart-type  $\sigma$ ;

Coefficient de variation  $\sigma/\mu$ ;

Coefficient d'asymétrie  $\nu = \mathbb{E}[(X - \mu)^3]/\sigma^3$

Coefficient d'aplatissement  $\kappa = \mathbb{E}[(X - \mu)^4]/\sigma^4$ .

$\nu = 0$ : loi symétrique.       $\nu > 0$ : asymétrique à droite.



## Quelques caractérisations des distributions.

Soit  $X$  une v.a. univariée.

Moyenne  $\mu = \mathbb{E}[X]$ ;

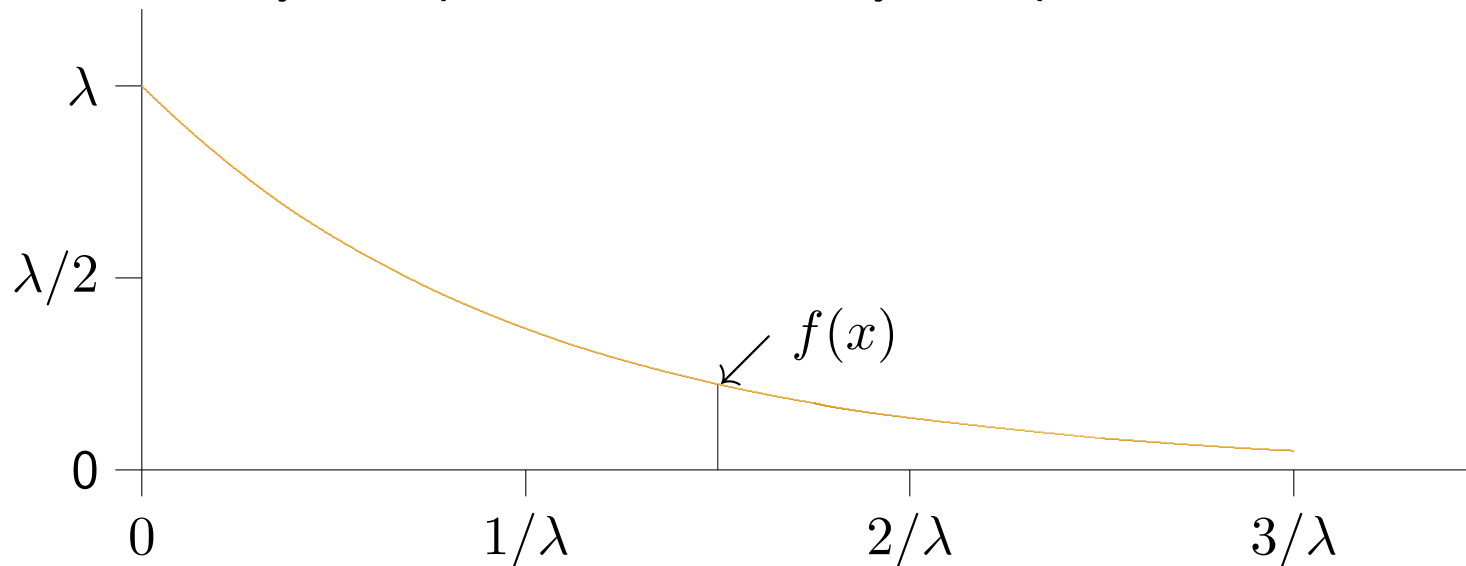
Variance  $\sigma^2 = \mathbb{E}[(X - \mu)^2]$ ;      Écart-type  $\sigma$ ;

Coefficient de variation  $\sigma/\mu$ ;

Coefficient d'asymétrie  $\nu = \mathbb{E}[(X - \mu)^3]/\sigma^3$

Coefficient d'aplatissement  $\kappa = \mathbb{E}[(X - \mu)^4]/\sigma^4$ .

$\nu = 0$ : loi symétrique.       $\nu > 0$ : asymétrique à droite.



Pour certaines familles de loi (e.g., Johnson), on peut choisir les paramètres pour obtenir des valeurs arbitraires de ces 4 coefficients.

Fonction de **répartition**:

$$F(x) = \mathbb{P}[X \leq x] = \begin{cases} \sum_{j=0}^x p_j & (\text{cas discret}), \\ \int_{-\infty}^x f(y) dy & (\text{cas continu}). \end{cases}$$

Fonction de **répartition**:

$$F(x) = \mathbb{P}[X \leq x] = \begin{cases} \sum_{j=0}^x p_j & (\text{cas discret}), \\ \int_{-\infty}^x f(y) dy & (\text{cas continu}). \end{cases}$$

Fonction de **survie**:

$$\bar{F}(x) = \mathbb{P}[X > x].$$

Fonction de **répartition**:

$$F(x) = \mathbb{P}[X \leq x] = \begin{cases} \sum_{j=0}^x p_j & (\text{cas discret}), \\ \int_{-\infty}^x f(y) dy & (\text{cas continu}). \end{cases}$$

Fonction de **survie**:

$$\bar{F}(x) = \mathbb{P}[X > x].$$

**Taux de panne**:

$$r(x) = \frac{f(x)}{1 - F(x)}$$

On a

$$\mathbb{P}[X < x + \epsilon | X > x] = \frac{\mathbb{P}[x < X < x + \epsilon]}{\mathbb{P}[X > x]} \approx \frac{\epsilon f(x)}{(1 - F(x))} = \epsilon r(x).$$

- Taux de panne **croissant** vs **décroissant**.
- Taux de panne en **baignoire**.

## Quelques lois de probabilité discrètes

Loi uniforme discrète  $(i, j)$

$$\mathbb{P}[X = x] = 1/(j - i + 1) \quad \text{pour } x = i, i + 1, \dots, j.$$



## Quelques lois de probabilité discrètes

### Loi uniforme discrète $(i, j)$

$$\mathbb{P}[X = x] = 1/(j - i + 1) \quad \text{pour } x = i, i + 1, \dots, j.$$

### Loi binomiale $(n, p)$

$$\mathbb{P}[X = x] = \binom{n}{x} p^x (1 - p)^{1-x} \quad \text{pour } x = 0, \dots, n.$$

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

## Loi géométrique ( $p$ )

$$\mathbb{P}[X = x] = p(1 - p)^x \quad \text{pour } x = 0, 1, \dots$$

Elle est **sans mémoire**:  $\mathbb{P}[X = y + x | X \geq y] = \mathbb{P}[X = x]$ .

## Loi géométrique ( $p$ )

$$\mathbb{P}[X = x] = p(1 - p)^x \quad \text{pour } x = 0, 1, \dots$$

Elle est **sans mémoire**:  $\mathbb{P}[X = y + x | X \geq y] = \mathbb{P}[X = x]$ .

## Loi binomiale négative ( $n, p$ )

$$\mathbb{P}[X = x] = \frac{\Gamma(n + x)}{\Gamma(n)\Gamma(x + 1)} p^n (1 - p)^x \quad \text{pour } x = 0, 1, \dots$$

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

Pour  $n = 1$ : loi géométrique.

Pour  $n$  entier:  $X$  = nombre d'échecs avant  $n$  succès.

## Loi de Poisson ( $\lambda$ )

$$\mathbb{P}[X = x] = \frac{\lambda^x e^{-\lambda}}{x!} \quad \text{pour } x = 0, 1, 2, \dots$$

$$\mathbb{E}[X] = \text{Var}[X] = \lambda.$$

## Loi de Poisson ( $\lambda$ )

$$\mathbb{P}[X = x] = \frac{\lambda^x e^{-\lambda}}{x!} \quad \text{pour } x = 0, 1, 2, \dots$$

$$\mathbb{E}[X] = \text{Var}[X] = \lambda.$$

Si  $X_1, \dots, X_q$  indep.,  $X_i \sim \text{Poisson}(\lambda_i)$ , alors  $X = X_1 + \dots + X_q \sim \text{Poisson}(\lambda)$   
où  $\lambda = \lambda_1 + \dots + \lambda_q$ .

## Loi de Poisson ( $\lambda$ )

$$\mathbb{P}[X = x] = \frac{\lambda^x e^{-\lambda}}{x!} \quad \text{pour } x = 0, 1, 2, \dots$$

$$\mathbb{E}[X] = \text{Var}[X] = \lambda.$$

Si  $X_1, \dots, X_q$  indep.,  $X_i \sim \text{Poisson}(\lambda_i)$ , alors  $X = X_1 + \dots + X_q \sim \text{Poisson}(\lambda)$   
où  $\lambda = \lambda_1 + \dots + \lambda_q$ .

Si  $X \sim \text{Binomiale}(n, p)$  et  $(n, np) \rightarrow (\infty, \lambda)$ , alors  $X \Rightarrow \text{Poisson}(\lambda)$ .

## Loi de Poisson ( $\lambda$ )

$$\mathbb{P}[X = x] = \frac{\lambda^x e^{-\lambda}}{x!} \quad \text{pour } x = 0, 1, 2, \dots$$

$$\mathbb{E}[X] = \text{Var}[X] = \lambda.$$

Si  $X_1, \dots, X_q$  indep.,  $X_i \sim \text{Poisson}(\lambda_i)$ , alors  $X = X_1 + \dots + X_q \sim \text{Poisson}(\lambda)$   
où  $\lambda = \lambda_1 + \dots + \lambda_q$ .

Si  $X \sim \text{Binomiale}(n, p)$  et  $(n, np) \rightarrow (\infty, \lambda)$ , alors  $X \Rightarrow \text{Poisson}(\lambda)$ .

**Théorème.** (Bornes explicites.) Soient  $X_1, \dots, X_n$  indép.,  
 $X_j \sim \text{Binomiale}(1, p_j)$ ,  $X = X_1 + \dots + X_n$  et  $\lambda = p_1 + \dots + p_n$ .  
Alors, pour tout  $x$ ,

$$\left| \mathbb{P}[X = x] - \frac{\lambda^x e^{-\lambda}}{x!} \right| \leq \sum_{j=1}^n p_j^2 \quad [= \lambda/n \text{ si } p_j = p \text{ pour tout } j].$$

## Loi de Poisson ( $\lambda$ )

$$\mathbb{P}[X = x] = \frac{\lambda^x e^{-\lambda}}{x!} \quad \text{pour } x = 0, 1, 2, \dots$$

$$\mathbb{E}[X] = \text{Var}[X] = \lambda.$$

Si  $X_1, \dots, X_q$  indep.,  $X_i \sim \text{Poisson}(\lambda_i)$ , alors  $X = X_1 + \dots + X_q \sim \text{Poisson}(\lambda)$   
où  $\lambda = \lambda_1 + \dots + \lambda_q$ .

Si  $X \sim \text{Binomiale}(n, p)$  et  $(n, np) \rightarrow (\infty, \lambda)$ , alors  $X \Rightarrow \text{Poisson}(\lambda)$ .

**Théorème.** (Bornes explicites.) Soient  $X_1, \dots, X_n$  indép.,  
 $X_j \sim \text{Binomiale}(1, p_j)$ ,  $X = X_1 + \dots + X_n$  et  $\lambda = p_1 + \dots + p_n$ .  
Alors, pour tout  $x$ ,

$$\left| \mathbb{P}[X = x] - \frac{\lambda^x e^{-\lambda}}{x!} \right| \leq \sum_{j=1}^n p_j^2 \quad [= \lambda/n \text{ si } p_j = p \text{ pour tout } j].$$

Exemples: nombre d'erreurs typographiques dans un livre, nombre de personnes achetant des actions de Bombardier un jour donné, etc.



## Loi de Zipf ( $\alpha$ )

$$\mathbb{P}[X = x] = x^{-\alpha} / \zeta(\alpha) \quad \text{pour } x = 1, 2, \dots$$

$$\text{où } \zeta(\alpha) = \sum_{x=1}^{\infty} x^{-\alpha}.$$

Particularité et intérêt: décroissance non exponentielle.

## Loi de Zipf ( $\alpha$ )

$$\mathbb{P}[X = x] = x^{-\alpha} / \zeta(\alpha) \quad \text{pour } x = 1, 2, \dots$$

où  $\zeta(\alpha) = \sum_{x=1}^{\infty} x^{-\alpha}$ .

Particularité et intérêt: décroissance non exponentielle.

## Loi multinomiale

## Loi de Zipf ( $\alpha$ )

$$\mathbb{P}[X = x] = x^{-\alpha} / \zeta(\alpha) \quad \text{pour } x = 1, 2, \dots$$

où  $\zeta(\alpha) = \sum_{x=1}^{\infty} x^{-\alpha}$ .

Particularité et intérêt: décroissance non exponentielle.

## Loi multinomiale

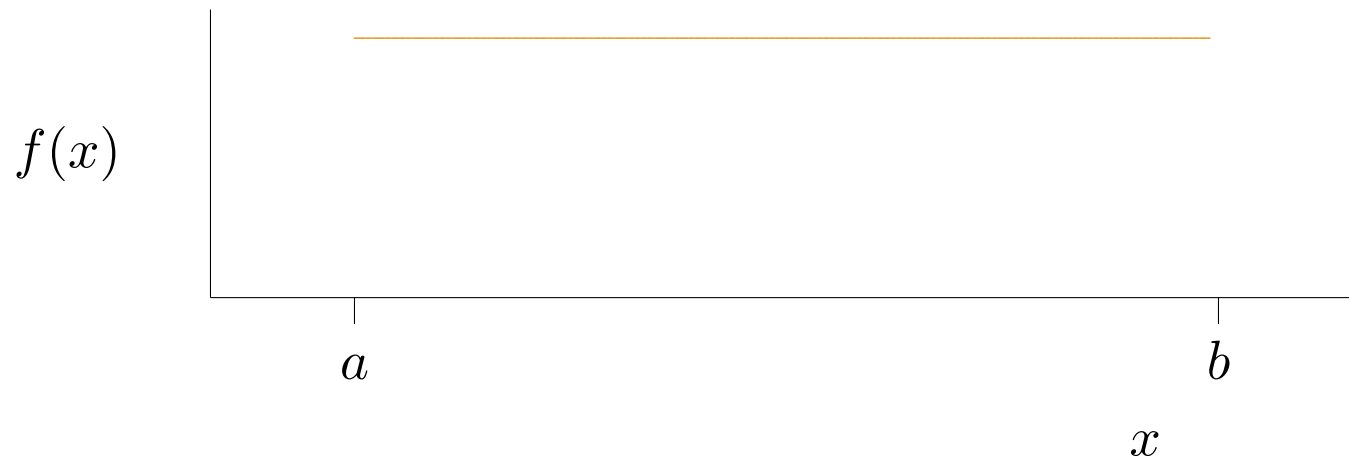
## Loi multinomiale négative

# Lois continues univariées

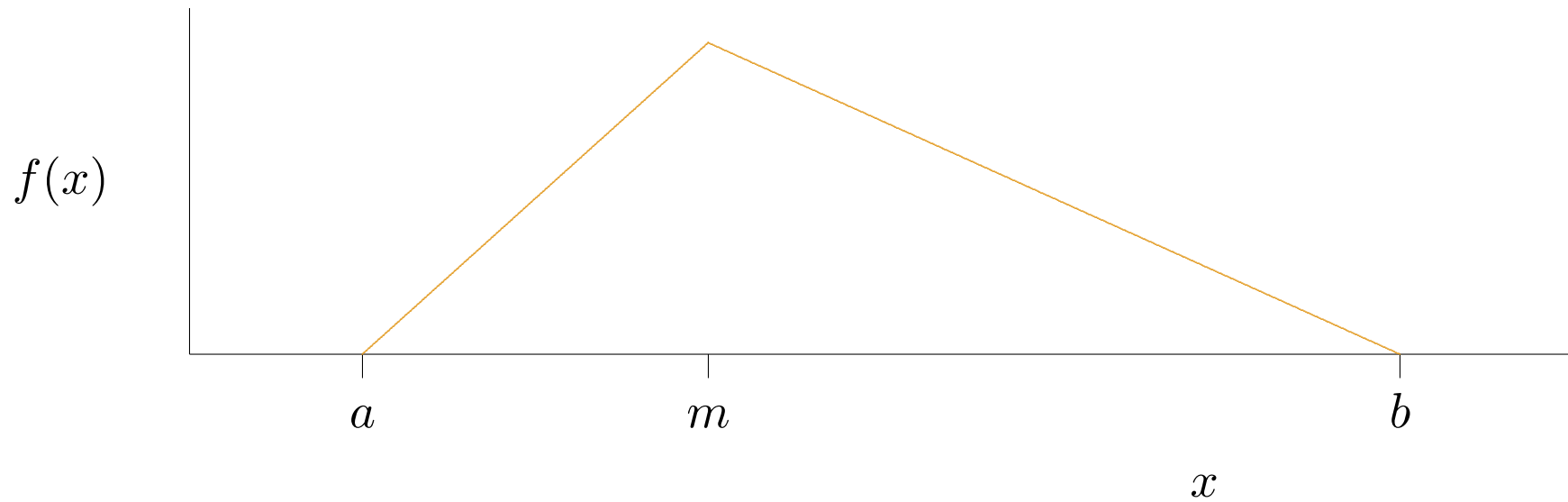
## Loi uniforme ( $a, b$ )

$$f(x) = 1/(b - a) \quad \text{pour } a < x < b.$$

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



## Loi triangulaire ( $a, m, b$ )

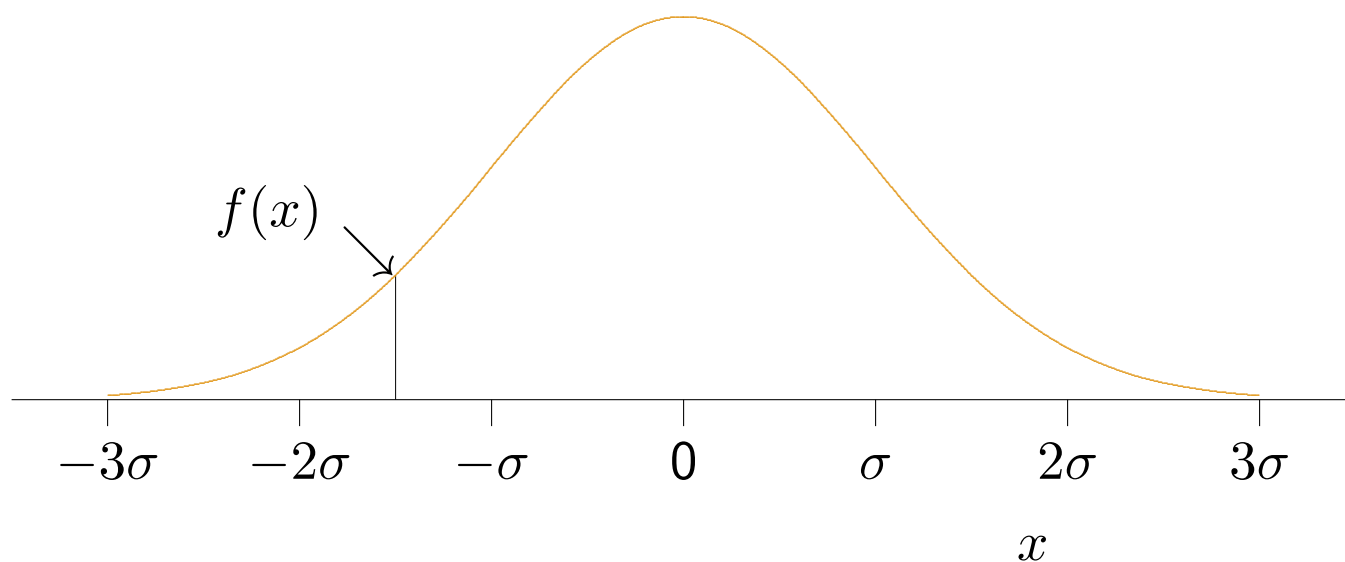


Souvent utilisée en l'absence de données...

**Loi normale**  $(\mu, \sigma^2)$

$$X \sim N(\mu, \sigma^2).$$

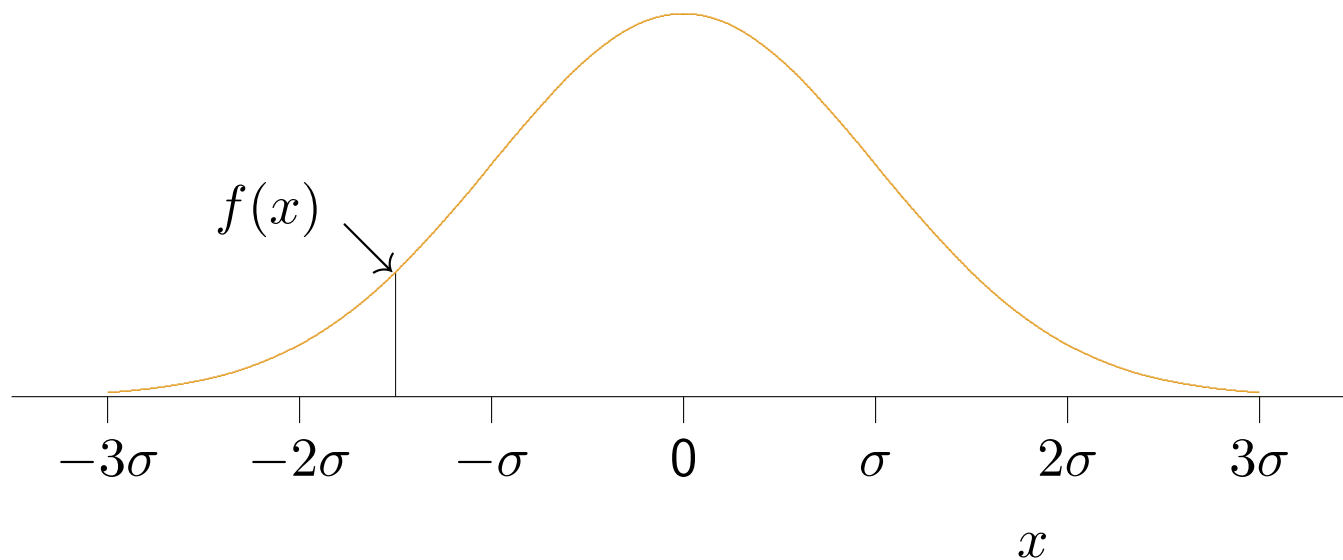
Densité:  $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\{-(x - \mu)^2 / (2\sigma^2)\}$  pour  $x \in \mathbb{R}$ .



**Loi normale**  $(\mu, \sigma^2)$

$$X \sim N(\mu, \sigma^2).$$

Densité:  $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\{-(x - \mu)^2/(2\sigma^2)\}$  pour  $x \in \mathbb{R}$ .

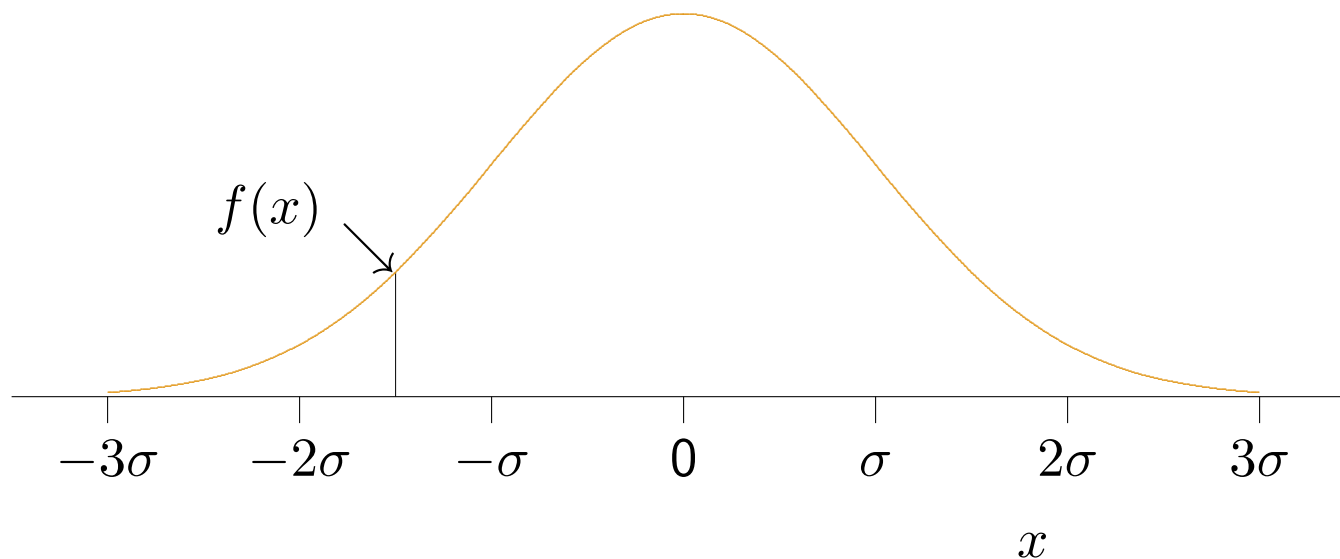


Si  $X \sim N(\mu, \sigma^2)$  alors  $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$ .

**Loi normale**  $(\mu, \sigma^2)$

$$X \sim N(\mu, \sigma^2).$$

Densité:  $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\{-(x - \mu)^2/(2\sigma^2)\}$  pour  $x \in \mathbb{R}$ .



Si  $X \sim N(\mu, \sigma^2)$  alors  $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$ .

$$\Phi(x) = \mathbb{P}[Z \leq x] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-s^2/2} ds.$$



Motivation: **théorèmes de la limite centrale** (TLC).

$X \approx$  normale si  $X$  est une somme de nombreux petits effets indépendants.

Plusieurs versions:  $X_i$  de lois différentes, dépendance, TLCs multivariés, TLC fonctionnels, etc.

Motivation: **théorèmes de la limite centrale** (TLC).

$X \approx$  normale si  $X$  est une somme de nombreux petits effets indépendants.

Plusieurs versions:  $X_i$  de lois différentes, dépendance, TLCs multivariés, TLC fonctionnels, etc.

Par exemple: **Théorème de Berry-Essen.**

Soient  $X_1, X_2, \dots$  indep.,  $\mathbb{E}[X_i] = \mu_i$  et  $\text{Var}[X_i] = \sigma_i^2 < \infty$ .

Posons  $s_n^2 = \sigma_1^2 + \dots + \sigma_n^2$ ,

$$Y_n = \frac{(X_1 - \mu_1) + \dots + (X_n - \mu_n)}{s_n},$$

et  $F_n(x) = \mathbb{P}[Y_n \leq x]$ . Alors,  $\mathbb{E}[Y_n] = 0$ ,  $\text{Var}[Y_n] = 1$ , et

$$\sup_{n \geq 1, x \in \mathbb{R}} |F_n(x) - \Phi(x)| \leq \kappa \frac{\mathbb{E}(|X_1 - \mu_1|^3) + \dots + \mathbb{E}(|X_n - \mu_n|^3)}{s_n^3}$$

où  $\kappa = 3$  si les  $X_i$  sont i.i.d. et  $\kappa = 6$  sinon.

La borne sur l'erreur dépend de l'**asymétrie**.

## Loi lognormale $(\mu, \sigma^2)$

Au lieu d'une somme, on a un produit de petits effets:  $Y_n = (Z_1 \cdot Z_2 \cdots Z_n)^{1/n}$ .

## Loi lognormale $(\mu, \sigma^2)$

Au lieu d'une somme, on a un **produit** de petits effets:  $Y_n = (Z_1 \cdot Z_2 \cdots Z_n)^{1/n}$ .

On a  $\ln Y_n = (\ln Z_1 + \cdots + \ln Z_n)/n$ .

## Loi lognormale $(\mu, \sigma^2)$

Au lieu d'une somme, on a un **produit** de petits effets:  $Y_n = (Z_1 \cdot Z_2 \cdots Z_n)^{1/n}$ .

On a  $\ln Y_n = (\ln Z_1 + \cdots + \ln Z_n)/n$ .

Sous les conditions du TLC:

$$\frac{\ln Y_n - \mathbb{E}[\ln Y_n]}{(\text{Var}[\ln Y_n])^{1/2}} \Rightarrow N(0, 1) \quad \text{pour } n \rightarrow \infty.$$

On dit que  $X \sim \text{lognormale}$  si  $\ln X \sim \text{normale}$ .

## Loi lognormale $(\mu, \sigma^2)$

Au lieu d'une somme, on a un **produit** de petits effets:  $Y_n = (Z_1 \cdot Z_2 \cdots Z_n)^{1/n}$ .

On a  $\ln Y_n = (\ln Z_1 + \cdots + \ln Z_n)/n$ .

Sous les conditions du TLC:

$$\frac{\ln Y_n - \mathbb{E}[\ln Y_n]}{(\text{Var}[\ln Y_n])^{1/2}} \Rightarrow N(0, 1) \quad \text{pour } n \rightarrow \infty.$$

On dit que  $X \sim$  **lognormale** si  $\ln X \sim$  normale.

Cette loi apparait souvent en économie et finance.

## Loi lognormale $(\mu, \sigma^2)$

Au lieu d'une somme, on a un **produit** de petits effets:  $Y_n = (Z_1 \cdot Z_2 \cdots Z_n)^{1/n}$ .

On a  $\ln Y_n = (\ln Z_1 + \cdots + \ln Z_n)/n$ .

Sous les conditions du TLC:

$$\frac{\ln Y_n - \mathbb{E}[\ln Y_n]}{(\text{Var}[\ln Y_n])^{1/2}} \Rightarrow N(0, 1) \quad \text{pour } n \rightarrow \infty.$$

On dit que  $X \sim$  **lognormale** si  $\ln X \sim$  normale.

Cette loi apparait souvent en économie et finance.

Si  $\ln X \sim N(\mu, \sigma^2)$ ,  $X \sim$  lognormale $(\mu, \sigma^2)$  a la densité

$$f(x) = \frac{1}{\sigma x \sqrt{2\pi}} e^{-(\ln x - \mu)^2 / (2\sigma^2)} \quad \text{pour } x > 0.$$

On a  $\mathbb{E}[X] = e^{\mu + \sigma^2/2}$  et  $\text{Var}[X] = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)$ .

## Loi lognormale $(\mu, \sigma^2)$

Au lieu d'une somme, on a un **produit** de petits effets:  $Y_n = (Z_1 \cdot Z_2 \cdots Z_n)^{1/n}$ .

On a  $\ln Y_n = (\ln Z_1 + \cdots + \ln Z_n)/n$ .

Sous les conditions du TLC:

$$\frac{\ln Y_n - \mathbb{E}[\ln Y_n]}{(\text{Var}[\ln Y_n])^{1/2}} \Rightarrow N(0, 1) \quad \text{pour } n \rightarrow \infty.$$

On dit que  $X \sim$  **lognormale** si  $\ln X \sim$  normale.

Cette loi apparait souvent en économie et finance.

Si  $\ln X \sim N(\mu, \sigma^2)$ ,  $X \sim$  lognormale $(\mu, \sigma^2)$  a la densité

$$f(x) = \frac{1}{\sigma x \sqrt{2\pi}} e^{-(\ln x - \mu)^2 / (2\sigma^2)} \quad \text{pour } x > 0.$$

On a  $\mathbb{E}[X] = e^{\mu + \sigma^2/2}$  et  $\text{Var}[X] = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)$ .

Attention:  $\mathbb{E}[\ln X] = \mu < \mu + \sigma^2/2 = \ln \mathbb{E}[X]$ .



**Exemple.** On place  $I_0$  dollars dans un fond mutuel, pour un an.  
Divisons l'année en  $n$  périodes égales.

**Exemple.** On place  $I_0$  dollars dans un fond mutuel, pour un an.  
Divisons l'année en  $n$  périodes égales.  
Soit  $R_j$  le taux de croissance (aléatoire et annuel) pour la période  $j$ .  
La valeur finale sera:

$$Y_n = I_0(1 + R_1)^{1/n}(1 + R_2)^{1/n} \cdots (1 + R_n)^{1/n}.$$

**Exemple.** On place  $I_0$  dollars dans un fond mutuel, pour un an.

Divisons l'année en  $n$  périodes égales.

Soit  $R_j$  le taux de croissance (aléatoire et annuel) pour la période  $j$ .

La valeur finale sera:

$$Y_n = I_0(1 + R_1)^{1/n}(1 + R_2)^{1/n} \cdots (1 + R_n)^{1/n}.$$

Si les  $R_j$  sont i.i.d., ou encore si  $\ln(1 + R_1), \dots, \ln(1 + R_n)$  obéissent au TLC, alors

$$\ln Y_n = \ln I_0 + \frac{1}{n} \sum_{j=1}^n \ln(1 + R_j) \approx \text{normal}$$

pour  $n$  grand. Donc  $Y_n \approx \text{lognormal}$ .

## Loi Gaussienne inverse

$X \sim \text{InvGaussian}(\mu, \lambda)$ , avec paramètre de localisation  $\mu > 0$  et paramètre de forme  $\lambda > 0$ , a la densité:

$$f(x) = \left( \frac{\lambda}{2\pi x^3} \right)^{1/2} \exp \left[ \frac{-\lambda(x - \mu)^2}{2\mu^2 x} \right] \quad \text{pour } x > 0,$$

et fonction de répartition

$$F(x) = \Phi(y) + e^{2\lambda/\mu}(1 - \Phi(y))$$

où  $y = (x/\mu - 1)\sqrt{\lambda/x}$ .

Si  $X = \{X(t), t \geq 0\}$  est un mouvement Brownien, alors le premier temps de passage au niveau  $\ell > 0$ ,  $T = \inf\{t > 0 : X(t) \geq \ell\}$  suit cette loi.

## Loi normale Gaussienne inverse

Si  $X \sim N(\mu + \beta Y, Y)$  où  $Y \sim \text{InvGaussian}(\delta, \gamma)$ ,  
alors  $X$  suit la loi normale Gaussienne inverse (NIG):  
 $X \sim \text{NIG}(\alpha, \beta, \mu, \delta)$ , où  $\alpha^2 = \beta^2 + \gamma^2$ .

## Loi normale Gaussienne inverse

Si  $X \sim N(\mu + \beta Y, Y)$  où  $Y \sim \text{InvGaussian}(\delta, \gamma)$ ,  
alors  $X$  suit la loi **normale Gaussienne inverse** (NIG):  
 $X \sim \text{NIG}(\alpha, \beta, \mu, \delta)$ , où  $\alpha^2 = \beta^2 + \gamma^2$ .

Utilisé récemment dans les processus à volatilité stochastique, en finance.  
Difficile à générer par inversion.

moyenne  $\mu + \delta\beta/\gamma$ ,

variance  $\delta\alpha^2/\gamma^3$ ,

coefficient d'asymétrie  $3\beta/(\alpha\sqrt{\delta\gamma})$ ,

coefficient d'aplatissement  $3(\alpha^2 + 4\beta^2)/(\delta\alpha^2\gamma)$ .

Le paramètre  $\alpha$  détermine l'épaisseur des queues.

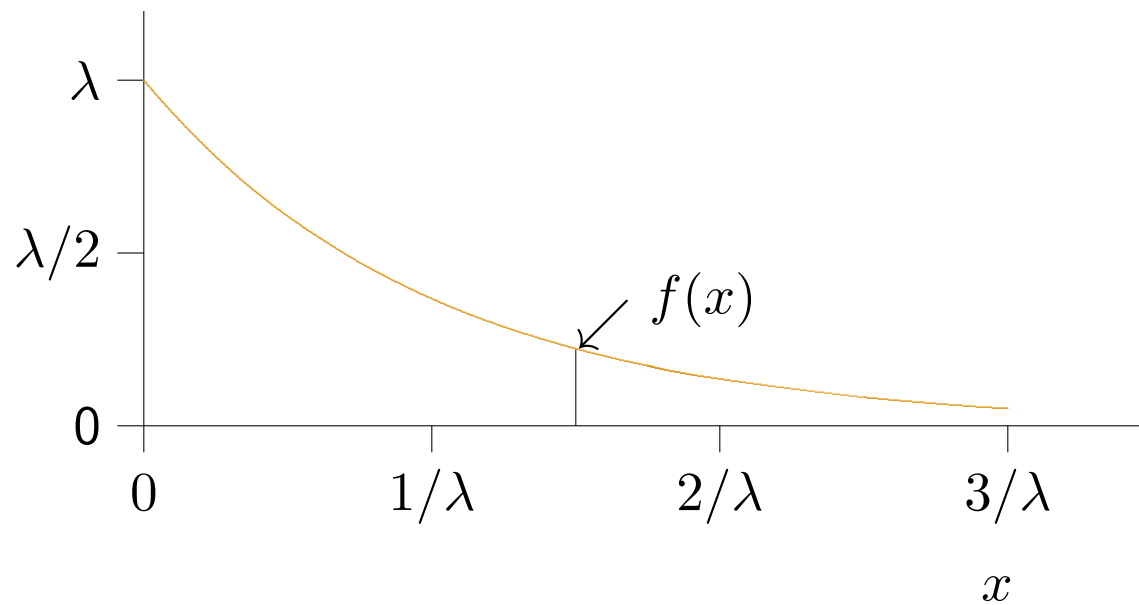
## Loi exponentielle ( $\lambda$ )

$$f(x) = \lambda e^{-\lambda x} \quad \text{pour } x > 0,$$

$$F(x) = 1 - e^{-\lambda x},$$

$$r(x) = \lambda.$$

$$\mathbb{E}[X] = 1/\lambda \quad \text{et} \quad \text{Var}[X] = 1/\lambda^2. \quad \text{CV} = 1.$$



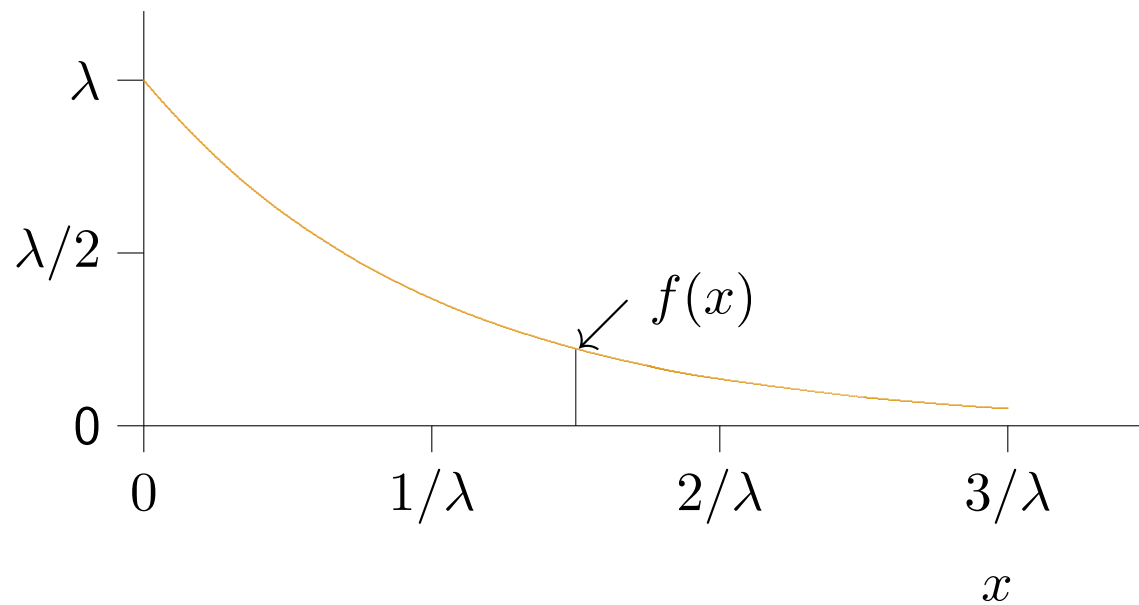
## Loi exponentielle ( $\lambda$ )

$$f(x) = \lambda e^{-\lambda x} \quad \text{pour } x > 0,$$

$$F(x) = 1 - e^{-\lambda x},$$

$$r(x) = \lambda.$$

$$\mathbb{E}[X] = 1/\lambda \quad \text{et} \quad \text{Var}[X] = 1/\lambda^2. \quad \text{CV} = 1.$$





Seule loi continue qui est **sans mémoire**:

$$\mathbb{P}[X > t + x \mid X > t] = \frac{\mathbb{P}[X > t + x]}{\mathbb{P}[X > t]}$$

Seule loi continue qui est **sans mémoire**:

$$\mathbb{P}[X > t + x \mid X > t] = \frac{\mathbb{P}[X > t + x]}{\mathbb{P}[X > t]} = \frac{e^{-\lambda(t+x)}}{e^{-\lambda t}}$$

Seule loi continue qui est **sans mémoire**:

$$\mathbb{P}[X > t + x \mid X > t] = \frac{\mathbb{P}[X > t + x]}{\mathbb{P}[X > t]} = \frac{e^{-\lambda(t+x)}}{e^{-\lambda t}} = e^{-\lambda x} = \mathbb{P}[X > x].$$

Seule loi continue qui est **sans mémoire**:

$$\mathbb{P}[X > t + x \mid X > t] = \frac{\mathbb{P}[X > t + x]}{\mathbb{P}[X > t]} = \frac{e^{-\lambda(t+x)}}{e^{-\lambda t}} = e^{-\lambda x} = \mathbb{P}[X > x].$$

Cette propriété simplifie beaucoup l'analyse mathématique et l'explique la grande popularité de la loi exponentielle.

Seule loi continue qui est **sans mémoire**:

$$\mathbb{P}[X > t + x \mid X > t] = \frac{\mathbb{P}[X > t + x]}{\mathbb{P}[X > t]} = \frac{e^{-\lambda(t+x)}}{e^{-\lambda t}} = e^{-\lambda x} = \mathbb{P}[X > x].$$

Cette propriété simplifie beaucoup l'analyse mathématique et l'explique la grande popularité de la loi exponentielle.

Exemple: durée de vie. Taux de panne constant.

Seule loi continue qui est **sans mémoire**:

$$\mathbb{P}[X > t + x \mid X > t] = \frac{\mathbb{P}[X > t + x]}{\mathbb{P}[X > t]} = \frac{e^{-\lambda(t+x)}}{e^{-\lambda t}} = e^{-\lambda x} = \mathbb{P}[X > x].$$

Cette propriété simplifie beaucoup l'analyse mathématique et l'explique la grande popularité de la loi exponentielle.

Exemple: durée de vie. Taux de panne constant.

**Exemple.** Programmation de l'événement `NextPeriod` dans `CallEv`.

Seule loi continue qui est **sans mémoire**:

$$\mathbb{P}[X > t + x \mid X > t] = \frac{\mathbb{P}[X > t + x]}{\mathbb{P}[X > t]} = \frac{e^{-\lambda(t+x)}}{e^{-\lambda t}} = e^{-\lambda x} = \mathbb{P}[X > x].$$

Cette propriété simplifie beaucoup l'analyse mathématique et l'explique la grande popularité de la loi exponentielle.

Exemple: durée de vie. Taux de panne constant.

**Exemple.** Programmation de l'événement NextPeriod dans CallEv.

Supposons que le taux d'arrivée passe de  $\lambda_{j-1}$  à  $\lambda_j$ .

Suffit de multiplier la durée résiduelle jusqu'à la prochaine arrivée par  $\lambda_{j-1}/\lambda_j$ .

```
class NextPeriod extends Event {
    checkQueue();
    nextArrival.reschedule ((nextArrival.time() - Sim.time())
                           * lambda[j-1] / lambda[j]);
    new NextPeriod(j+1).schedule (1.0 * HOUR);
}
```

Si  $X_1, \dots, X_k$  indep.,  $X_j \sim \text{exponentielle}(\lambda_j)$ ,  
alors  $X = \min(X_1, \dots, X_k) \sim \text{exponentielle}(\lambda)$  où  $\lambda = \lambda_1 + \dots + \lambda_k$ .



Si  $X_1, \dots, X_k$  indep.,  $X_j \sim \text{exponentielle}(\lambda_j)$ ,  
alors  $X = \min(X_1, \dots, X_k) \sim \text{exponentielle}(\lambda)$  où  $\lambda = \lambda_1 + \dots + \lambda_k$ .

**Exemple.** File  $M/M/1$ , taux d'arrivée  $\lambda$ , taux de service  $\mu$ .

Si  $X_1, \dots, X_k$  indep.,  $X_j \sim \text{exponentielle}(\lambda_j)$ ,  
alors  $X = \min(X_1, \dots, X_k) \sim \text{exponentielle}(\lambda)$  où  $\lambda = \lambda_1 + \dots + \lambda_k$ .

**Exemple.** File  $M/M/1$ , taux d'arrivée  $\lambda$ , taux de service  $\mu$ .  
Soit  $X$  = durée jusqu'au prochain événement.

Si  $X_1, \dots, X_k$  indep.,  $X_j \sim \text{exponentielle}(\lambda_j)$ ,  
alors  $X = \min(X_1, \dots, X_k) \sim \text{exponentielle}(\lambda)$  où  $\lambda = \lambda_1 + \dots + \lambda_k$ .

**Exemple.** File  $M/M/1$ , taux d'arrivée  $\lambda$ , taux de service  $\mu$ .  
Soit  $X$  = durée jusqu'au prochain événement.  
Serveur libre:  $X \sim \text{exponentielle}(\lambda)$ .

Si  $X_1, \dots, X_k$  indep.,  $X_j \sim \text{exponentielle}(\lambda_j)$ ,  
alors  $X = \min(X_1, \dots, X_k) \sim \text{exponentielle}(\lambda)$  où  $\lambda = \lambda_1 + \dots + \lambda_k$ .

**Exemple.** File  $M/M/1$ , taux d'arrivée  $\lambda$ , taux de service  $\mu$ .

Soit  $X$  = durée jusqu'au prochain événement.

Serveur libre:  $X \sim \text{exponentielle}(\lambda)$ .

Serveur occupé:  $X \sim \text{exponentielle}(\lambda + \mu)$ ;

Si  $X_1, \dots, X_k$  indep.,  $X_j \sim \text{exponentielle}(\lambda_j)$ ,  
alors  $X = \min(X_1, \dots, X_k) \sim \text{exponentielle}(\lambda)$  où  $\lambda = \lambda_1 + \dots + \lambda_k$ .

**Exemple.** File  $M/M/1$ , taux d'arrivée  $\lambda$ , taux de service  $\mu$ .

Soit  $X$  = durée jusqu'au prochain événement.

Serveur libre:  $X \sim \text{exponentielle}(\lambda)$ .

Serveur occupé:  $X \sim \text{exponentielle}(\lambda + \mu)$ ;

c'est une arrivée avec prob.  $p = \lambda/(\lambda + \mu)$ ,

une fin de service avec probabilité  $1 - p$ .

Si  $X_1, \dots, X_k$  indep.,  $X_j \sim \text{exponentielle}(\lambda_j)$ ,  
alors  $X = \min(X_1, \dots, X_k) \sim \text{exponentielle}(\lambda)$  où  $\lambda = \lambda_1 + \dots + \lambda_k$ .

**Exemple.** File  $M/M/1$ , taux d'arrivée  $\lambda$ , taux de service  $\mu$ .

Soit  $X$  = durée jusqu'au prochain événement.

Serveur libre:  $X \sim \text{exponentielle}(\lambda)$ .

Serveur occupé:  $X \sim \text{exponentielle}(\lambda + \mu)$ ;

c'est une arrivée avec prob.  $p = \lambda/(\lambda + \mu)$ ,

une fin de service avec probabilité  $1 - p$ .

Facile à implanter sans liste d'événement!

Si  $X_1, \dots, X_k$  indep.,  $X_j \sim \text{exponentielle}(\lambda_j)$ ,  
alors  $X = \min(X_1, \dots, X_k) \sim \text{exponentielle}(\lambda)$  où  $\lambda = \lambda_1 + \dots + \lambda_k$ .

**Exemple.** File  $M/M/1$ , taux d'arrivée  $\lambda$ , taux de service  $\mu$ .

Soit  $X$  = durée jusqu'au prochain événement.

Serveur libre:  $X \sim \text{exponentielle}(\lambda)$ .

Serveur occupé:  $X \sim \text{exponentielle}(\lambda + \mu)$ ;

c'est une arrivée avec prob.  $p = \lambda/(\lambda + \mu)$ ,

une fin de service avec probabilité  $1 - p$ .

Facile à implanter sans liste d'événement!

**Généralisations:** réseau de files d'attentes; chaînes de Markov en temps continu.

## Loi de Weibull $(\alpha, \lambda)$

$$\begin{aligned} f(x) &= \alpha \lambda^\alpha (x - \delta)^{\alpha-1} \exp[-(\lambda(x - \delta))^\alpha] && \text{pour } x > \delta, \\ F(x) &= 1 - \exp[-(\lambda(x - \delta))^\alpha]. \end{aligned}$$

Paramètres:  $\delta \in \mathbb{R}$  (localisation),  $\alpha > 0$  (forme),  $\lambda > 0$  (échelle).



## Loi de Weibull $(\alpha, \lambda)$

$$\begin{aligned} f(x) &= \alpha \lambda^\alpha (x - \delta)^{\alpha-1} \exp[-(\lambda(x - \delta))^\alpha] && \text{pour } x > \delta, \\ F(x) &= 1 - \exp[-(\lambda(x - \delta))^\alpha]. \end{aligned}$$

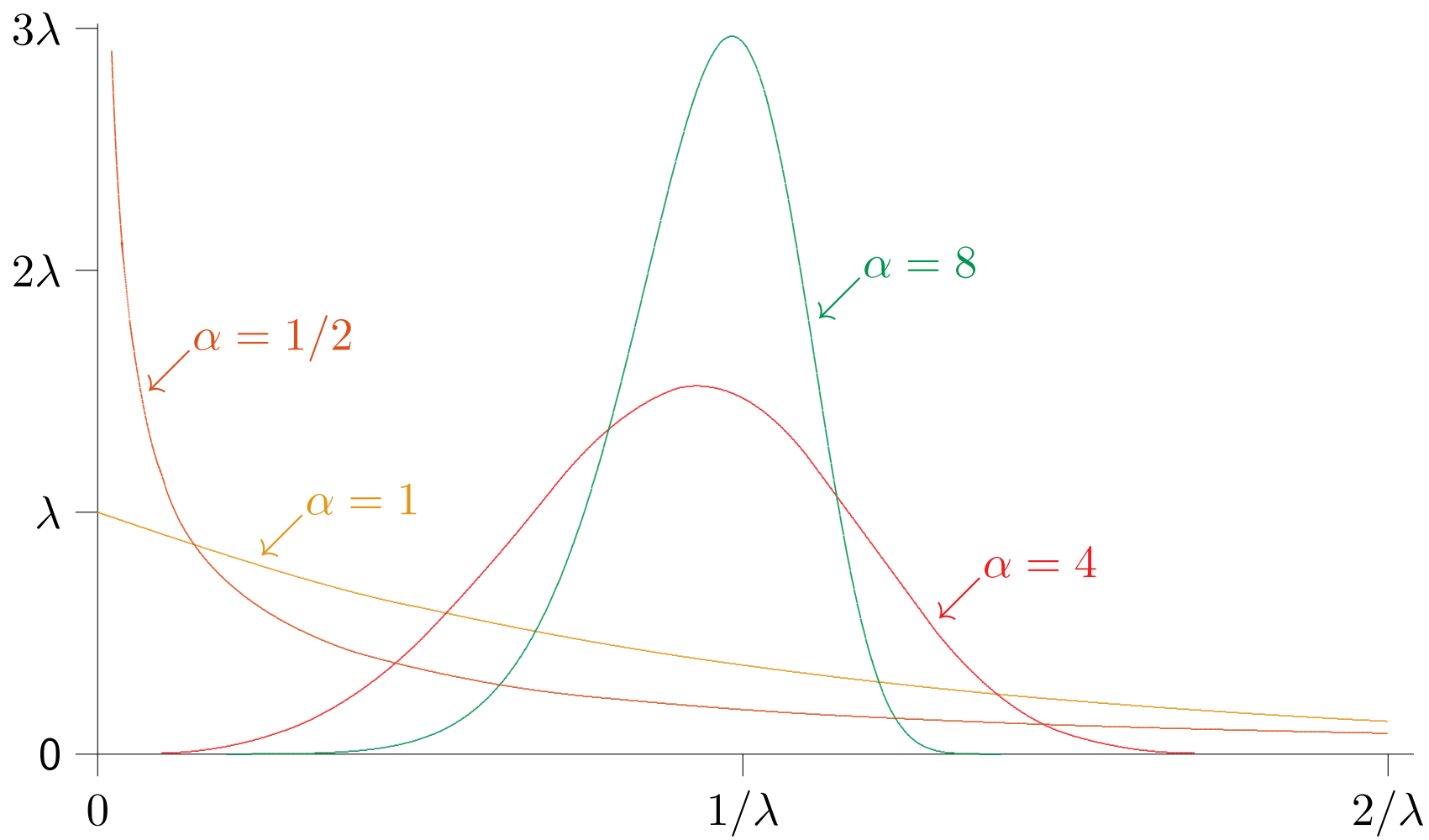
Paramètres:  $\delta \in \mathbb{R}$  (localisation),  $\alpha > 0$  (forme),  $\lambda > 0$  (échelle).

Souvent utilisé (avec  $\delta = 0$ ) pour modéliser les durées de vie.

Pour  $\alpha = 1$ : loi exponentielle.

Pour  $\alpha > 1$ : taux de panne **croissant**,  $CV < 1$ .

Pour  $\alpha < 1$ : taux de panne **décroissant**,  $CV > 1$ .



**Motivation:** le **minimum** de plusieurs v.a. i.i.d. bornées inférieurement suit approx. une loi de Weibull.

Interprétation: la panne d'un système survient lors de l'occurrence du premier événement parmi plusieurs événements indépendants.

**Motivation:** le **minimum** de plusieurs v.a. i.i.d. bornées inférieurement suit approx. une loi de Weibull.

Interprétation: la panne d'un système survient lors de l'occurrence du premier événement parmi plusieurs événements indépendants.

Autre application: Pour une fonction ayant des millions de minima locaux, la valeur retournée par un algorithme de recherche aléatoire qui trouve un minimum local à partir de chacun de  $n$  points de départ, puis retourne le meilleur, suit souvent (approx.) une loi Weibull. Ici,  $\delta$  est la valeur au minimum global.

**Motivation:** le **minimum** de plusieurs v.a. i.i.d. bornées inférieurement suit approx. une loi de Weibull.

Interprétation: la panne d'un système survient lors de l'occurrence du premier événement parmi plusieurs événements indépendants.

Autre application: Pour une fonction ayant des millions de minima locaux, la valeur retournée par un algorithme de recherche aléatoire qui trouve un minimum local à partir de chacun de  $n$  points de départ, puis retourne le meilleur, suit souvent (approx.) une loi Weibull. Ici,  $\delta$  est la valeur au minimum global.

S'applique aussi au maximum si on change le signe après.

**Théorème.** Soient  $X_1, \dots, X_n$  i.i.d.,  $\mathbb{P}[X_j \leq x] = G(x)$ ,  
 $\delta = \inf\{x \mid G(x) > 0\} > -\infty$ .

**Théorème.** Soient  $X_1, \dots, X_n$  i.i.d.,  $\mathbb{P}[X_j \leq x] = G(x)$ ,

$\delta = \inf\{x \mid G(x) > 0\} > -\infty$ .

Posons  $W_n = \min(X_1, X_2, \dots, X_n)$  et  $d_n$  tel que  $G(\delta + d_n) = 1/n$ . Si

$$\lim_{t \rightarrow \infty} \frac{G(\delta + 1/(tx))}{G(\delta + 1/t)} = x^{-\alpha}.$$

pour  $\alpha > 0$ , alors

$$\lim_{n \rightarrow \infty} \mathbb{P}[W_n - \delta < d_n x] = \begin{cases} 1 - \exp[-x^\alpha] & \text{si } x > 0, \\ 0 & \text{si } x \leq 0. \end{cases}$$

I.e.,  $(W_n - \delta)/d_n \Rightarrow \text{Weibull}(\alpha, 1)$  pour  $n \rightarrow \infty$ .

La valeur de  $\alpha$  dépend de la forme de  $G$  près de  $x = \delta$ .

## Loi de Gumbel

$X \sim \text{Gumbel}(\delta, \beta)$  a la fonction de répartition

$$F(x) = \begin{cases} \exp \left[ -e^{(\delta-x)/\beta} \right] & \text{si } \beta > 0, \\ 1 - \exp \left[ -e^{(\delta-x)/\beta} \right] & \text{si } \beta < 0. \end{cases}$$

pour  $x \in \mathbb{R}$ . Paramètre de localisation  $\delta$  et d'échelle  $\beta \neq 0$ .

**Gumbel standard:**  $\delta = 0$  et  $\beta = 1$ .

$X \sim \text{Gumbel}(\delta, \beta)$  ssi  $2\delta - X \sim \text{Gumbel}(\delta, -\beta)$ , donc on peut se limiter à  $\beta > 0$ .



## Loi de Gumbel

$X \sim \text{Gumbel}(\delta, \beta)$  a la fonction de répartition

$$F(x) = \begin{cases} \exp \left[ -e^{(\delta-x)/\beta} \right] & \text{si } \beta > 0, \\ 1 - \exp \left[ -e^{(\delta-x)/\beta} \right] & \text{si } \beta < 0. \end{cases}$$

pour  $x \in \mathbb{R}$ . Paramètre de localisation  $\delta$  et d'échelle  $\beta \neq 0$ .

**Gumbel standard:**  $\delta = 0$  et  $\beta = 1$ .

$X \sim \text{Gumbel}(\delta, \beta)$  ssi  $2\delta - X \sim \text{Gumbel}(\delta, -\beta)$ , donc on peut se limiter à  $\beta > 0$ .

Si  $Y \sim \text{Weibull}(\alpha, \lambda)$  et  $0 \neq c \in \mathbb{R}$ , alors  $X = c \ln Y \sim \text{Gumbel}(\delta, \beta)$  avec  $\delta = (\ln \lambda)/\alpha$  et  $\beta = -1/(c\alpha)$ .

On peut donc interpréter une Weibull comme une **logGumbel** (avec  $\beta < 0$ ).

## Loi de Gumbel

$X \sim \text{Gumbel}(\delta, \beta)$  a la fonction de répartition

$$F(x) = \begin{cases} \exp \left[ -e^{(\delta-x)/\beta} \right] & \text{si } \beta > 0, \\ 1 - \exp \left[ -e^{(\delta-x)/\beta} \right] & \text{si } \beta < 0. \end{cases}$$

pour  $x \in \mathbb{R}$ . Paramètre de localisation  $\delta$  et d'échelle  $\beta \neq 0$ .

**Gumbel standard:**  $\delta = 0$  et  $\beta = 1$ .

$X \sim \text{Gumbel}(\delta, \beta)$  ssi  $2\delta - X \sim \text{Gumbel}(\delta, -\beta)$ , donc on peut se limiter à  $\beta > 0$ .

Si  $Y \sim \text{Weibull}(\alpha, \lambda)$  et  $0 \neq c \in \mathbb{R}$ , alors  $\textcolor{brown}{X} = c \ln Y \sim \text{Gumbel}(\delta, \beta)$  avec  $\delta = (\ln \lambda)/\alpha$  et  $\beta = -1/(c\alpha)$ .

On peut donc interpréter une Weibull comme une **logGumbel** (avec  $\beta < 0$ ).

Si  $X_1, \dots, X_n$  i.i.d.,  $\mathbb{P}[X_j \leq x] = \textcolor{brown}{G}(x)$ , et  $G$  a une queue infinie à décroissance exponentielle, et  $\textcolor{brown}{M}_n = \max(X_1, \dots, X_n)$ , alors  $X_n = (M_n - d_n)/c_n$  converge vers une Gumbel pour certaines suites  $\{d_n\}$  et  $\{c_n\}$ .

## Loi de Fréchet

Paramètres de localisation  $\delta \in \mathbb{R}$ , d'échelle  $\beta > 0$ , et de forme  $\alpha \in \mathbb{R}$ . On a

$$F(x) = \exp \left[ -((x - \delta)/\beta)^{-\alpha} \right]$$

pour  $x > \delta$ .

Si  $X_1, \dots, X_n$  i.i.d.,  $\mathbb{P}[X_j \leq x] = G(x)$ , et  $G$  a une queue infinie à décroissance sous-exponentielle, et  $M_n = \max(X_1, \dots, X_n)$ , alors  $(M_n - d_n)/c_n$  converge vers une loi de Fréchet pour certaines suites  $\{d_n\}$  et  $\{c_n\}$ .

## Loi gamma ( $\alpha, \lambda$ )

$$f(x) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} \quad \text{pour } x > 0,$$

où

$$\Gamma(y) = \int_0^\infty x^{y-1} e^{-x} dx \quad \text{pour } y > 0$$

( $\Gamma(k) = (k-1)!$  si  $k$  entier).

Souvent utilisé pour modéliser des durées d'activités.

## Loi gamma ( $\alpha, \lambda$ )

$$f(x) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} \quad \text{pour } x > 0,$$

où

$$\Gamma(y) = \int_0^\infty x^{y-1} e^{-x} dx \quad \text{pour } y > 0$$

( $\Gamma(k) = (k-1)!$  si  $k$  entier).

Souvent utilisé pour modéliser des durées d'activités.

Pour  $\alpha = 1$ : exponentielle.

## Loi gamma ( $\alpha, \lambda$ )

$$f(x) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} \quad \text{pour } x > 0,$$

où

$$\Gamma(y) = \int_0^\infty x^{y-1} e^{-x} dx \quad \text{pour } y > 0$$

( $\Gamma(k) = (k-1)!$  si  $k$  entier).

Souvent utilisé pour modéliser des durées d'activités.

Pour  $\alpha = 1$ : exponentielle.

Pour  $2\alpha$  entier et  $\lambda = 1/2$ : chi-deux.

## Loi gamma ( $\alpha, \lambda$ )

$$f(x) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} \quad \text{pour } x > 0,$$

où

$$\Gamma(y) = \int_0^\infty x^{y-1} e^{-x} dx \quad \text{pour } y > 0$$

( $\Gamma(k) = (k-1)!$  si  $k$  entier).

Souvent utilisé pour modéliser des durées d'activités.

Pour  $\alpha = 1$ : exponentielle.

Pour  $2\alpha$  entier et  $\lambda = 1/2$ : chi-deux.

Pour  $\alpha = k$  entier:  $\text{gamma}(\alpha, \lambda) \equiv \text{Erlang}(k, \lambda)$ .

## Loi gamma ( $\alpha, \lambda$ )

$$f(x) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} \quad \text{pour } x > 0,$$

où

$$\Gamma(y) = \int_0^\infty x^{y-1} e^{-x} dx \quad \text{pour } y > 0$$

( $\Gamma(k) = (k-1)!$  si  $k$  entier).

Souvent utilisé pour modéliser des durées d'activités.

Pour  $\alpha = 1$ : exponentielle.

Pour  $2\alpha$  entier et  $\lambda = 1/2$ : chi-deux.

Pour  $\alpha = k$  entier:  $\text{gamma}(\alpha, \lambda) \equiv \text{Erlang}(k, \lambda)$ .

C'est une somme de  $k$  exponentielles( $\lambda$ ) i.i.d.



## Loi gamma ( $\alpha, \lambda$ )

$$f(x) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} \quad \text{pour } x > 0,$$

où

$$\Gamma(y) = \int_0^\infty x^{y-1} e^{-x} dx \quad \text{pour } y > 0$$

( $\Gamma(k) = (k-1)!$  si  $k$  entier).

Souvent utilisé pour modéliser des durées d'activités.

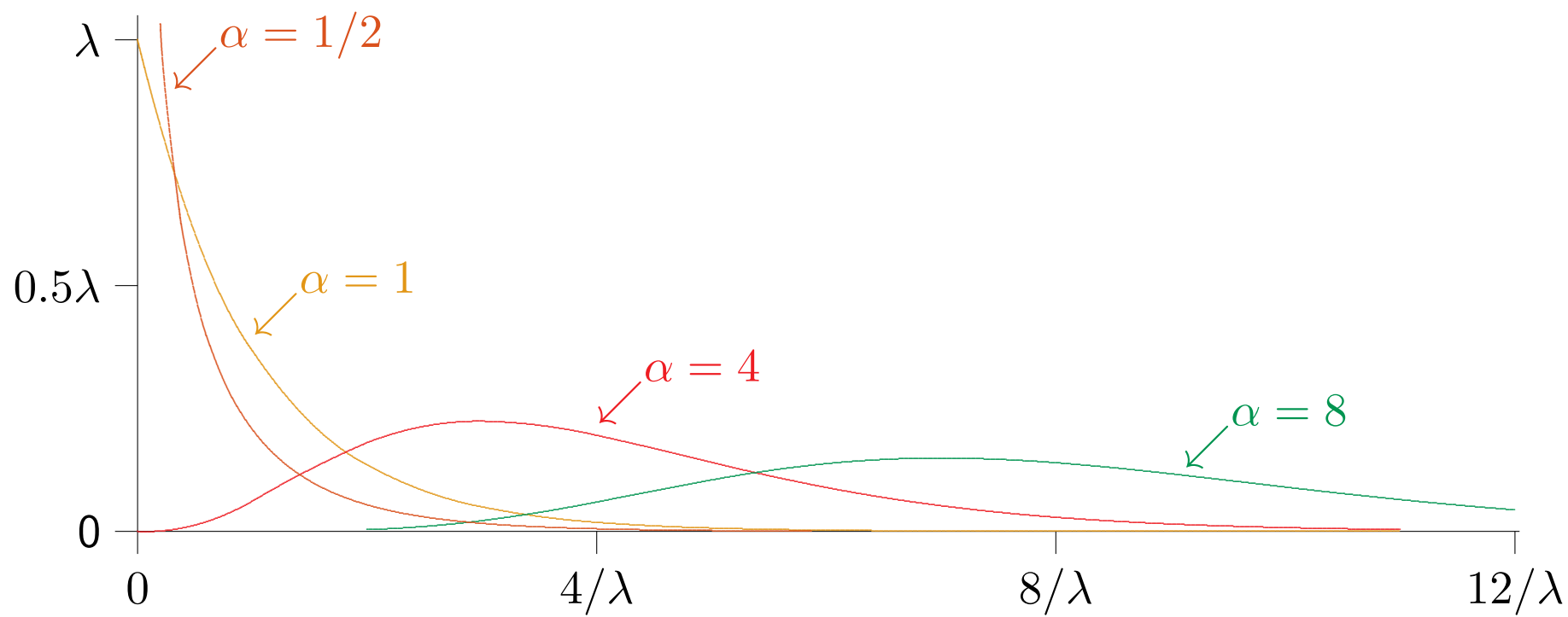
Pour  $\alpha = 1$ : exponentielle.

Pour  $2\alpha$  entier et  $\lambda = 1/2$ : chi-deux.

Pour  $\alpha = k$  entier:  $\text{gamma}(\alpha, \lambda) \equiv \text{Erlang}(k, \lambda)$ .

C'est une somme de  $k$  exponentielles( $\lambda$ ) i.i.d.

Si  $X \sim \text{Erlang}(k, \lambda)$  et  $Y \sim \text{Poisson}(\lambda x)$ , alors  $\mathbb{P}[X \leq x] = \mathbb{P}[Y \geq k]$ .



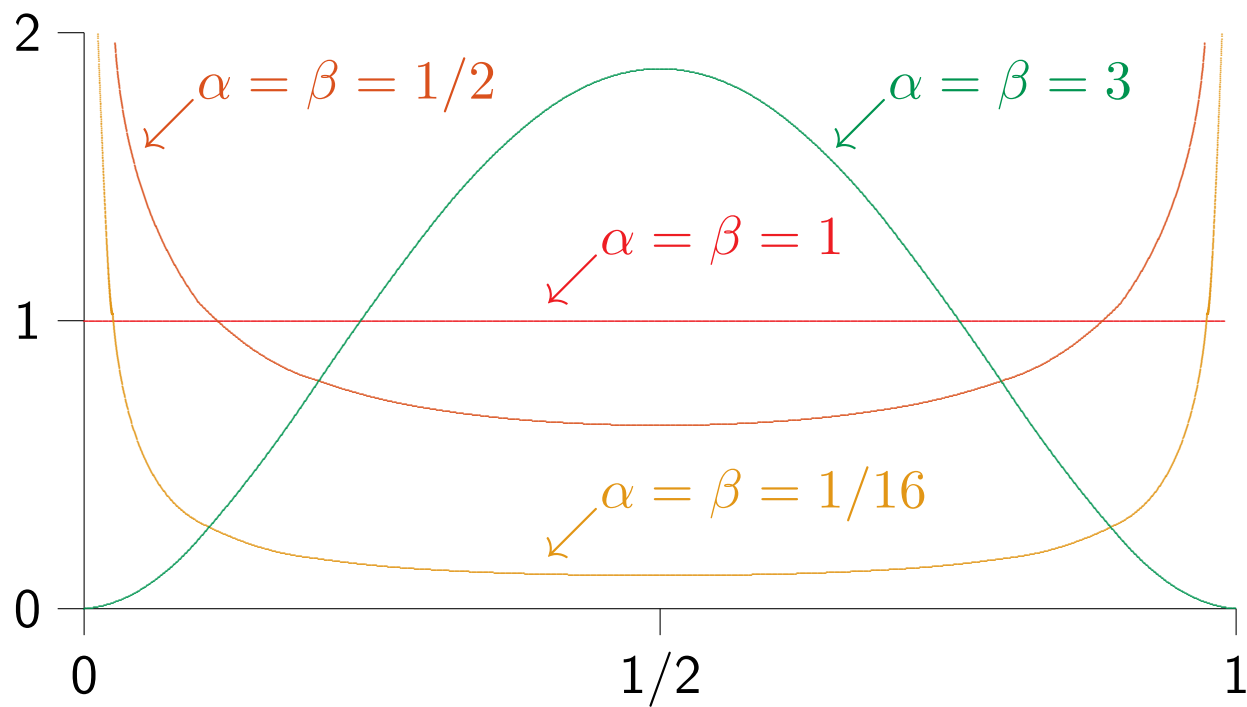
## Loi beta ( $\alpha, \beta$ )

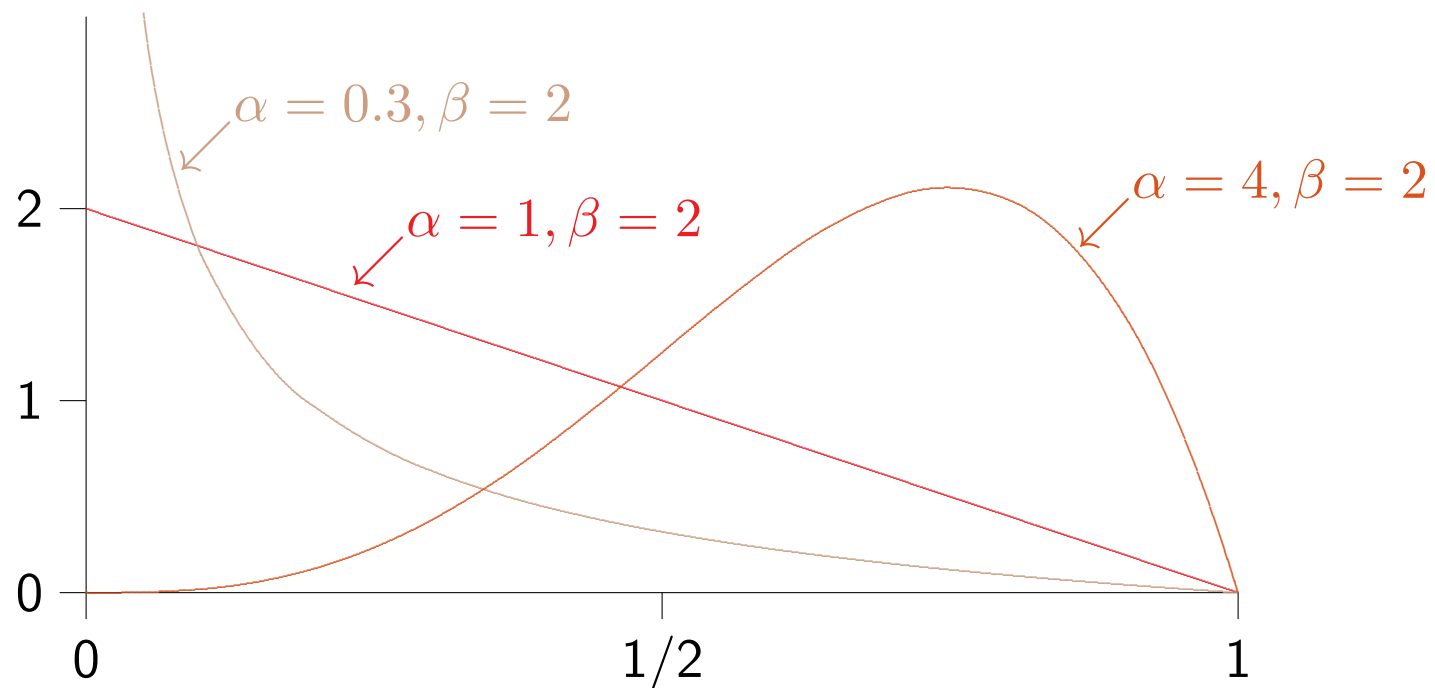
$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} \quad \text{pour } 0 < x < 1.$$

Densité sur un intervalle borné, très flexible.

On peut ajouter des paramètres d'échelle et de localisation:

$$Y = a + (b - a)X.$$





Voir Law et Kelton (2000) pour d'autres figures.

## Loi du chi-deux ( $k$ )

$$f(x) = \frac{x^{(k/2)-1} e^{-x/2}}{2^{k/2} \Gamma(k/2)} \quad \text{pour } x > 0.$$

Cas particulier de la loi gamma.

$$\mathbb{E}[X] = k \quad \text{et} \quad \text{Var}[X] = 2k.$$

## Loi du chi-deux ( $k$ )

$$f(x) = \frac{x^{(k/2)-1} e^{-x/2}}{2^{k/2} \Gamma(k/2)} \quad \text{pour } x > 0.$$

Cas particulier de la loi gamma.

$$\mathbb{E}[X] = k \quad \text{et} \quad \text{Var}[X] = 2k.$$

Si  $X_1, \dots, X_k$  i.i.d.  $N(0, 1)$ , alors  $X = X_1^2 + \dots + X_k^2 \sim \text{chi-deux}(k)$ .

## Loi du chi-deux ( $k$ )

$$f(x) = \frac{x^{(k/2)-1} e^{-x/2}}{2^{k/2} \Gamma(k/2)} \quad \text{pour } x > 0.$$

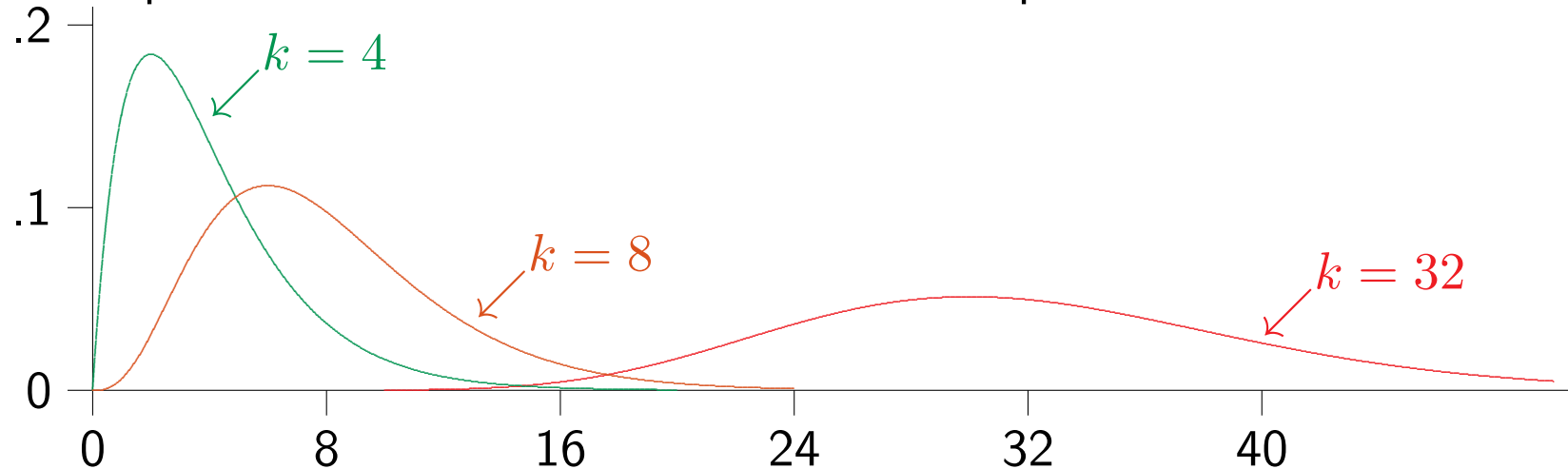
Cas particulier de la loi gamma.

$$\mathbb{E}[X] = k \quad \text{et} \quad \text{Var}[X] = 2k.$$

Si  $X_1, \dots, X_k$  i.i.d.  $N(0, 1)$ , alors  $X = X_1^2 + \dots + X_k^2 \sim \text{chi-deux}(k)$ .

Utilisé dans les tests de chi-deux.

Aussi pour calculer un intervalle de confiance pour la variance.





## Loi de Pareto ( $\alpha, \beta$ )

Il y en a plusieurs types.

Propriété:  $\bar{F}(x) \approx cx^{-\alpha}$ .

## Loi de Pareto ( $\alpha, \beta$ )

Il y en a plusieurs types.

Propriété:  $\bar{F}(x) \approx cx^{-\alpha}$ .

Par exemple:

$$\begin{aligned} f(x) &= \alpha\beta^\alpha x^{-(\alpha+1)} && \text{pour } x > \beta, \\ F(x) &= 1 - (\beta/x)^\alpha. \end{aligned}$$

## Loi de Pareto $(\alpha, \beta)$

Il y en a plusieurs types.

Propriété:  $\bar{F}(x) \approx cx^{-\alpha}$ .

Par exemple:

$$\begin{aligned} f(x) &= \alpha\beta^\alpha x^{-(\alpha+1)} && \text{pour } x > \beta, \\ F(x) &= 1 - (\beta/x)^\alpha. \end{aligned}$$

$\mathbb{E}[X] = \alpha\beta/(\alpha - 1)$  pour  $\alpha > 1$ , infini sinon.

$\text{Var}[X] = \alpha\beta^2/[(\alpha - 2)(\alpha - 1)]$  pour  $\alpha > 2$ , infini sinon.

## Loi de Pareto ( $\alpha, \beta$ )

Il y en a plusieurs types.

Propriété:  $\bar{F}(x) \approx cx^{-\alpha}$ .

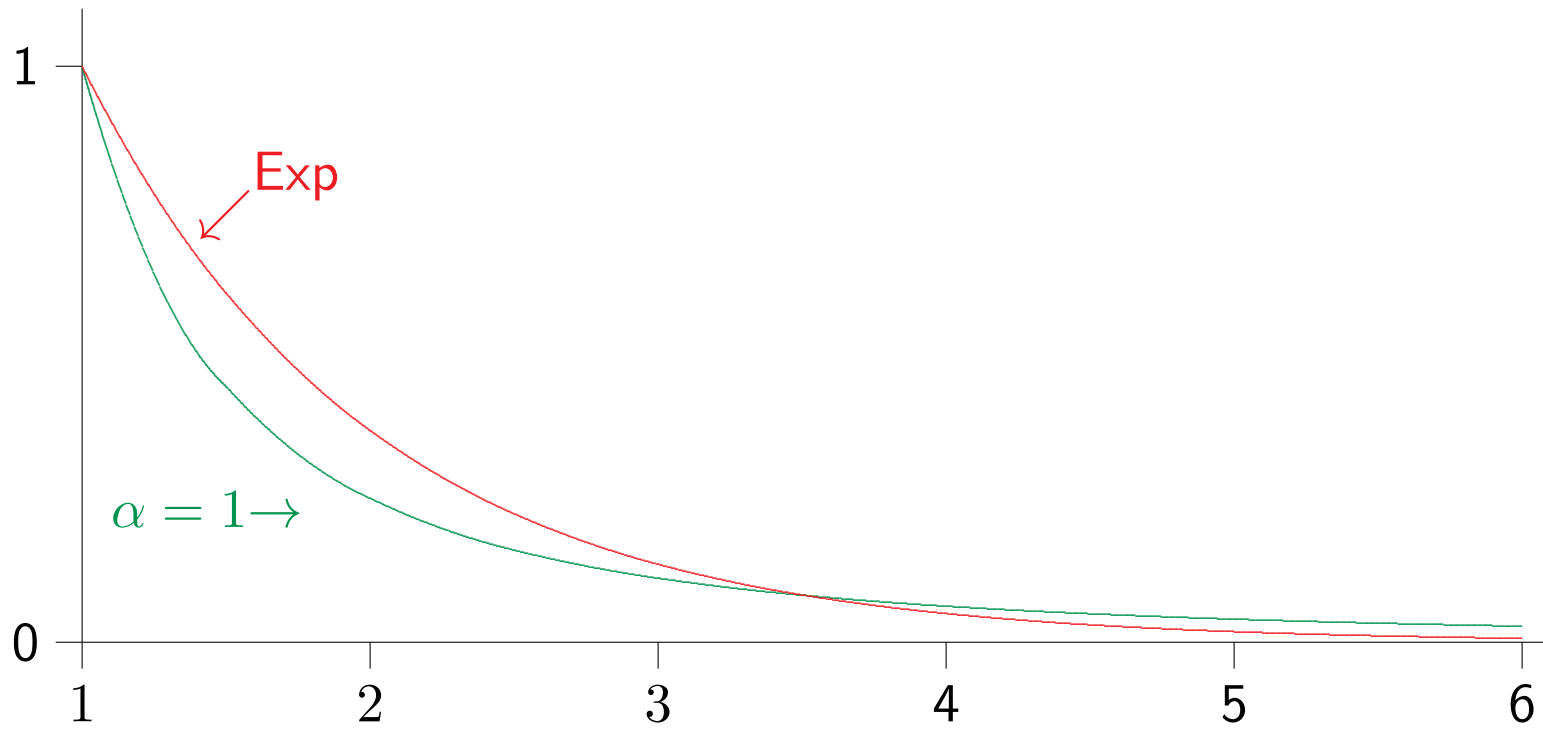
Par exemple:

$$\begin{aligned} f(x) &= \alpha\beta^\alpha x^{-(\alpha+1)} && \text{pour } x > \beta, \\ F(x) &= 1 - (\beta/x)^\alpha. \end{aligned}$$

$\mathbb{E}[X] = \alpha\beta/(\alpha - 1)$  pour  $\alpha > 1$ , infini sinon.

$\text{Var}[X] = \alpha\beta^2/[(\alpha - 2)(\alpha - 1)]$  pour  $\alpha > 2$ , infini sinon.

Utilisé pour modéliser la taille des réclamations en assurance, et la taille des “messages” dans les réseaux de communication.



## Familles de Johnson

Très flexibles, grande variété de formes.

Peut imiter la plupart des lois standard.

Forme générale:

$$F(x) = \Phi[\gamma + \delta g((x - \xi)/\lambda)], \quad \text{for } x \in \mathbb{R}.$$

Paramètres:  $\gamma$  et  $\delta > 0$  (forme),  $\xi$  (localisation),  $\lambda > 0$  (échelle).

Transformation  $g$  choisie parmi:

$$g(y) = \begin{cases} \ln y & \text{famille lognormale,} \\ \sinh^{-1}(y) = \ln(y + \sqrt{y^2 + 1}) & \text{famille non bornée,} \\ \ln(y/(1 - y)) & \text{famille bornée,} \\ y & \text{famille normale.} \end{cases}$$

## Familles de Johnson

Très flexibles, grande variété de formes.

Peut imiter la plupart des lois standard.

Forme générale:

$$F(x) = \Phi[\gamma + \delta g((x - \xi)/\lambda)], \quad \text{for } x \in \mathbb{R}.$$

Paramètres:  $\gamma$  et  $\delta > 0$  (forme),  $\xi$  (localisation),  $\lambda > 0$  (échelle).

Transformation  $g$  choisie parmi:

$$g(y) = \begin{cases} \ln y & \text{famille lognormale,} \\ \sinh^{-1}(y) = \ln(y + \sqrt{y^2 + 1}) & \text{famille non bornée,} \\ \ln(y/(1 - y)) & \text{famille bornée,} \\ y & \text{famille normale.} \end{cases}$$

Dans chaque cas,  $Z = \gamma + \delta g((X - \xi)/\lambda) \sim N(0, 1)$ .

Facile de générer des v.a.:  $X = \xi + \lambda g^{-1}((Z - \gamma)/\delta)$  où

$$g^{-1}(z) = \begin{cases} e^z & \text{pour la famille lognormale,} \\ (e^z - e^{-z})/2 & \text{pour la famille non bornée,} \\ 1/(1 + e^{-z}) & \text{pour la famille bornée,} \\ z & \text{pour la famille normale.} \end{cases}$$

Densité:

$$f(x) = \frac{\delta}{\lambda\sqrt{2\pi}} f'((x - \xi)/\lambda) \exp \left[ -(\gamma + \delta f((x - \xi)/\lambda))^2 / 2 \right] \quad \text{pour } x \in H,$$

$H = (-\infty, \infty)$  pour normale et non bornée,

$H = [\xi, \xi + \lambda]$  pour bornée,

$H = [\xi, \infty)$  pour lognormale.



## Lois à phases exponentielles

Approxime une loi quelconque par une somme ou un mélange d'exponentielles.

Motivation: avantages de la propriété sans mémoire de l'exponentielle.

## Lois à phases exponentielles

Approxime une loi quelconque par une somme ou un mélange d'exponentielles.

Motivation: avantages de la propriété sans mémoire de l'exponentielle.

A. Somme de  $k$  exponentielles i.i.d.: Erlang.

Coefficient de variation  $1/\sqrt{k}$ .

## Lois à phases exponentielles

Approxime une loi quelconque par une somme ou un mélange d'exponentielles.

Motivation: avantages de la propriété sans mémoire de l'exponentielle.

A. Somme de  $k$  exponentielles i.i.d.: Erlang.

Coefficient de variation  $1/\sqrt{k}$ .

B. Somme de  $k$  exponentielles de taux  $\lambda_1, \dots, \lambda_k$ : hypoexponentielle( $\lambda_1, \dots, \lambda_k$ ).

Coefficient de variation  $< 1$ .

Idée: décompose une durée en une somme de durées plus courtes, indép.

## Lois à phases exponentielles

Approxime une loi quelconque par une somme ou un mélange d'exponentielles.

Motivation: avantages de la propriété sans mémoire de l'exponentielle.

A. Somme de  $k$  exponentielles i.i.d.: Erlang.

Coefficient de variation  $1/\sqrt{k}$ .

B. Somme de  $k$  exponentielles de taux  $\lambda_1, \dots, \lambda_k$ : hypoexponentielle( $\lambda_1, \dots, \lambda_k$ ).

Coefficient de variation  $< 1$ .

Idée: décompose une durée en une somme de durées plus courtes, indép.

C. Mélange de  $k$  exponentielles (i.e., en parallèle).

$X$  exponentielle( $\lambda_j$ ) avec probabilité  $p_j$ ,  $1 \leq j \leq k$ :

$X \sim \text{hyperexponentielle}(p_1, \lambda_1, \dots, p_k, \lambda_k)$ .

## Lois à phases exponentielles

Approxime une loi quelconque par une somme ou un mélange d'exponentielles.

Motivation: avantages de la propriété sans mémoire de l'exponentielle.

A. Somme de  $k$  exponentielles i.i.d.: **Erlang**.

Coefficient de variation  $1/\sqrt{k}$ .

B. Somme de  $k$  exponentielles de taux  $\lambda_1, \dots, \lambda_k$ : **hypoexponentielle** $(\lambda_1, \dots, \lambda_k)$ .

Coefficient de variation  $< 1$ .

Idée: décompose une durée en une somme de durées plus courtes, indép.

C. Mélange de  $k$  exponentielles (i.e., en parallèle).

$X$  exponentielle( $\lambda_j$ ) avec probabilité  $p_j$ ,  $1 \leq j \leq k$ :

$X \sim$  **hyperexponentielle** $(p_1, \lambda_1, \dots, p_k, \lambda_k)$ .

$\mathbb{E}[X] = \sum_{j=1}^k p_j / \lambda_j$  et  $\text{Var}[X] = 2 \sum_{j=1}^k p_j / \lambda_j^2 - \mathbb{E}[X]^2$ .

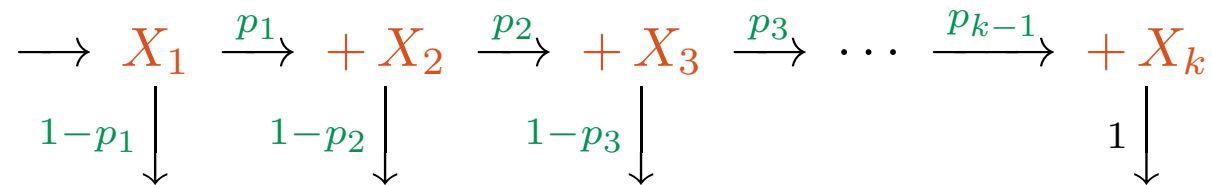
Coefficient de variation  $> 1$ , augmente avec  $k$  (en gros).

Exemple: temps de CPU consommé par un programme.

## D. Loi à phases exponentielles.

$X = X_1 + \cdots + X_R$  où  $X_1, \dots, X_k, R$  v.a. indep.,  $X_j \sim \text{exponentielle}(\lambda_j)$ ,

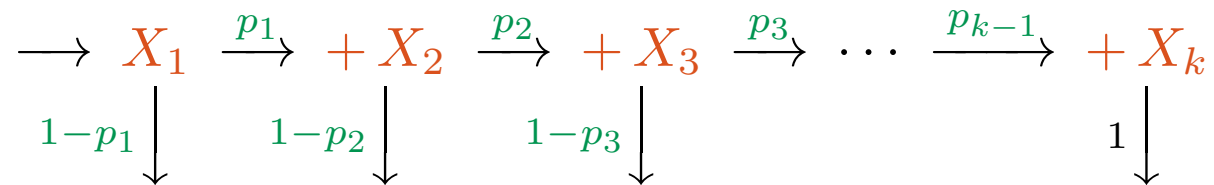
$$\mathbb{P}[R = r] = \begin{cases} 1 - p_1 & \text{if } r = 1, \\ p_1 \cdots p_{r-1}(1 - p_r) & \text{if } 2 \leq r \leq k - 1, \\ p_1 \cdots p_{k-1} & \text{if } r = k. \end{cases}$$



## D. Loi à phases exponentielles.

$X = X_1 + \cdots + X_R$  où  $X_1, \dots, X_k, R$  v.a. indep.,  $X_j \sim \text{exponentielle}(\lambda_j)$ ,

$$\mathbb{P}[R = r] = \begin{cases} 1 - p_1 & \text{if } r = 1, \\ p_1 \cdots p_{r-1}(1 - p_r) & \text{if } 2 \leq r \leq k - 1, \\ p_1 \cdots p_{k-1} & \text{if } r = k. \end{cases}$$

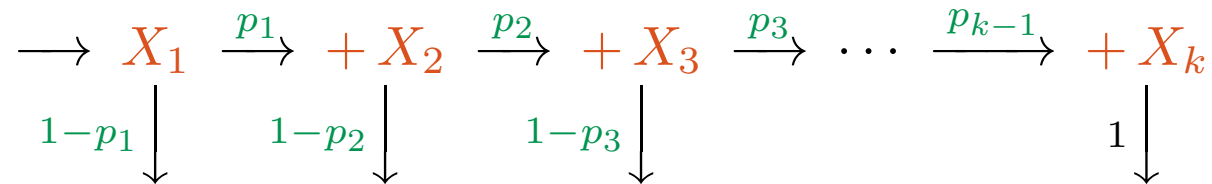


**Théorème.** Si  $\textcolor{brown}{F}$  est une fonction de répartition continue sur  $[0, \infty)$ , pour tout  $\epsilon > 0$ , il existe une loi exponentielle à phases  $H$  telle que  $|H(t) - \textcolor{brown}{F}(t)| < \epsilon$  pour tout  $t > 0$ .

## D. Loi à phases exponentielles.

$X = X_1 + \cdots + X_R$  où  $X_1, \dots, X_k, R$  v.a. indep.,  $X_j \sim \text{exponentielle}(\lambda_j)$ ,

$$\mathbb{P}[R = r] = \begin{cases} 1 - p_1 & \text{if } r = 1, \\ p_1 \cdots p_{r-1}(1 - p_r) & \text{if } 2 \leq r \leq k - 1, \\ p_1 \cdots p_{k-1} & \text{if } r = k. \end{cases}$$

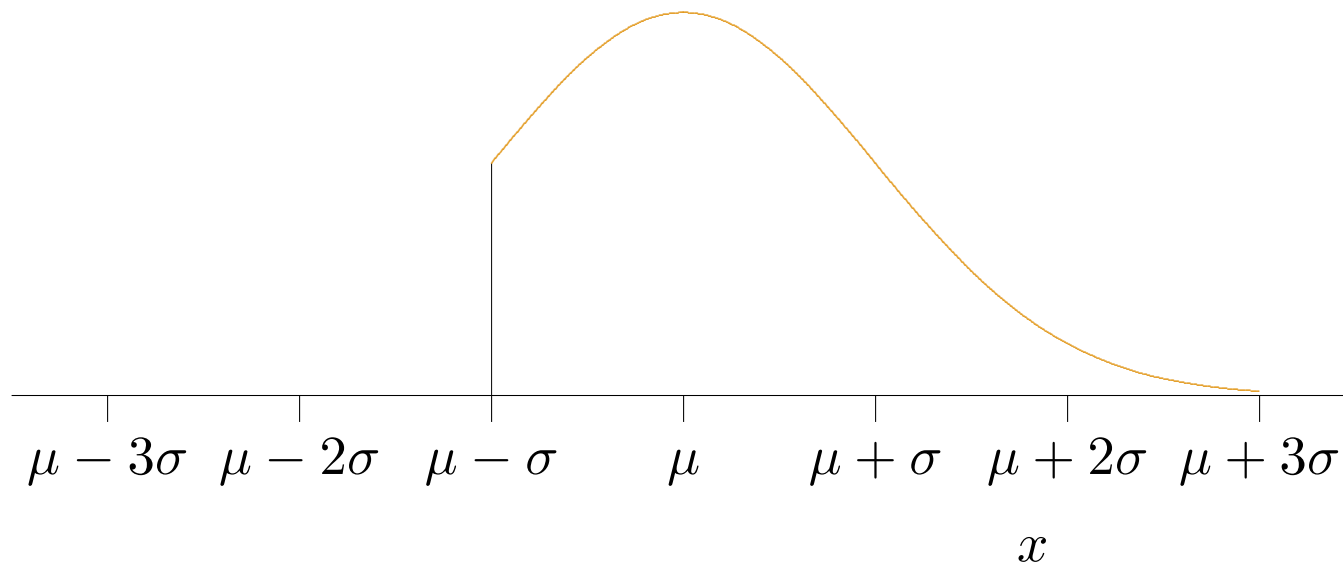


**Théorème.** Si  $F$  est une fonction de répartition continue sur  $[0, \infty)$ , pour tout  $\epsilon > 0$ , il existe une loi exponentielle à phases  $H$  telle que  $|H(t) - F(t)| < \epsilon$  pour tout  $t > 0$ .

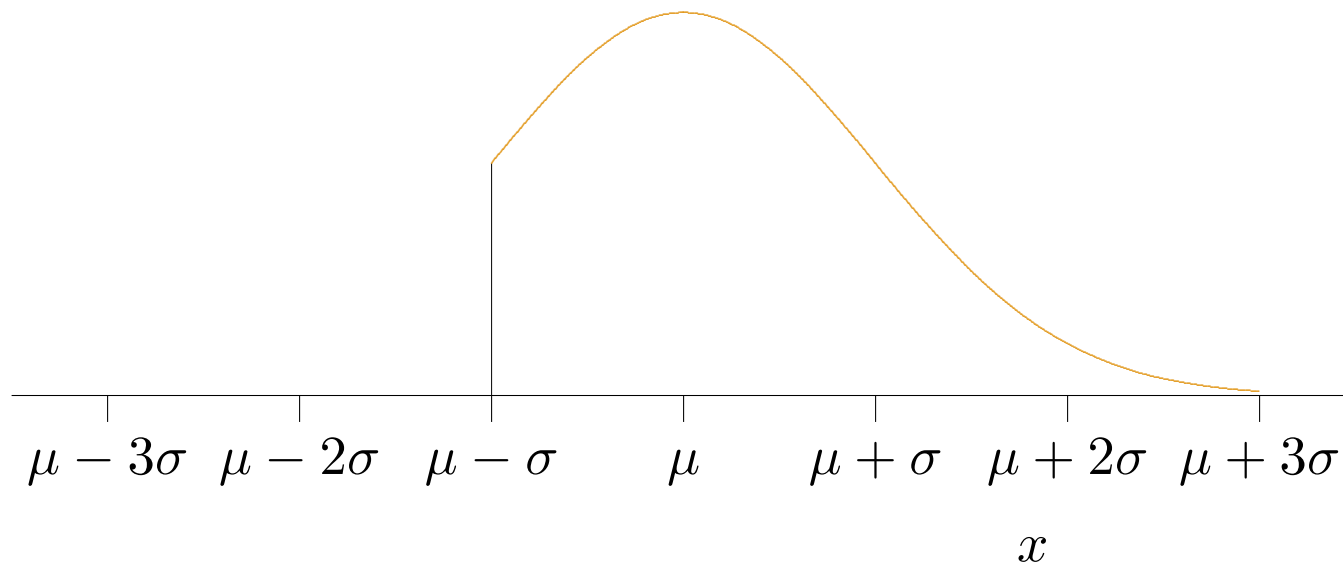
**Application.** Dans une simulation, si on approxime les lois de toutes les durées par des lois à phases exponentielles, alors plus besoin de liste d'événements!



## Lois tronquées



## Lois tronquées

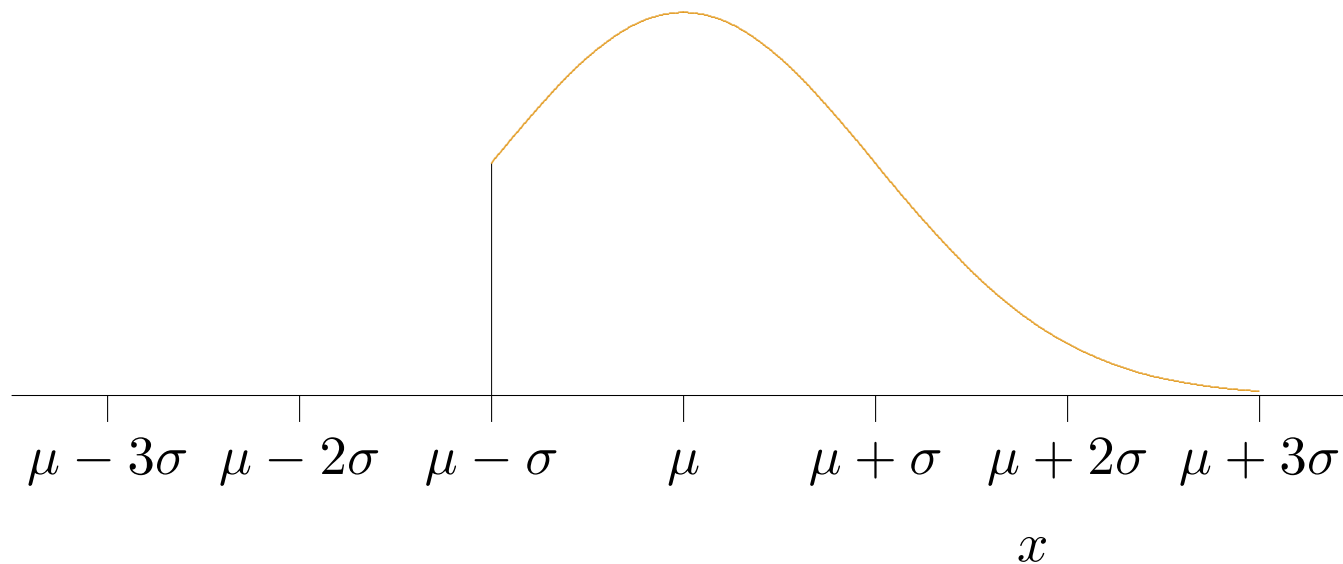


Pour tronquer  $f(x)$  sur un intervalle  $(a, b)$ :

$$\tilde{f}(x) = \begin{cases} f(x)/K & \text{pour } a < x < b, \\ 0 & \text{ailleurs,} \end{cases}$$

où  $K = \int_a^b f(u)du$ .

## Lois tronquées



Pour tronquer  $f(x)$  sur un intervalle  $(a, b)$ :

$$\tilde{f}(x) = \begin{cases} f(x)/K & \text{pour } a < x < b, \\ 0 & \text{ailleurs,} \end{cases}$$

où  $K = \int_a^b f(u)du$ .

Pour générer des v.a. selon  $\tilde{f}$ :  $U \sim \text{Uniforme}(F(a), F(b))$  et  $X = F^{-1}(U)$ .

## Lois décalées

$X = X_0 + a$  où  $X_0$  a une loi standard.

Exemples: exponentielle, gamma, lognormale, ...

## Lois décalées

$X = X_0 + a$  où  $X_0$  a une loi standard.

Exemples: exponentielle, gamma, lognormale, ...

Estimer le paramètre  $a$  à partir des données est parfois difficile!

## Mélanges de lois

$$f(x) = \sum_{j=1}^{\infty} w_j f_j(x),$$

où chaque  $f_j$  est une densité et les **poids**  $w_j$  somment à 1.

## Mélanges de lois

$$f(x) = \sum_{j=1}^{\infty} w_j f_j(x),$$

où chaque  $f_j$  est une densité et les **poids**  $w_j$  somment à 1.

Population divisée en sous-ensembles, chacun a sa propre distribution.

## Mélanges de lois

$$f(x) = \sum_{j=1}^{\infty} w_j f_j(x),$$

où chaque  $f_j$  est une densité et les **poids**  $w_j$  somment à 1.

Population divisée en sous-ensembles, chacun a sa propre distribution.

Mélange non dénombrable:

$$f(x) = \int_{\Theta} f_{\theta}(x) w(\theta) d\theta,$$

où  $\Theta \subseteq \mathbb{R}^d$ , et  $w$  une densité sur  $\Theta$ .



## Mélanges de lois

$$f(x) = \sum_{j=1}^{\infty} w_j f_j(x),$$

où chaque  $f_j$  est une densité et les **poids**  $w_j$  somment à 1.

Population divisée en sous-ensembles, chacun a sa propre distribution.

Mélange non dénombrable:

$$f(x) = \int_{\Theta} f_{\theta}(x) w(\theta) d\theta,$$

où  $\Theta \subseteq \mathbb{R}^d$ , et  $w$  une densité sur  $\Theta$ .

Pour générer  $X$ , générer d'abord  $Y$  selon les poids  $w_j$  ou la densité  $w$ , puis  $X$  sous la densité  $f_Y$ .

Exemple:  $X \sim \text{binomiale}(n, p)$  où  $p \sim \text{beta}(\alpha, \beta)$ .

Exemple:  $X \sim \text{Poisson}(\Lambda)$  où  $\Lambda \sim \text{Gamma}(\alpha, \beta)$ .