Les Processus Empiriques et leurs Applications

RAHMANI Naceur

Département de Mathématiques

2025-2026

Les Processus Empiriques

Les processus empiriques font référence à des modèles qui sont construits à partir de données réelles observées, contrairement aux modèles théoriques qui reposent sur des hypothèses idéalisées. Les processus empiriques sont largement utilisés pour analyser, prédire et modéliser des phénomènes dans des domaines comme l'économie, la finance, la biologie, les sciences sociales, etc.

Voici une explication détaillée des processus empiriques et de leurs applications dans plusieurs domaines :

Qu'est-ce qu'un Processus Empirique ?

Un processus empirique est un modèle qui repose sur l'observation directe des phénomènes. Contrairement aux processus théoriques, qui sont dérivés de principes théoriques ou de lois naturelles, un processus empirique est construit à partir de données observées dans le monde réel.
Exemple : Un modèle prédictif des tendances boursières qui utilise les données passées des prix pour prédire les prix futurs.

Qu'est-ce qu'un Processus Empirique ?

Les processus empiriques peuvent inclure **des modèles stochastiques** (qui incorporent de l'incertitude) ou **des modèles déterministes**. Ces modèles sont utilisés dans des situations où il est difficile de construire un modèle théorique complet ou lorsque les données sont trop complexes pour être expliquées uniquement par des modèles théoriques.

Remarque:

Modèle déterministe: les résultats sont entièrement déterminés par les paramètres d'entrée et il n'y a pas de place pour l'incertitude.

Modèle stochastique: l'incertitude est intégrée dans le modèle. Les résultats ne sont pas uniques et varient en fonction des variables aléatoires ou probabilistes.

2. Applications des Processus Empiriques
A. Économie et Finance

Les processus empiriques sont largement utilisés en économie et finance pour prédire des variables économiques ou des comportements financiers à partir des données passées.

Modélisation des Séries Temporelles : L'analyse des séries temporelles est une application majeure des processus empiriques. Les économistes utilisent des séries temporelles pour prédire des variables comme les taux de croissance économique, l'inflation, les taux d'intérêt, ou encore les prix des actifs financiers. Des modèles comme **ARIMA** (AutoRegressive Integrated Moving Average) sont souvent utilisés pour la prédiction des séries temporelles.

Modèles de Volatilité: Dans les marchés financiers, les processus empiriques sont utilisés pour modéliser la volatilité des actifs financiers. Les modèles **ARCH** (Autoregressive Conditional Heteroskedasticity) et **GARCH** (Generalized ARCH) permettent d'analyser et de prédire la volatilité des marchés financiers.

Prévisions Économiques: L'usage de données empiriques pour prédire la

2. Applications des Processus Empiriques
B. Sciences Sociales et Psychologie

Les processus empiriques sont également largement utilisés dans les sciences sociales et la psychologie pour étudier des comportements humains et sociaux.

Modélisation du Comportement Social: Les sociologues et les psychologues utilisent des processus empiriques pour étudier des phénomènes tels que les comportements de groupe, la dynamique familiale, ou les relations interpersonnelles. Par exemple, l'utilisation de modèles empiriques pour analyser les effets des politiques publiques sur la pauvreté ou l'éducation.

Études de Comportement et Prédiction: En psychologie, des processus empiriques peuvent être utilisés pour prédire des comportements humains, comme les réponses émotionnelles à certains stimuli ou les réactions des individus à différents types de thérapies.

2. Applications des Processus Empiriques
C. Biologie et Médecine

Les modèles empiriques sont cruciaux dans les sciences biologiques et médicales, où les phénomènes sont souvent trop complexes pour être modélisés uniquement par des théories.

Modélisation des Maladies Épidémiques: Par exemple, des modèles empiriques peuvent être utilisés pour étudier la propagation des maladies infectieuses comme la grippe ou COVID-19, en se basant sur des données d'infection et de transmission collectées dans le monde réel.

Études Cliniques et Prédictions de Réponse au Traitement: Dans les essais cliniques, des modèles empiriques peuvent être utilisés pour analyser l'efficacité des traitements médicaux, en tenant compte de facteurs comme l'âge, le sexe, et les antécédents médicaux des patients.

2. Applications des Processus Empiriques
D. Sciences de l'Environnement

Les modèles de processus empiriques sont aussi appliqués dans les sciences de l'environnement pour modéliser des phénomènes complexes comme les changements climatiques, la gestion des ressources naturelles, et la pollution.

Modélisation de la Pollution : Par exemple, des modèles empiriques sont utilisés pour prédire l'impact de la pollution de l'air sur la santé publique ou pour estimer les niveaux futurs de pollution dans différentes régions géographiques.

Gestion des Ressources Naturelles: Les modèles empiriques sont utilisés pour prédire les rendements agricoles, la disponibilité de l'eau, ou la gestion durable des forêts en fonction des observations passées des ressources naturelles.

2. Applications des Processus Empiriques

E. Ingénierie et Technologie

Les processus empiriques sont utilisés dans le domaine de l'ingénierie pour modéliser des systèmes complexes à partir des données observées.

Modélisation de Systèmes de Transport : Des modèles empiriques peuvent être utilisés pour étudier les flux de trafic, prédire les embouteillages, et optimiser les systèmes de transport en fonction des données sur les déplacements des véhicules.

Conception et Amélioration de Produits : Les ingénieurs utilisent des données empiriques pour améliorer la conception de produits ou de systèmes. Par exemple, dans l'aéronautique, des tests empiriques sont effectués pour analyser la performance des matériaux ou des moteurs.

3. Méthodes Utilisées dans les Processus Empiriques

Les processus empiriques utilisent diverses techniques statistiques et économétriques pour analyser les données et établir des modèles :

Régression Linéaire et Non-linéaire: Utilisées pour modéliser des relations simples ou complexes entre les variables.

Méthodes de Séries Temporelles: Comme ARIMA et GARCH pour analyser les tendances et la volatilité des séries chronologiques.

3. Méthodes Utilisées dans les Processus Empiriques

Méthodes de Simulation: Des techniques comme la Monte Carlo sont utilisées pour simuler des scénarios et analyser les incertitudes.
Échantillonnage et Estimation: Utilisation de techniques d'échantillonnage et d'estimation pour déduire les caractéristiques de la population à partir d'un échantillon de données.
Modèles Stochastiques: Utilisés pour modéliser des phénomènes incertains ou aléatoires.

4.Fonction de Répartition Empirique (FRE)

La fonction de répartition empirique (ou fonction de distribution empirique, FRE) est une méthode statistique utilisée pour estimer la fonction de répartition cumulative d'une variable aléatoire à partir d'un échantillon de données observées. Elle donne la proportion des observations qui sont inférieures ou égales à une valeur donnée. Cette fonction est une estimation non paramétrique de la fonction de répartition théorique.

4.Fonction de Répartition Empirique (FRE)

Définition

Soit un échantillon de données indépendantes $X_1, X_2, ..., X_n$ provenant d'une population avec une distribution inconnue. La fonction de répartition empirique $\hat{F}_n(x)$ pour une valeur x est définie comme la fraction d'observations dans l'échantillon qui sont inférieures ou égales à x. Formellement:

$$\hat{\mathbf{F}}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{X_i \leq x\}}$$

4.Fonction de Répartition Empirique (FRE)

Définition

Où :

- $\mathbf{1}_{\{X_i \leq x\}}$: est la fonction indicatrice qui vaut 1 si $\{X_i \leq x\}$, et 0 sinon.
- n est la taille de l'échantillon.
- x est la valeur pour laquelle on souhaite estimer la fonction de répartition.

Ainsi, $\hat{\mathbf{F}}_n(x)$ représente la proportion d'observations X_i dans l'échantillon qui sont inférieures ou égales à x.

4.Fonction de Répartition Empirique (FRE)

Propriétés de la Fonction de Répartition Empirique

a. Monotonie

La fonction de répartition empirique est croissante:

- Si $x_1 < x_{2,}$ alors $\hat{\mathbf{F}}_n(x_1) \leq \hat{\mathbf{F}}_n(x_2)$

Cela signifie que plus x est grand, plus la proportion des données inférieures ou égales à x augmente.

4.Fonction de Répartition Empirique (FRE)

Propriétés de la Fonction de Répartition Empirique

b. Discontinuité

La FRE est une fonction en escaliers, contrairement à la fonction de répartition théorique qui est continue. À chaque observation X_i dans l'échantillon, il y a un saut de la FRE. Si X_i est une observation unique dans l'échantillon, la fonction aura un saut de $\frac{1}{n}$ à cet endroit.

4.Fonction de Répartition Empirique (FRE)

Propriétés de la Fonction de Répartition Empirique

c Convergence

La **FRE** converge vers la fonction de répartition théorique $F(x)$ Lorsque la taille de l'échantillon n devient grande. Cela est démontré par le théorème de **Glivenko-Cantelli**, qui affirme que:

$$\hat{\mathbf{F}}_n(x) \overset{n \to \infty}{\longrightarrow} F(x) \quad \textbf{en probabilité}$$

Cela signifie que la FRE devient une approximation de plus en plus précise de la fonction de répartition théorique à mesure que le nombre d'observations augmente.

4.Fonction de Répartition Empirique (FRE)

Propriétés de la Fonction de Répartition Empirique

d.Convergence presque sûre (Loi des grands nombres):
La fonction de répartition empirique $\hat{\mathbf{F}}_n(x)$ converge presque sûrement vers la fonction de répartition théorique $F(x)$ de la variable aléatoire X. Cela peut être exprimé comme suit:

$$\hat{\mathbf{F}}_n(x) \overset{p.s}{\longrightarrow} F(x) \quad \textit{lorsque } n \longrightarrow \infty$$

Cela signifie qu'avec une probabilité de 1, la fonction empirique converge vers la fonction de répartition théorique pour chaque x.

4.Fonction de Répartition Empirique (FRE)

Propriétés de la Fonction de Répartition Empirique

e.Convergence uniforme

La convergence de la fonction de répartition empirique vers la fonction de répartition théorique se fait aussi de manière uniforme. Cela signifie que la différence entre $\hat{\mathbf{F}}_n(x)$ et $F(x)$ est uniformément petite à mesure que n augmente. Cette convergence uniforme peut être étudiée en termes de **convergence en probabilité** ou de **convergence presque sûre**, selon les conditions.

4.Fonction de Répartition Empirique (FRE)

Propriétés de la Fonction de Répartition Empirique

f.Convergence en distribution (Théorème de Donsker) :

Le théorème de Donsker (ou théorème de la fonction de répartition empirique) affirme que, après une normalisation appropriée, la fonction de répartition empirique $\hat{\mathbf{F}}_n(x)$ converge en distribution vers un processus de Brownien. Plus précisément, si on considère la distance de Kolmogorov D_n entre la fonction de répartition empirique et la fonction de répartition théorique, alors :

$$\sqrt{n} \sup_{x \in \mathbb{R}} \left| \hat{\mathbf{F}}_n(x) - F(x) \right| \xrightarrow{d} \sup_{x \in \mathbb{R}} |B(x)|$$

où $B(x)$ est un processus de Brownien standard.

5.Fonction des Quantiles

Un quantile est un point qui divise une distribution de données en sous-ensembles de probabilités égales. Il représente la valeur de la variable aléatoire à laquelle une proportion donnée des observations sont inférieures ou égales. Par exemple, le **médian** est le quantile qui divise la distribution en deux parties égales (**50 %** des observations sont inférieures ou égales à la médiane).

5.Fonction des Quantiles

La fonction des quantiles est une fonction **inverse** de **la fonction de répartition cumulative (FRC)**, qui donne la valeur du quantile associé à un certain niveau de probabilité.

5.Fonction des Quantiles
Définition Formelle

La fonction des quantiles, notée $Q(p)$, est définie comme l'inverse de la fonction de répartition $F(x)$, où p est une probabilité entre 0 et 1, et x est une valeur de la variable aléatoire. Plus précisément:

$$Q(p) = F^{-1}(p)$$

Cela signifie que pour un quantile p, $Q(p)$ est la valeur de x telle que:

$$F(Q(p)) = P$$

5.Fonction des Quantiles

Types de Quantiles

Les quantiles peuvent être classés en plusieurs types, en fonction de la proportion p qu'ils représentent:

Médiane (quantile 0.5): Divise la distribution en deux parties égales. 50 % des données sont inférieures ou égales à la médiane, et 50 % sont supérieures.

Quartiles: Divisent les données en quatre parties égales.

Le premier quartile (Q_1) correspond au 25e percentile (25 % des données sont inférieures ou égales à cette valeur).

Le deuxième quartile (Q_2) correspond à la médiane (50e percentile).

Le troisième quartile (Q_3) correspond au 75e percentile (75 % des données sont inférieures ou égales à cette valeur).

Déciles : Divisent les données en dix parties égales, chaque décile représentant 10 % des données.

Percentiles: Divisent les données en 100 parties égales. Le 10e percentile correspond à la valeur en dessous de laquelle se trouvent 10 % des

6.Quantile Empirique

Le quantile empirique est l'estimation d'un quantile basé sur un échantillon de données. Il est calculé de manière non paramétrique et est utilisé pour estimer les quantiles d'une distribution de données lorsque la distribution théorique est inconnue.

6.Quantile Empirique

Calcul du Quantile Empirique

Soit un échantillon de données $X_1, X_2, ..., X_n$ (de taille n) tiré d'une population. Pour calculer le quantile empirique $Q_n(p)$, voici les étapes à suivre:

1. **Trier les données**: On commence par trier les données dans l'ordre croissant:

$$X_{(1)} \leq X\left(_2\right) \leq ... \leq X\left(_n\right)$$

6.Quantile Empirique

Calcul du Quantile Empirique

Soit un échantillon de données $X_1, X_2, ..., X_n$ (de taille n) tiré d'une population. Pour calculer le quantile empirique $Q_n(p)$, voici les étapes à suivre:

1. **Trier les données**: On commence par trier les données dans l'ordre croissant:

$$X_{(1)} \leq X\left(_2\right) \leq ... \leq X\left(_n\right)$$

2. **Trouver la position du quantile**: La position k_p du quantile empirique pour un quantile p avec $(0 \leq p \leq 1)$ est donnée par:

$$k_p = \lceil p * n \rceil$$

   où $\lceil x \rceil$ est la fonction "plafond", qui donne le plus petit entier supérieur ou égal à x

6.Quantile Empirique
Calcul du Quantile Empirique

3 **Extraire la valeur du quantile empirique**:
La valeur du quantile empirique est la donnée située à la position k_p dans l'échantillon trié:

$$Q_n(p) = Q_{emp}(p) = X_{(k_p)}$$

Par exemple, pour un échantillon de taille 10 et $p = 0.25$, la position du premier quartile serait:

$$k_{0.25} = \lceil 0.25 * 10 \rceil = 3$$

Ainsi,

$$Q_{10}(0.25) = Q_{emp}(0.25) = X_3$$

6.Quantile Empirique

Relation avec la fonction de répartition empirique

La fonction de répartition empirique $\hat{\mathbf{F}}_n(x)$ est étroitement liée aux quantiles empiriques. En effet, un quantile empirique $Q_n(p)$ est la valeur x pour laquelle la fonction de répartition empirique $\hat{\mathbf{F}}_n(x)$atteint une probabilité p

$$Q_n(p) = \hat{\mathbf{F}}_n^{-1}(p)$$

Les quantiles permettent donc de résumer de manière concise la distribution empirique des données et sont utilisés pour visualiser la dispersion et la forme de la distribution, notamment à travers des diagrammes en boîte (boxplots).

6.Quantile Empirique

Convergence des Quantiles Empiriques

La convergence des quantiles empiriques désigne le phénomène par lequel, à mesure que **la taille de l'échantillon augmente**, les quantiles empiriques d'un échantillon convergent vers les quantiles théoriques de la distribution sous-jacente de la variable aléatoire. Cette convergence est fondamentale dans le cadre de l'estimation non paramétrique et elle est garantie par certains résultats théoriques, notamment **la loi des grands nombres** et **le théorème de la convergence des quantiles empiriques.**

6.Quantile Empirique

1.Convergence en Probabilité des Quantiles Empiriques

La convergence en probabilité des quantiles empiriques est définie comme suit:

$$\mathbf{Q}_n\left(p\right) \xrightarrow{p} Q\left(p\right) \quad lorsque \quad n \longrightarrow \infty$$

Cela signifie que, pour un grand échantillon, la valeur du quantile empirique $\mathbf{Q}_n\left(p\right)$ va converger vers le quantile théorique $Q\left(p\right)$ de manière probabiliste.

6.Quantile Empirique

1.Convergence en Probabilité des Quantiles Empiriques

$$\mathbf{Q}_n(p) \xrightarrow{p} Q(p) \quad \textit{lorsque} \quad n \longrightarrow \infty$$

Explication:
$\mathbf{Q}(p)$: est le quantile théorique basé sur la fonction de répartition de la population.
$\mathbf{Q}_n(p)$: est le quantile empirique estimé à partir de l'échantillon.
À mesure que la taille de l'échantillon n augmente, la fonction de répartition empirique $\hat{\mathbf{F}}_n(p)$ devient une meilleure approximation de la fonction de répartition réelle $F(x)$, et donc $\mathbf{Q}_n(p)$ se rapproche de $\mathbf{Q}(p)$

6.Quantile Empirique

2. Convergence Forte des Quantiles Empiriques

La convergence forte des quantiles empiriques signifie que les quantiles empiriques convergent presque sûrement vers les quantiles théoriques lorsque la taille de l'échantillon tend vers l'infini. Formulée de manière plus précise, cela se traduit par la relation suivante:

$$P\left(\lim_{n\to+\infty}\left|\mathbf{Q}_n\left(p\right)-Q\left(p\right)\right|=0\right)=1$$

6.Quantile Empirique
Application

Les quantiles empiriques sont utilisés dans de nombreuses situations statistiques et analytiques :

Boxplot : Les quantiles sont utilisés pour tracer des boîtes et des moustaches (boxplots), qui permettent de visualiser la dispersion des données et de détecter les valeurs aberrantes.

Tests statistiques: Les quantiles empiriques sont utilisés dans les tests non paramétriques pour comparer des échantillons ou tester des hypothèses sur les distributions des données.

Analyse de la répartition des données: Calculer les quantiles permet de mieux comprendre la distribution des données (par exemple, connaître les percentiles dans une population).