



**HAL**  
open science

# **Théorème d'approximation de Stone-Weierstrass en dimension finie sur un compact : Artiffaille sur les probabilités et les polynômes de Bernstein**

David Pigeon

## ► To cite this version:

David Pigeon. Théorème d'approximation de Stone-Weierstrass en dimension finie sur un compact : Artiffaille sur les probabilités et les polynômes de Bernstein. 2025. <hal-04939275>

**HAL Id: hal-04939275**

**<https://hal.science/hal-04939275v1>**

Preprint submitted on 10 Feb 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

# Théorème d'approximation de Stone-Weierstrass en dimension finie sur un compact : Artifaille sur les probabilités et les polynômes de Bernstein

David Pigeon\*

10 février 2025

**Mots clés** : polynômes de Bernstein, approximation uniforme de Stone-Weierstrass, probabilités, inégalité de Bienaymé-Tchebychev.

## Résumé

On donne une preuve simple du théorème d'approximation de Stone-Weierstrass en dimension  $d$  d'une fonction définie sur un compact  $K$  de  $\mathbb{R}^d$ . On démontre d'abord le résultat dans le cas  $K := [0, 1]^d$ , puis par un résultat de prolongement par continuité à tout l'espace  $\mathbb{R}^d$  d'une fonction définie sur un compact quelconque, on en déduit le résultat général.

## Table des matières

<b>1</b>	<b>Cas sur l'hypercube <math>[0, 1]^d</math></b>	<b>3</b>
1.1	Fonctions polynômiales simplifiées de Bernstein . . . . .	3
1.2	Théorème . . . . .	4
<b>2</b>	<b>Cas sur un compact</b>	<b>8</b>
2.1	Prolongement continue . . . . .	8
2.2	Théorème . . . . .	11
<b>A</b>	<b>Artifaille en probabilités</b>	<b>13</b>
A.1	Existence de variables aléatoires binomiales indépendantes . . . . .	13
A.2	Deux inégalités . . . . .	14
<b>B</b>	<b>Polynômes de Bernstein généralisés</b>	<b>18</b>
B.1	Cas sur l'hypercube $[0, 1]^d$ . . . . .	18
B.2	Cas sur un pavé . . . . .	19
	<b>Références</b>	<b>21</b>

---

\*<https://sites.google.com/view/davidpigeon/>

On fixe un entier  $d \in \mathbb{N}_1 := \mathbb{N} - \{0\}$ .

On munit  $\mathbb{R}^d$  de la **norme infinie** :

$$\forall \mathbf{x} := (x_1, \dots, x_d), \mathbf{y} := (y_1, \dots, y_d) \in [0, 1]^d, \|\mathbf{x} - \mathbf{y}\|_\infty := \max_{i \in [1, d]} |x_i - y_i|.$$

On munit l'ensemble des fonctions continues  $\mathcal{C}(K, \mathbb{R})$  sur un compact  $K$  de  $\mathbb{R}^d$  de la **norme infinie** :

$$\forall f \in \mathcal{C}(K, \mathbb{R}), \|f\|_{\infty, K} := \max_{\mathbf{x} \in [0, 1]^d} |f(\mathbf{x})|.$$

### Bibliographie commentée.

- Dans [3] et [6], on trouvera des informations complémentaires sur la preuve du théorème de Weierstrass dans l'hypercube  $[0, 1]^d$  et une définition plus générale des polynômes de Bernstein à plusieurs variables.
- Dans [1] et [2], on trouvera des informations complémentaires sur le prolongement sur  $\mathbb{R}^d$  d'une fonction continue définie sur un compact. La preuve du résultat de prolongation d'une fonction continue se trouve aussi dans [1].
- Dans [4] et [5], on trouvera les preuves originales de Stone et Weierstrass.

Le présent texte vise à donner une preuve simple dans le cas d'un compact de  $\mathbb{R}^d$  en utilisant une version généralisée des polynômes de Bernstein.

## 1 Cas sur l'hypercube $[0, 1]^d$

La preuve usuelle de l'approximation de Weierstrass en dimension 1 s'adapte bien en dimension  $d$ .

### 1.1 Fonctions polynômiales simplifiées de Bernstein

On commence par une définition simplifiée des fonctions polynômiales de Bernstein généralisées (pour une définition plus générale des fonctions polynômiales de Bernstein, voir la section B de l'annexe).

#### Définition 1.1

Soit  $f \in \mathcal{C}([0, 1]^d, \mathbb{R})$  et  $n \in \mathbb{N}_1$ . La **fonction polynômiale simplifiée de Bernstein de paramètres  $d$  et  $n$  associée à  $f$**  est définie par :

$$B_{n,f}^{(d)} : [0, 1]^d \longrightarrow \mathbb{R}$$

$$(x_1, \dots, x_d) \longmapsto \sum_{(k_1, \dots, k_d) \in \llbracket 0, n \rrbracket^d} f\left(\frac{k_1}{n}, \dots, \frac{k_d}{n}\right) \prod_{i=1}^d \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i}$$

On a alors les trois exemples simples suivants.

#### Exemple 1.2

(1) **Cas  $d := 1$ .** En dimension 1, on retrouve les polynômes usuels de Bernstein :

$$\forall x \in [0, 1], B_{n,f}(x) := B_{n,f}^{(1)}(x) = \sum_{k=0}^n f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}.$$

(2) **Cas  $f := 1$  sur  $[0, 1]^d$ .** On a pour tout  $(x_1, \dots, x_d) \in [0, 1]^d$  :

$$\begin{aligned} B_{n,1}^{(d)}(x_1, \dots, x_d) &= \sum_{(k_1, \dots, k_d) \in \llbracket 0, n \rrbracket^d} \prod_{i=1}^d \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} \\ &= \prod_{i=1}^d \left( \sum_{k_i \in \llbracket 0, n \rrbracket} \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} \right) \\ &= \prod_{i=1}^d 1 \\ &= 1 \end{aligned}$$

(3) **Cas  $f$  à variables séparées sur  $[0, 1]^d$ .** Supposons qu'il existe des fonctions continues  $f_1, \dots, f_d \in \mathcal{C}([0, 1], \mathbb{R})$  telles que pour tout  $(x_1, \dots, x_d) \in [0, 1]^d$  :

$$f(x_1, \dots, x_d) = \prod_{i=1}^d f_i(x_i).$$

On a pour tout  $(x_1, \dots, x_d) \in [0, 1]^d$  :

$$\begin{aligned} B_{n,f}^{(d)}(x_1, \dots, x_d) &= \sum_{(k_1, \dots, k_d) \in \llbracket 0, n \rrbracket^d} f\left(\frac{k_1}{n}, \dots, \frac{k_d}{n}\right) \prod_{i=1}^d \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} \\ &= \sum_{(k_1, \dots, k_d) \in \llbracket 0, n \rrbracket^d} \prod_{i=1}^d f_i\left(\frac{k_i}{n}\right) \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} \end{aligned}$$

$$\begin{aligned}
&= \prod_{i=1}^d \left( \sum_{k_i=0}^n f_i \left( \frac{k_i}{n} \right) \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} \right) \\
&= \prod_{i=1}^d B_{n, f_i}(x_i)
\end{aligned}$$

## 1.2 Théorème

On a alors le théorème de Stone-Weierstrass dans la cas de l'hypercube  $[0,1]^d$ .

### Théorème 1.3: Stone-Weierstrass - cas de l'hypercube

Soit  $f \in \mathcal{C}([0,1]^d, \mathbb{R})$ . Alors la suite  $(B_{n,f}^{(d)})_n$  converge uniformément vers  $f$  sur  $[0,1]^d$  i.e. on a :

$$\lim_{n \rightarrow \infty} \|B_{n,f}^{(d)} - f\|_{\infty, [0,1]^d} = 0.$$

*Démonstration.* Posons :

$$K := [0,1]^d.$$

Soit  $\varepsilon > 0$ . Par le théorème de Heine, la fonction  $f$  est uniformément continue sur  $K$ . Donc il existe  $\alpha > 0$  tel que :

$$\forall \mathbf{x}, \mathbf{y} \in [0,1]^d, \|\mathbf{x} - \mathbf{y}\|_{\infty} \leq \alpha \implies |f(\mathbf{x}) - f(\mathbf{y})| \leq \varepsilon$$

Posons :

$$N := \left\lfloor \frac{d\|f\|_{\infty, K}}{2\varepsilon\alpha^2} \right\rfloor + 1.$$

Soit  $\mathbf{x} := (x_1, \dots, x_d) \in K^d$ ,  $n \in \mathbb{N}_1$  tel que  $n \geq N$  et  $\mathbf{X} := (X_1, \dots, X_d)$  une suite de  $d$  variables aléatoires indépendantes de lois binomiales  $\mathcal{B}_{n, x_1}, \dots, \mathcal{B}_{n, x_d}$  respectivement sur un espace probabilisé  $(\Omega, \mathbb{P})$  (voir le théorème A.2 pour l'existence d'une telle suite de variables aléatoires). Posons :

$$\Gamma_{\alpha} := \left\{ \mathbf{k} \in \llbracket 0, n \rrbracket^d, \left\| \frac{\mathbf{k}}{n} - \mathbf{x} \right\|_{\infty} \leq \alpha \right\}, \quad \Gamma_{\alpha}^c := \llbracket 0, n \rrbracket^d - \Gamma_{\alpha}.$$

Par indépendance des  $X_1, \dots, X_d$ , on a pour tout  $\mathbf{k} := (k_1, \dots, k_d) \in \llbracket 0, n \rrbracket^d$  :

$$\prod_{i=1}^d \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} = \prod_{i=1}^d \mathbb{P}(X_i = k_i) = \mathbb{P}(\mathbf{X} = \mathbf{k}).$$

Ainsi on a :

$$\begin{aligned}
|B_{n,f}^{(d)}(\mathbf{x}) - f(\mathbf{x})| &\stackrel{(1)}{=} \left| \sum_{\mathbf{k} \in \llbracket 0, n \rrbracket^d} f\left(\frac{\mathbf{k}}{n}\right) \mathbb{P}(\mathbf{X} = \mathbf{k}) + f(\mathbf{x}) \sum_{\mathbf{k} \in \llbracket 0, n \rrbracket^d} \mathbb{P}(\mathbf{X} = \mathbf{k}) \right| \\
&\stackrel{(2)}{=} \left| \sum_{\mathbf{k} \in \llbracket 0, n \rrbracket^d} \left( f\left(\frac{\mathbf{k}}{n}\right) - f(\mathbf{x}) \right) \mathbb{P}(\mathbf{X} = \mathbf{k}) \right| \\
&\stackrel{(3)}{=} \left| \sum_{\mathbf{k} \in \Gamma_{\alpha}} \left( f\left(\frac{\mathbf{k}}{n}\right) - f(\mathbf{x}) \right) \mathbb{P}(\mathbf{X} = \mathbf{k}) + \sum_{\mathbf{k} \in \Gamma_{\alpha}^c} \left( f\left(\frac{\mathbf{k}}{n}\right) - f(\mathbf{x}) \right) \mathbb{P}(\mathbf{X} = \mathbf{k}) \right| \\
&\stackrel{(4)}{\leq} \sum_{\mathbf{k} \in \Gamma_{\alpha}} \left| f\left(\frac{\mathbf{k}}{n}\right) - f(\mathbf{x}) \right| \mathbb{P}(\mathbf{X} = \mathbf{k}) + \sum_{\mathbf{k} \in \Gamma_{\alpha}^c} \left| f\left(\frac{\mathbf{k}}{n}\right) - f(\mathbf{x}) \right| \mathbb{P}(\mathbf{X} = \mathbf{k}) \\
&\stackrel{(5)}{\leq} \varepsilon \sum_{\mathbf{k} \in \Gamma_{\alpha}} \mathbb{P}(\mathbf{X} = \mathbf{k}) + 2\|f\|_{\infty, K} \sum_{\mathbf{k} \in \Gamma_{\alpha}^c} \mathbb{P}(\mathbf{X} = \mathbf{k})
\end{aligned}$$

$$\begin{aligned}
 (6) &\leq \varepsilon \sum_{\mathbf{k} \in \llbracket 0, n \rrbracket^d} \mathbb{P}(\mathbf{X} = \mathbf{k}) + 2\|f\|_{\infty, K} \mathbb{P}(\|\mathbf{X} - n\mathbf{x}\|_{\infty} > n\alpha) \\
 (7) &= \varepsilon + 2\|f\|_{\infty, K} \mathbb{P}(\|\mathbf{X} - \mathbb{E}(\mathbf{X})\|_{\infty} > n\alpha) \\
 (8) &= \varepsilon + 2\|f\|_{\infty, K} \mathbb{P}\left(\max_{i \in \llbracket 1, d \rrbracket} |X_i - \mathbb{E}(X_i)| > n\alpha\right) \\
 (9) &= \varepsilon + 2\|f\|_{\infty, K} \mathbb{P}\left(\bigcup_{i=1}^d \{|X_i - \mathbb{E}(X_i)| > n\alpha\}\right) \\
 (10) &\leq \varepsilon + 2\|f\|_{\infty, K} \left(\sum_{i=1}^d \mathbb{P}(|X_i - \mathbb{E}(X_i)| > n\alpha)\right) \\
 (11) &\leq \varepsilon + 2\|f\|_{\infty, K} \left(\sum_{i=1}^d \frac{\mathbb{V}(X_i)}{n^2\alpha^2}\right) \\
 (12) &= \varepsilon + 2\|f\|_{\infty, K} \left(\sum_{i=1}^d \frac{nx_i(1-x_i)}{n^2\alpha^2}\right) \\
 (13) &\leq \varepsilon + 2\|f\|_{\infty, K} \left(\sum_{i=1}^d \frac{n/4}{n^2\alpha^2}\right) \\
 (14) &= \varepsilon + \frac{d\|f\|_{\infty, K}}{2n\alpha^2} \\
 (15) &\leq 2\varepsilon
 \end{aligned}$$

Donc on a pour tout  $n \in \mathbb{N}_1$  tel que  $n \geq N$  :

$$\|B_{n,f}^{(d)} - f\|_{\infty, K} \leq 2\varepsilon.$$

Ainsi on a :

$$\lim_{n \rightarrow \infty} \|B_{n,f}^{(d)} - f\|_{\infty, K} = 0.$$

Voici quelques explications pour chaque égalité ou inégalité :

(1) la famille  $\{\mathbf{X} = \mathbf{k}, \mathbf{k} \in \llbracket 0, n \rrbracket^d\}$  forme un système complet d'événements et donc :

$$\sum_{\mathbf{k} \in \llbracket 0, n \rrbracket^d} \mathbb{P}(\mathbf{X} = \mathbf{k}) = 1;$$

(2) en regroupant et factorisant ;

(3) en faisant le partitionnage :

$$\llbracket 0, n \rrbracket^d = \Gamma_{\alpha} \sqcup \Gamma_{\alpha}^c;$$

(4) en utilisant l'inégalité triangulaire ;

(5) on a pour tout  $\mathbf{k} \in \Gamma_{\alpha}$  (car  $\mathbf{k}/n \in [0, 1]^d$  et  $|\mathbf{k}/n - \mathbf{x}| \leq \alpha$ ) :

$$\left|f\left(\frac{\mathbf{k}}{n}\right) - f(\mathbf{x})\right| \leq \varepsilon$$

et pour tout  $\mathbf{k} \in \Gamma_{\alpha}^c$  (car  $\mathbf{k}/n \in [0, 1]^d$ ) :

$$\left|f\left(\frac{\mathbf{k}}{n}\right) - f(\mathbf{x})\right| \leq \left|f\left(\frac{\mathbf{k}}{n}\right)\right| + |f(\mathbf{x})| \leq 2 \max_{\mathbf{x} \in [0, 1]^d} |f(\mathbf{x})| = 2\|f\|_{\infty, K};$$

(6) on a :

$$\bigsqcup_{\mathbf{k} \in \Gamma_{\alpha}^c} \{\mathbf{X} = \mathbf{k}\} = \left\{\left\|\frac{\mathbf{X}}{n} - \mathbf{x}\right\|_{\infty} > \alpha\right\} = \{\|\mathbf{X} - n\mathbf{x}\|_{\infty} > n\alpha\};$$

(7) pour tout  $i \in \llbracket 1, d \rrbracket$  l'espérance de  $X_i$  vaut  $nx_i$  car :

$$\begin{aligned}
\mathbb{E}(X_i) &= \sum_{k=0}^n k \mathbb{P}(X_i = k) \\
&= \sum_{k=0}^n k \binom{n}{k} x_i^k (1-x_i)^{n-k} \\
&= \sum_{k=0}^n n \binom{n-1}{k-1} x_i^k (1-x_i)^{n-k} \\
&= nx_i \sum_{k=0}^{n-1} \binom{n-1}{k} x_i^{k-1} (1-x_i)^{n-1-(k-1)} \\
&= nx_i (x_i + (1-x_i))^{n-1} \\
&= nx_i
\end{aligned}$$

et donc l'espérance de  $\mathbf{X}$  vaut  $n\mathbf{x}$  (voir aussi les notations A.4 pour la définition de l'espérance de  $\mathbf{X}$ );

(8) par définition de la norme infini sur  $\mathbb{R}^d$ , on a :

$$\|\mathbf{X} - \mathbb{E}(\mathbf{X})\|_\infty = \max_{i \in \llbracket 1, d \rrbracket} |X_i - \mathbb{E}(X_i)|;$$

(9) on a les égalités d'évènements :

$$\begin{aligned}
\left\{ \max_{i \in \llbracket 1, d \rrbracket} |X_i - \mathbb{E}(X_i)| > n\alpha \right\} &= \{ \omega \in \Omega, \exists i \in \llbracket 1, d \rrbracket, |X_i(\omega) - \mathbb{E}(X_i)| > n\alpha \} \\
&= \bigcup_{i=1}^d \{ |X_i - \mathbb{E}(X_i)| > n\alpha \};
\end{aligned}$$

(10) la probabilité d'une union d'évènements est inférieure à la somme des probabilités de ces évènements :

$$\mathbb{P}\left(\bigcup_{i=1}^d \{ |X_i - \mathbb{E}(X_i)| > n\alpha \}\right) \leq \sum_{i=1}^d \mathbb{P}(|X_i - \mathbb{E}(X_i)| > n\alpha);$$

(11) en appliquant  $d$  fois l'inégalité de Bienaymé-Tchebychev pour les variables aléatoires  $X_1, \dots, X_d$  (voir le point (ii) du théorème A.3), on a pour tout  $i \in \llbracket 1, d \rrbracket$  :

$$\mathbb{P}(|X_i - \mathbb{E}(X_i)| > n\alpha) \leq \frac{\mathbb{V}(X_i)}{n^2\alpha^2};$$

(on peut aussi appliquer directement le corollaire A.5 pour les étapes (8) à (11))

(12) pour tout  $i \in \llbracket 1, d \rrbracket$ , la variance de la loi binomiale  $X_i$  vaut  $nx_i(1-x_i)$  car on a :

$$\begin{aligned}
\mathbb{V}(X_i) &= \mathbb{E}(X_i^2) - \mathbb{E}(X_i)^2 \\
&= \sum_{k=0}^n k^2 \binom{n}{k} x_i^k (1-x_i)^{n-k} - (nx_i)^2 \\
&= \sum_{k=1}^n kn \binom{n-1}{k-1} x_i^k (1-x_i)^{n-k} - (nx_i)^2 \\
&= nx_i \sum_{k=0}^{n-1} (k+1) \binom{n-1}{k} x_i^k (1-x_i)^{n-1-k} - (nx_i)^2 \\
&= nx_i \sum_{k=0}^{n-1} \binom{n-1}{k} x_i^k (1-x_i)^{n-1-k} + nx_i \sum_{k=0}^{n-1} k \binom{n-1}{k} x_i^k (1-x_i)^{n-1-k} - (nx_i)^2 \\
&= nx_i - (nx_i)^2 + nx_i \sum_{k=1}^{n-1} k \binom{n-1}{k} x_i^k (1-x_i)^{n-1-k}
\end{aligned}$$

$$\begin{aligned} &= nx_i - (nx_i)^2 + nx_i \sum_{k=1}^{n-1} (n-1) \binom{n-2}{k-1} x_i^k (1-x_i)^{n-1-k} \\ &= nx_i - (nx_i)^2 + n(n-1)x_i^2 \sum_{k=0}^{n-2} \binom{n-2}{k} x_i^k (1-x_i)^{n-2-k} \\ &= nx_i - (nx_i)^2 + n(n-1)x_i^2 \\ &= nx_i(1-x_i); \end{aligned}$$

(13) la courbe  $t \mapsto t(1-t)$  est une parabole tournée vers le bas s'annulant en 0 et 1, donc son maximum est au milieu du segment  $[0,1]$  en  $t := 1/2$  i.e. :

$$\forall t \in [0,1], t(1-t) \leq \frac{1}{2} \left(1 - \frac{1}{2}\right) = \frac{1}{4}.$$

(14) on a :

$$\sum_{i=1}^d \frac{n/4}{n^2 \alpha^2} = \frac{d}{4n\alpha^2}$$

(15) comme :

$$n \geq N > \frac{d\|f\|_{\infty,K}}{2\varepsilon\alpha^2},$$

on a :

$$\frac{d\|f\|_{\infty,K}}{2n\alpha^2} \leq \frac{d\|f\|_{\infty,K}}{2N\alpha^2} < \frac{d\|f\|_{\infty,K}}{2(d\|f\|_{\infty,K}/(2\varepsilon\alpha^2))\alpha^2} = \varepsilon.$$

□

## 2 Cas sur un compact

### 2.1 Prolongement continue

Commençons par la définition d'un compact. On note  $\|\cdot\|$  une norme quelconque sur  $\mathbb{R}^d$ . Pour tout  $\mathbf{x} \in \mathbb{R}^d$  et tout  $r \in \mathbb{R}_+^*$ , on note  $\mathbb{B}(\mathbf{x}, r)$  la **boule ouverte de centre  $\mathbf{x}$  et de rayon  $r$  pour la norme  $\|\cdot\|$**  :

$$\mathbb{B}(x, r) := \{y \in \mathbb{R}^d, \|x - y\| < r\}.$$

#### Définition 2.1: Axiome de Borel-Lebesgue

Soit un sous-ensemble  $E$  de  $\mathbb{R}^d$ . On dit que  $E$  est **compact** si pour tout recouvrement d'ouverts de  $E$  possède un sous-recouvrement fini *i.e.* pour tout ensemble  $I$  et toute famille d'ouverts  $(U_i)_{i \in I}$  de  $\mathbb{R}^d$  telle que :

$$E \subset \bigcup_{i \in I} U_i$$

il existe un sous-ensemble fini  $I'$  de  $I$  tel que :

$$E \subset \bigcup_{i \in I'} U_i.$$

On a alors les équivalences usuelles suivantes en dimension finie pour  $E$  un sous-ensemble de  $\mathbb{R}^d$  :

- (i)  $E$  est compact ;
- (ii)  $E$  est fermé et borné.

On fixe pour la suite  $K$  un compact de  $\mathbb{R}^d$ . Commençons par un lemme usuel.

#### Lemma 2.2

Il existe un sous-ensemble  $J$  de  $\mathbb{N}$  au plus dénombrable et un ensemble  $\mathcal{S} := \{s_k\}_{k \in J}$  dense et au plus dénombrable dans  $K$ .

*Démonstration.* On rappelle que la densité de  $\mathcal{S}$  dans  $K$  signifie que pour tout  $\varepsilon \in \mathbb{R}_+^*$  et pour tout  $s \in K$ , la boule ouverte  $\mathbb{B}(s, \varepsilon)$  de centre  $s$  et de rayon  $\varepsilon$  rencontre  $\mathcal{S}$  *i.e.* il existe  $k \in J$  tel que  $s_k \in \mathbb{B}(s, \varepsilon)$ .

**Construction de  $\mathcal{S}$ .** Soit  $r \in \mathbb{Q}_+^*$ . Considérons un recouvrement par les boules ouvertes  $B(s, r)$  avec  $s \in K$  :

$$K \subset \bigcup_{s \in K} \mathbb{B}(s, r).$$

Par compacité, il existe un sous-recouvrement fini de  $K$  par ces boules *i.e.* il existe une partie finie  $K_r$  de  $K$  telle que :

$$(*) : K \subset \bigcup_{s \in K_r} \mathbb{B}(s, r).$$

Comme  $\mathbb{Q}_+^*$  est dénombrable et comme l'union dénombrable d'ensembles finis est au plus dénombrable, l'union :

$$\mathcal{S} := \bigcup_{r \in \mathbb{Q}_+^*} K_r$$

est au plus dénombrable. Donc il existe une partie  $J$  de  $\mathbb{N}$  au plus dénombrable et une bijection  $\phi_{\mathcal{S}}$  de  $J$  vers  $\mathcal{S}$  *i.e.*

$$\mathcal{S} = \{s_k := \phi_{\mathcal{S}}(k)\}_{k \in J}.$$

**Densité de  $\mathcal{S}$ .** Montrons que  $\mathcal{S}$  est dense dans  $K$ . Soit  $\varepsilon \in \mathbb{R}_+^*$  et  $s \in K$ . Comme la suite  $(1/n)_n$  tend vers 0, il existe  $n \in \mathbb{N}_1$  tel que  $0 < 1/n < \varepsilon$ . Par (\*), il existe  $s' \in K_{1/n} \subset \mathcal{S}$  tel que  $s \in \mathbb{B}(s', 1/n)$ . Et donc on a :

$$\|s - s'\| < 1/n < \varepsilon$$

i.e.  $s' \in \mathbb{B}(s, \varepsilon) \cap \mathcal{S}$ . □

On utilisera dans le théorème suivant l'ensemble  $\mathcal{S} := \{s_k := \phi_{\mathcal{S}}(k)\}_{k \in J}$  construit dans le lemme et on pose :

$$U := \mathbb{R}^d - K.$$

### Théorème 2.3: Prolongement continue

Soit  $f \in \mathcal{C}(K, \mathbb{R})$ . Alors il existe une fonction  $g \in \mathcal{C}(\mathbb{R}^d, \mathbb{R})$  telle que :

$$g|_K = f.$$

*Démonstration.* Définissons trois fonctions avant de définir la fonction  $g$ .

(1) **Fonctions  $u_k$ .** Soit  $k \in J$ . La fonction  $u_k$  est définie par :

$$u_k : U \longrightarrow \mathbb{R} \\ x \longmapsto \max \left\{ 2 - \frac{\|x - s_k\|}{d(x, K)}, 0 \right\}$$

où pour tout  $x \in U$ , on a noté  $d(x, K)$  la **distance de  $x$  à  $K$**  définie par :

$$d(x, K) := \inf_{s \in K} \|x - s\| = \min_{s \in K} \|x - s\|$$

(la borne inférieure est un minimum car la fonction  $s \mapsto \|x - s\|$  est continue sur le compact  $K$  donc elle atteint ses bornes). On a deux propriétés importantes.

(a) **Montrons que  $u_k \in \mathcal{C}(U, [0, 1])$ .** La fonction  $u_k$  est le maximum de deux fonctions continues sur  $U$  donc elle est continue sur  $U$ . De plus pour tout  $x \in U$ , on a :

$$d(x, K) \leq d(x, s_k) = \|x - s_k\|$$

i.e. on a :

$$0 \leq 2 - \frac{\|x - s_k\|}{d(x, K)} \leq 1.$$

et donc  $u_k(x) \in [0, 1]$ .

(b) **Calcul de  $u_k^{-1}(\{0\})$ .** Pour tout  $x \in U$ , on a équivalence entre :

$$u_k(x) = 0 \iff 2 - \frac{\|x - s_k\|}{d(x, K)} \leq 0 \iff \|x - s_k\| \geq 2d(x, K).$$

(2) **Fonction  $\sigma$ .** La fonction  $\sigma$  est définie par :

$$\sigma : U \longrightarrow \mathbb{R} \\ x \longmapsto \sum_{j \in J} 2^{-j} u_j(x)$$

On a deux propriétés importantes.

(a) **Montrons que  $\sigma$  est définie et continue sur  $U$ .** Pour tout  $x \in U$ , on a :

$$(*) : 0 \leq \sigma(x) = \sum_{j \in J} 2^{-j} u_j(x) \leq \sum_{j \in J} 2^{-j} \leq \sum_{j=1}^{\infty} 2^{-j} = 1$$

donc  $\sigma$  est définie sur  $U$ . De plus comme pour tout  $j \in J$ , la fonction  $x \mapsto 2^{-j} u_j(x)$  est continue sur  $U$  et de norme infinie inférieure à  $2^{-j}$  sur  $U$ , on a par convergence normale que la fonction  $\sigma$  est continue sur  $U$  (car comme on a vu la série  $\sum_J 2^{-j}$  converge par  $(*)$ ).

(b) **Montrons que  $\sigma > 0$  sur  $U$ .** Raisonnons par l'absurde, supposons qu'il existe  $x \in U$  tel que  $\sigma(x) = 0$ . Comme  $K$  est compact, la distance entre  $x$  et  $K$  est atteinte *i.e.* il existe  $s \in K$  tel que :

$$\ell := d(x, K) = \|x - s\| > 0.$$

La somme définissant  $\sigma(x)$  est une somme de termes positifs, donc la condition  $\sigma(x) = 0$  est équivalente à pour tout  $j \in J$ ,  $u_j(x) = 0$  *i.e.*  $\|x - s_j\| \geq 2d(x, K) = 2\ell$ . Donc on a :

$$\inf_{j \in J} \|x - s_j\| \geq 2d(x, K) = 2\ell$$

Donc la boule ouverte  $\mathbb{B}(x, 2\ell)$  de centre  $x$  et de rayon  $2\ell$  ne rencontre pas  $\mathcal{S}$  *i.e.* :

$$\mathbb{B}(x, 2\ell) \cap \mathcal{S} = \emptyset$$

(en particulier  $s \notin \mathcal{S}$ ).

Ainsi la boule ouverte de centre  $s$  et de rayon  $\ell/2$  ne rencontre pas  $\mathcal{S}$  (mais contient l'élément  $s$  de  $K$ ) *i.e.*  $\mathbb{B}(s, \ell/2) \cap \mathcal{S} = \emptyset$  car on a par inégalité triangulaire renversée et pour tout  $j \in J$  :

$$\|s - s_j\| \geq \|s - x\| - \|x - s_j\| = \ell - \|x - s_j\| \geq \ell$$

*i.e.* on a  $s_j \notin \mathbb{B}(s, \ell/2)$ . Ce qui est contradictoire avec la densité de  $\mathcal{S}$  dans  $K$  car la boule ouverte  $\mathbb{B}(s, \ell/2)$  doit rencontrer  $\mathcal{S}$  par densité.

(3) **Fonctions  $v_k$ .** Soit  $k \in J$ . La fonction  $v_k$  est définie par :

$$\begin{aligned} v_k : U &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{1}{\sigma(x)} 2^{-k} u_k(x) \end{aligned}$$

Pour tout  $x \in U$ , on a :

$$\sum_{k \in J} v_k(x) = \frac{1}{\sigma(x)} \sum_{k \in J} 2^{-k} u_k(x) = \frac{\sigma(x)}{\sigma(x)} = 1.$$

**Fonction  $g$ .** On va définir la fonction  $g$  comme une série sur  $J$ . Commençons par expliquer la convergence. Pour tout  $k \in J$ , on a :

$$\forall x \in U, |v_k(x) f(s_k)| \leq v_k(x) \|f\|_{\infty, K}$$

et donc

$$(**) : \sum_{k \in J} v_k(x) \|f\|_{\infty, K} = \|f\|_{\infty, K} \sum_{k \in J} v_k(x) = \|f\|_{\infty, K}$$

La fonction  $g$  est donc définie par :

$$\begin{aligned} g : \mathbb{R}^d &\longrightarrow \mathbb{R} \\ x &\longmapsto \begin{cases} f(x) & \text{si } x \in K \\ \sum_{k \in J} v_k(x) f(s_k) & \text{si } x \in U. \end{cases} \end{aligned}$$

Ainsi la fonction  $g$  est un prolongement de  $f$  à  $\mathbb{R}^d$ . Montrons que  $g$  est continue sur  $\mathbb{R}^d$  *i.e.*  $g$  est continue sur  $U$  et sur  $K$ .

• **Continuité de  $g$  sur  $U$ .** Pour tout  $k \in J$ , la fonction  $x \mapsto v_k(x)f(s_k)$  est continue sur  $U$  et par (\*\*), on a la convergence normale sur  $U$  de la série définissant  $g$ . Donc la fonction  $g$  est continue sur  $U$ .

• **Continuité de  $g$  sur  $K$ .** Il ne reste plus qu'à montrer que  $g$  est continue sur  $K$  i.e. que pour tout  $a \in K \cap \bar{U}$  :

$$\lim_{\substack{x \rightarrow a \\ x \in U}} g(x) = f(a).$$

Soit  $\varepsilon \in \mathbb{R}_+^*$  et  $a \in K \cap \bar{U}$ . Par continuité de  $f$  en  $a$ , il existe  $\delta > 0$  tel que :

$$\forall k \in J, \|a - s_k\| < \delta \implies |f(a) - f(s_k)| < \varepsilon.$$

Soit  $x \in U$  tel que  $\|x - a\| < \delta/4$  (on peut trouver un tel  $x$  car  $a \in \bar{U}$  et donc toute boule ouverte de centre  $a$  rencontre  $U$ ). Remarquons d'abord que pour tout  $k \in J$  tel que  $\|a - s_k\| \geq \delta$ , on a par inégalité triangulaire renversée :

$$\|x - s_k\| \geq |\|x - a\| - \|a - s_k\|| \geq \frac{3}{4}\delta > 2\|x - a\| \geq 2d(x, K).$$

i.e.  $v_k(x) = 0$  (car  $u_k(x) = 0$  par (1.b)). On a donc par inégalité triangulaire :

$$\begin{aligned} |g(x) - f(a)| &= \left| \sum_{k \in J} v_k(x)f(s_k) - f(a) \sum_{k \in J} v_k(x) \right| \\ &\leq \sum_{k \in J} v_k(x) |f(s_k) - f(a)| \\ &= \sum_{\substack{k \in J \\ \|a - s_k\| < \delta}} v_k(x) |f(s_k) - f(a)| + \sum_{\substack{k \in J \\ \|a - s_k\| \geq \delta}} v_k(x) |f(s_k) - f(a)| \\ &= \sum_{\substack{k \in J \\ \|a - s_k\| < \delta}} v_k(x) |f(s_k) - f(a)| \\ &< \varepsilon \sum_{\substack{k \in J \\ \|a - s_k\| < \delta}} v_k(x) \\ &\leq \varepsilon \sum_{k \in J} v_k(x) \\ &= \varepsilon. \end{aligned}$$

□

## 2.2 Théorème

On en déduit alors le résultat suivant.

### Théorème 2.4: Stone-Weierstrass - cas d'un compact

Soit  $f \in \mathcal{C}(K, \mathbb{R})$ . Alors il existe une suite de fonctions polynômiales  $(p_n)_n$  qui converge uniformément vers  $f$  sur  $K$ .

*Démonstration.* Traitons deux cas.

(1) **Cas  $K := [-M, M]^d$  avec  $M > 0$  i.e.  $f \in \mathcal{C}([-M, M]^d, \mathbb{R})$ .** Posons (voir aussi la définition B.4) :

$$\begin{aligned} \mathfrak{h}(f) : [0, 1]^d &\longrightarrow \mathbb{R} \\ \mathbf{x} &\longmapsto f(M \cdot (2\mathbf{x} - \mathbf{1})) \end{aligned}$$

avec  $\mathbf{1} := (1, \dots, 1) \in \mathbb{R}^d$ . Alors on a  $\mathfrak{h}(f) \in \mathcal{C}([0, 1]^d, \mathbb{R})$ . Par le théorème 1.3, la suite  $(B_{n, \mathfrak{h}(f)}^{(d)})_n$  converge uniformément vers  $\mathfrak{h}(f)$  sur  $[0, 1]^d$ .

Posons pour tout  $\mathbf{x} \in [-M, M]^d$  (voir aussi la définition B.5) :

$$B_{n,f}^{(d)}(\mathbf{x}) := B_{n,\mathfrak{h}(f)}^{(d)}\left(\frac{1}{2M}(\mathbf{x} + M.\mathbf{1})\right).$$

Alors pour tout  $\mathbf{x} \in [-M, M]^d$  :

$$\left|B_{n,f}^{(d)}(\mathbf{x}) - f(\mathbf{x})\right| = \left|B_{n,\mathfrak{h}(f)}^{(d)}\left(\frac{1}{2M}(\mathbf{x} + M.\mathbf{1})\right) - \mathfrak{h}(f)\left(\frac{1}{2M}(\mathbf{x} + M.\mathbf{1})\right)\right| \leq \left\|B_{n,\mathfrak{h}(f)}^{(d)} - \mathfrak{h}(f)\right\|_{\infty,K}.$$

Donc on a :

$$\lim_{n \rightarrow \infty} \left\|B_{n,f}^{(d)} - f\right\| \leq \lim_{n \rightarrow \infty} \left\|B_{n,\mathfrak{h}(f)}^{(d)} - \mathfrak{h}(f)\right\|_{\infty,K} = 0.$$

Donc la suite de fonctions polynômiales  $\left(B_{n,f}^{(d)}\right)_n$  converge uniformément vers  $f$  sur  $[-M, M]^d$ .

(2) **Cas général.** Comme  $K$  est compact *i.e.* fermé et borné, il existe  $M > 0$  tel que :

$$K \subset [-M, M]^d.$$

Par le théorème 2.3, il existe un prolongement continue  $g$  de  $f$  à  $\mathbb{R}^d$ . Posons :

$$\tilde{f} := g|_{[-M, M]^d}.$$

On a  $\tilde{f} \in \mathcal{C}([-M, M]^d, \mathbb{R})$ . Par le point (1), la suite de fonctions polynômiales  $(q_n)_n := \left(B_{n,\tilde{f}}^{(d)}\right)_n$  converge uniformément vers  $\tilde{f}$  sur  $[-M, M]^d$ .

Donc par restriction à  $K$ , la suite de fonctions polynômiales  $(p_n)_n := (q_n|_K)_n$  converge uniformément vers  $f$  sur  $K$ .

□

Pour  $f \in \mathcal{C}(K, \mathbb{R})$ , notons  $\text{pr}(f, \phi_{\mathcal{S}})$  le prolongement continue de  $f$  à  $\mathbb{R}^d$  construit dans la preuve du théorème 2.3 (ce prolongement dépend de  $\mathcal{S}$  et précisément de la bijection  $\phi_{\mathcal{S}}$  de  $J$  vers  $\mathcal{S}$  définie dans le lemme 2.2). On peut décrire explicitement la suite de fonctions polynômiales  $(p_n)_n$  qui converge uniformément vers  $f$  sur  $K$ , elle est donnée pour tout  $n \in \mathbb{N}_1$  par :

$$p_n := \left(B_{n,\text{pr}(f,\phi_{\mathcal{S}})}^{(d)}\right)|_K.$$

## A Artifaillle en probabilités

### A.1 Existence de variables aléatoires binomiales indépendantes

Dans cette sous-section, on justifie l'existence d'une suite finie de variables aléatoires indépendantes suivant des lois binomiales.

#### Définition A.1

Soit  $n \in \mathbb{N}$  et  $p \in [0, 1]$ . La loi binomiale de paramètres  $n$  et  $p$  est l'application  $\mathcal{B}_{n,p}$  définie par :

$$\begin{aligned} \mathcal{B}_{n,p}: \quad [0, n] &\longrightarrow [0, 1] \\ k &\longmapsto \binom{n}{k} p^k (1-p)^{n-k}. \end{aligned}$$

Pour  $n \in \mathbb{N}$  et  $p \in [0, 1]$ , on a par le binôme de Newton :

$$\sum_{k=0}^n \mathcal{B}_{n,p}(k) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n = 1.$$

#### Théorème A.2

Soit  $d, n \in \mathbb{N}_1$  et  $\mathbf{x} := (x_1, \dots, x_d) \in [0, 1]^d$ . Alors il existe un espace probabilisé  $(\Omega, \mathbb{P})$  et une suite de  $d$  variables aléatoires indépendantes  $\mathbf{X} := (X_1, \dots, X_d)$  sur  $(\Omega, \mathbb{P})$  de lois binomiales  $\mathcal{B}_{n,x_1}, \dots, \mathcal{B}_{n,x_d}$  respectivement.

*Démonstration.* Posons pour tout  $k \in [1, d]$  :

$$\mathcal{B}_k := \mathcal{B}_{n,x_k}.$$

On définit les objets suivants.

- **(Univers)** Posons  $\Omega := [0, n]^d$ .
- **(Probabilités)** Comme :

$$\sum_{\omega \in \Omega} \left( \prod_{k=1}^d \mathcal{B}_k(\omega_k) \right) = \prod_{k=1}^d \left( \sum_{\omega \in \Omega} \mathcal{B}_k(\omega_k) \right) = 1.$$

il existe un unique probabilité  $\mathbb{P}$  sur  $\Omega$  telle que :

$$\forall \omega \in \Omega, \quad \mathbb{P}(\{\omega\}) := \prod_{k=1}^d \mathcal{B}_k(\omega_k).$$

- **(Variables aléatoires)** Pour tout  $k \in [1, d]$ , on définit la variable aléatoire  $X_k$  par :

$$\begin{aligned} X_k: \quad \Omega &\longrightarrow [0, n] \\ \omega &\longmapsto \omega_k \end{aligned}$$

Soit  $l \in [1, d]$ . Montrons que  $X_l \sim \mathcal{L}_l$ . On a pour tout  $\omega \in [0, n]^{d-1}$  et tout  $x \in [0, n]$  :

$$X_l((\omega_1, \dots, \omega_{l-1}, x, \omega_l, \dots, \omega_d)) = x.$$

Ainsi on a  $X_l(\Omega) = [0, n]$ .

Pour tout  $x \in [0, n]$ , on a :

$$\mathbb{P}(X_l = x) = \mathbb{P}(\{\omega \in \Omega, \omega_l = x\})$$

$$\begin{aligned}
&= \mathbb{P} \left( \bigsqcup_{\omega \in \llbracket 0, n \rrbracket^{d-1}} \{(\omega_1, \dots, \omega_{l-1}, x, \omega_l, \dots, \omega_d)\} \right) \\
&= \sum_{\omega \in \llbracket 0, n \rrbracket^{d-1}} \mathbb{P}(\{(\omega_1, \dots, \omega_{l-1}, x, \omega_l, \dots, \omega_d)\}) \\
&= \sum_{\omega \in \llbracket 0, n \rrbracket^{d-1}} \mathcal{B}_l(x) \prod_{\substack{k=1 \\ k \neq l}}^d \mathcal{B}_k(\omega_k) \\
&= \mathcal{B}_l(x) \sum_{\omega \in \llbracket 0, n \rrbracket^{d-1}} \prod_{\substack{k=1 \\ k \neq l}}^d \mathcal{B}_k(\omega_k) \\
&= \mathcal{B}_l(x) \prod_{\substack{k=1 \\ k \neq l}}^d \left( \sum_{\omega_k \in \llbracket 0, n \rrbracket} \mathcal{B}_k(\omega_k) \right) \\
&= \mathcal{B}_l(x).
\end{aligned}$$

Ainsi,  $X_l \sim \mathcal{L}_l$ .

Montrons que les variables aléatoires  $X_1, \dots, X_d$  sont mutuellement indépendantes. Soit  $I \subset \llbracket 1, d \rrbracket$ . Posons  $I^c := \llbracket 1, d \rrbracket - I$ . Pour tout  $x \in \llbracket 0, n \rrbracket^I$ , on a :

$$\begin{aligned}
\mathbb{P} \left( \bigcap_{i \in I} (X_i = x(i)) \right) &= \mathbb{P}(\{\omega \in \Omega, \forall i \in I, \omega_i = x(i)\}) \\
&= \mathbb{P} \left( \bigsqcup_{y \in \llbracket 0, n \rrbracket^{I^c}} \{\omega \in \Omega, \forall i \in I, \omega_i = x(i) \wedge \forall k \in I^c, \omega_k = y(k)\} \right) \\
&= \sum_{y \in \llbracket 0, n \rrbracket^{I^c}} \mathbb{P}(\{\omega \in \Omega, \forall i \in I, \omega_i = x(i) \wedge \forall k \in I^c, \omega_k = y(k)\}) \\
&= \sum_{y \in \llbracket 0, n \rrbracket^{I^c}} \mathcal{B}_i(x(i)) \cdot \prod_{k \in I^c} \mathcal{B}_k(y(k)) \\
&= \prod_{i \in I} \mathcal{B}_i(x(i)) \left( \sum_{y \in \llbracket 0, n \rrbracket^{I^c}} \prod_{k \in I^c} \mathcal{B}_k(y(k)) \right) \\
&= \prod_{i \in I} \mathcal{B}_i(x(i)) \cdot \prod_{k \in I^c} \left( \sum_{y_k=0}^n \mathcal{B}_k(y(k)) \right) \\
&= \prod_{i \in I} \mathcal{B}_i(x(i)) \\
&= \prod_{i \in I} \mathbb{P}(X_i = x(i))
\end{aligned}$$

D'où le résultat. □

## A.2 Deux inégalités

Dans cette deuxième sous-section, on démontre deux inégalités connues en probabilités (voir les notations A.4 ci-après pour les définitions de l'espérance et de la variance d'une variable aléatoire).

### Théorème A.3

Soit  $X$  une variable aléatoire sur un espace probabilité fini  $(\Omega, \mathbb{P})$  d'espérance  $m$  et de variance  $\sigma$ .

(i) **(Inégalité de Markov)** Pour tout  $a > 0$ , on a :

$$\mathbb{P}(|X| \geq a) \leq \frac{m}{a}.$$

(ii) **(Inégalité de Bienaymé-Tchebychev)** Pour tout  $a > 0$ , on a :

$$\mathbb{P}(|X - m| \geq a) \leq \frac{\sigma}{a^2}.$$

*Démonstration.* (i) Soit  $a > 0$ . Posons

$$A := \{|X| \geq a\} := \{\omega \in \Omega, |X(\omega)| \geq a\}.$$

Soit  $\omega \in \Omega$ . On a deux cas.

(1) **Cas**  $\omega \in A$ . On a :

$$\begin{aligned} a1_A(\omega) &= a \cdot 1 = a \\ |X(\omega)|1_A(\omega) &= |X(\omega)| \cdot 1 = |X(\omega)| \geq a. \end{aligned}$$

Donc on a :

$$a1_A(\omega) \leq |X(\omega)|1_A(\omega) = |X(\omega)|$$

(2) **Cas**  $\omega \in A^c$ . On a :

$$\begin{aligned} a1_A(\omega) &= 0 \\ |X(\omega)|1_A(\omega) &= 0 \\ |X(\omega)| &\geq 0. \end{aligned}$$

Donc on a :

$$a1_A(\omega) = |X(\omega)|1_A(\omega) \leq |X(\omega)|.$$

Ainsi on a dans les deux cas :

$$a1_A(\omega) \leq |X(\omega)|1_A(\omega) \leq |X(\omega)|.$$

Donc on a :

$$a1_A \leq |X|1_A \leq |X|.$$

Les trois fonctions  $a1_A$ ,  $|X|1_A$ , et  $|X|$  sont des variables aléatoires sur  $\Omega$ . Donc par croissance de l'espérance, on a :

$$\mathbb{E}(a1_A) \leq \mathbb{E}(|X|1_A) \leq \mathbb{E}(|X|)$$

Or on a par linéarité de l'espérance :

$$\mathbb{E}(a1_A) = a\mathbb{E}(1_A) = a\mathbb{P}(A).$$

Et donc on a :

$$a\mathbb{P}(A) \leq \mathbb{E}(|X|)$$

*i.e.*

$$\mathbb{P}(A) \leq \frac{\mathbb{E}(|X|)}{a}.$$

(ii) Soit  $a > 0$ . Posons :

$$Y := (X - \mathbb{E}(X))^2.$$

On a  $Y \geq 0$  *i.e.*  $|Y| = Y$ . Et ainsi par (i), on a :

$$\mathbb{P}(Y \geq a^2) \leq \frac{\mathbb{E}(Y)}{a^2}.$$

*i.e.*

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{\mathbb{V}(X)}{a^2}.$$

□

On en déduit alors une inégalité générale en dimension  $d$  (voir [7] pour d'autres types d'inégalités). Commençons par quelques notations sur les suites de variables aléatoires.

### Notation A.4

Soit  $\mathbf{X} := (X_1, \dots, X_d)$  une suite de  $d$  variables aléatoires sur un espace probabilisé fini  $(\Omega, \mathbb{P})$ .

(1) L'espérance de  $\mathbf{X}$  est définie par :

$$\mathbb{E}(\mathbf{X}) := (\mathbb{E}(X_1), \dots, \mathbb{E}(X_d)).$$

avec pour tout  $i \in \llbracket 1, d \rrbracket$  :

$$\mathbb{E}(X_i) := \sum_{x \in X_i(\Omega)} x \mathbb{P}(X_i = x) = \sum_{\omega \in \Omega} X_i(\omega) \mathbb{P}(\{\omega\}).$$

(2) La matrice de covariance de  $\mathbf{X}$  est définie par la matrice symétrique :

$$\mathbb{V}(\mathbf{X}) := (\text{Cov}(X_i, X_j))_{1 \leq i, j \leq d} = \begin{pmatrix} \mathbb{V}(X_1) & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_d) \\ \text{Cov}(X_1, X_2) & \mathbb{V}(X_2) & \cdots & \text{Cov}(X_2, X_d) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_1, X_d) & \text{Cov}(X_2, X_d) & \cdots & \mathbb{V}(X_d) \end{pmatrix}$$

avec pour tous  $i, j \in \llbracket 1, d \rrbracket$  :

$$\begin{aligned} \text{Cov}(X_i, X_j) &:= \mathbb{E}((X_i - \mathbb{E}(X_i))(X_j - \mathbb{E}(X_j))) \\ \mathbb{V}(X_i) &:= \text{Cov}(X_i, X_i) = \mathbb{E}((X_i - \mathbb{E}(X_i))^2) = \mathbb{E}(X_i^2) - \mathbb{E}(X_i)^2 \end{aligned}$$

la dernière égalité s'appelle la **formule de König-Huygens**.

On a alors une généralisation possible de l'inégalité de Bienaymé-Tchebychev en dimension  $d$ .

### Corollary A.5: Une inégalité de Bienaymé-Tchebychev en dimension $d$

Soit  $\delta > 0$  et  $\mathbf{X} := (X_1, \dots, X_d)$  une suite de  $d$  variables aléatoires sur un espace probabilisé fini  $(\Omega, \mathbb{P})$ . On a :

$$\mathbb{P}(\|\mathbf{X} - \mathbb{E}(\mathbf{X})\|_\infty > \delta) \leq \frac{\text{Tr}(\mathbb{V}(\mathbf{X}))}{\delta^2}.$$

*Démonstration.* On reprend les raisonnements des points (7) à (11) de la preuve du théorème 1.3. On a par la point (ii) du théorème A.3 :

$$\begin{aligned} \mathbb{P}(\|\mathbf{X} - \mathbb{E}(\mathbf{X})\|_\infty > \delta) &= \mathbb{P}\left(\max_{i \in \llbracket 1, d \rrbracket} |X_i - \mathbb{E}(X_i)| > \delta\right) \\ &= \mathbb{P}\left(\bigcup_{i=1}^d \{|X_i - \mathbb{E}(X_i)| > \delta\}\right) \\ &\leq \sum_{i=1}^d \mathbb{P}(|X_i - \mathbb{E}(X_i)| > \delta) \\ &\leq \sum_{i=1}^d \frac{\mathbb{V}(X_i)}{\delta^2} \\ &= \frac{\text{Tr}(\mathbb{V}(\mathbf{X}))}{\delta^2}. \end{aligned}$$

□

On aurait pu prendre une autre norme  $\|\cdot\|$  sur  $\mathbb{R}^d$ . Comme toutes les normes sont équivalentes, il existe  $\alpha, \beta > 0$  tels que :

$$\forall \mathbf{x} \in \mathbb{R}^d, \alpha \|\mathbf{x}\|_\infty \leq \|\mathbf{x}\| \leq \beta \|\mathbf{x}\|_\infty.$$

**Corollary A.6: Une inégalité de Bienaymé-Tchebychev en dimension  $d$** 

Soit  $\delta > 0$  et  $\mathbf{X} := (X_1, \dots, X_d)$  une suite de  $d$  variables aléatoires sur un espace probabilisé fini  $(\Omega, \mathbb{P})$ . On a :

$$\mathbb{P}(\|\mathbf{X} - \mathbb{E}(\mathbf{X})\| > \delta) \leq \frac{\beta^2 \text{Tr}(\mathbb{V}(\mathbf{X}))}{\delta^2}.$$

*Démonstration.* Comme :

$$\{\|\mathbf{X} - \mathbb{E}(\mathbf{X})\| > \delta\} \subset \{\beta \|\mathbf{X} - \mathbb{E}(\mathbf{X})\|_\infty > \delta\}$$

on a par croissance de  $\mathbb{P}$  :

$$\begin{aligned} \mathbb{P}(\|\mathbf{X} - \mathbb{E}(\mathbf{X})\| > \delta) &\leq \mathbb{P}(\beta \|\mathbf{X} - \mathbb{E}(\mathbf{X})\|_\infty > \delta) \\ &= \mathbb{P}\left(\|\mathbf{X} - \mathbb{E}(\mathbf{X})\|_\infty > \frac{\delta}{\beta}\right) \\ &\leq \frac{\beta^2 \text{Tr}(\mathbb{V}(\mathbf{X}))}{\delta^2} \end{aligned}$$

□

## B Polynômes de Bernstein généralisés

Dans cette sous-section, on généralise la définition 1.1 des fonctions polynômiales simplifiées de Bernstein (voir aussi [6]).

### B.1 Cas sur l'hypercube $[0, 1]^d$

On commence par le cas de l'hypercube  $[0, 1]^d$ .

#### Définition B.1

Soit  $f \in \mathcal{C}([0, 1]^d, \mathbb{R})$  et  $\mathbf{n} := (n_1, \dots, n_d) \in (\mathbb{N}_1)^d$ . La **fonction polynômiale de Bernstein de paramètres  $d$  et  $\mathbf{n}$  associée à  $f$**  est définie par :

$$B_{\mathbf{n},f}^{(d)} : [0, 1]^d \longrightarrow \mathbb{R}$$

$$(x_1, \dots, x_d) \longmapsto \sum_{(k_1, \dots, k_d) \in \llbracket 0, n \rrbracket^d} f\left(\frac{k_1}{n_1}, \dots, \frac{k_d}{n_d}\right) \prod_{i=1}^d \binom{n_i}{k_i} x_i^{k_i} (1 - x_i)^{n - k_i}$$

On retrouve la définition 1.1 en prenant  $\mathbf{n} := (n, \dots, n)$  avec  $n \in \mathbb{N}_1$  :

$$B_{(n, \dots, n), f}^{(d)} = B_{n, f}^{(d)}.$$

On reprend les exemples 1.2.

#### Exemple B.2

Soit  $\mathbf{n} := (n_1, \dots, n_d) \in (\mathbb{N}_1)^d$ . Posons :

$$\Gamma := \prod_{i=1}^d \llbracket 0, n_i \rrbracket$$

(1) **Cas  $f := 1$  sur  $[0, 1]^d$ .** On a pour tout  $(x_1, \dots, x_d) \in [0, 1]^d$  :

$$\begin{aligned} B_{\mathbf{n},1}^{(d)}(x_1, \dots, x_d) &= \sum_{(k_1, \dots, k_d) \in \Gamma} \prod_{i=1}^d \binom{n_i}{k_i} x_i^{k_i} (1 - x_i)^{n_i - k_i} \\ &= \prod_{i=1}^d \left( \sum_{k_i \in \llbracket 0, n_i \rrbracket} \binom{n_i}{k_i} x_i^{k_i} (1 - x_i)^{n_i - k_i} \right) \\ &= \prod_{i=1}^d 1 \\ &= 1 \end{aligned}$$

(2) **Cas  $f$  à variables séparées sur  $[0, 1]^d$ .** Supposons qu'il existe des fonctions continues  $f_1, \dots, f_d \in \mathcal{C}([0, 1], \mathbb{R})$  telles que pour tout  $(x_1, \dots, x_d) \in [0, 1]^d$  :

$$f(x_1, \dots, x_d) = \prod_{i=1}^d f_i(x_i).$$

On a pour tout  $(x_1, \dots, x_d) \in [0, 1]^d$  :

$$\begin{aligned} B_{\mathbf{n},f}^{(d)}(x_1, \dots, x_d) &= \sum_{(k_1, \dots, k_d) \in \Gamma} f\left(\frac{k_1}{n_1}, \dots, \frac{k_d}{n_d}\right) \prod_{i=1}^d \binom{n_i}{k_i} x_i^{k_i} (1 - x_i)^{n_i - k_i} \\ &= \sum_{(k_1, \dots, k_d) \in \Gamma} \prod_{i=1}^d f_i\left(\frac{k_i}{n_i}\right) \binom{n_i}{k_i} x_i^{k_i} (1 - x_i)^{n_i - k_i} \end{aligned}$$

$$\begin{aligned}
 &= \prod_{i=1}^d \left( \sum_{k_i=0}^n f_i \left( \frac{k_i}{n_i} \right) \binom{n_i}{k_i} x_i^{k_i} (1-x_i)^{n_i-k_i} \right) \\
 &= \prod_{i=1}^d B_{n_i, f_i}(x_i)
 \end{aligned}$$

## B.2 Cas sur un pavé

Commençons par la définition d'un segment et d'un pavé en dimension  $n$ .

### Définition B.3

- (i) Soit  $(a, b) \in \mathbb{R}^2$  tels que  $a \leq b$ . Le **segment**  $[a, b]$  est l'ensemble des  $x \in \mathbb{R}$  tel que  $a \leq x \leq b$ .
- (ii) Soit une partie  $I$  de  $\mathbb{R}$ . On dit que  $I$  est un **segment de**  $\mathbb{R}$  s'il existe un couple  $(a, b) \in \mathbb{R}^2$  tels que  $a \leq b$  et  $I := [a, b]$ .
- (iii) Soit une partie  $K$  de  $\mathbb{R}^d$ . On dit que  $I$  est un **pavé de**  $\mathbb{R}^d$  s'il existe une suite de couple  $(a_i, b_i)_{i \in \llbracket 1, d \rrbracket} \in (\mathbb{R}^2)^d$  telle que pour tout  $i \in \llbracket 1, d \rrbracket$ ,  $a_i \leq b_i$  et :

$$K = \prod_{i=1}^d [a_i, b_i].$$

On dit que le pavé est **non trivial** si pour tout  $i \in \llbracket 1, d \rrbracket$ ,  $a_i < b_i$ .

Tous les segments et pavés considérés dans la suite seront non triviaux.

On a pour tout  $a \in \mathbb{R}$  :

$$[a, a] = \{a\}.$$

Comme tout segment  $[a, b]$  (non trivial) peut s'écrire comme l'ensemble des barycentres positifs de  $a$  et  $b$  :

$$[a, b] = \{xb + (1-x)a, 0 \leq x \leq 1\} = \{xb + (1-x)a, x \in [0, 1]\}$$

on a une bijection naturelle :

$$\begin{aligned}
 [0, 1] &\longrightarrow [a, b] \\
 x &\longmapsto xb + (1-x)a
 \end{aligned}$$

de réciproque :

$$\begin{aligned}
 [a, b] &\longrightarrow [0, 1] \\
 x &\longmapsto \frac{x-a}{b-a}
 \end{aligned}$$

et pour tout pavé  $\prod_{i=1}^d [a_i, b_i]$  (non trivial) de  $\mathbb{R}^d$ , on a une bijection :

$$\begin{aligned}
 [0, 1]^d &\longrightarrow \prod_{i=1}^d [a_i, b_i] \\
 (x_1, \dots, x_d) &\longmapsto (x_i b_i + (1-x_i) a_i)_{i \in \llbracket 1, d \rrbracket}
 \end{aligned}$$

de réciproque :

$$\begin{aligned}
 \prod_{i=1}^d [a_i, b_i] &\longrightarrow [0, 1]^d \\
 (x_1, \dots, x_d) &\longmapsto \left( \frac{x_i - a_i}{b_i - a_i} \right)_{i \in \llbracket 1, d \rrbracket}
 \end{aligned}$$

On peut alors définir les polynômes de Bernstein sur des pavés quelconques de  $\mathbb{R}^d$ . D'où la définition suivante.

#### Définition B.4

Soit  $f \in \mathcal{C}(\prod_{i=1}^d [a_i, b_i], \mathbb{R})$ . La **fonction continue sur  $[0, 1]^d$  associée à  $f$**  est définie par :

$$\begin{aligned} \mathfrak{h}(f) : [0, 1]^d &\longrightarrow \mathbb{R} \\ \mathbf{x} &\longmapsto f\left(\left(x_i b_i + (1 - x_i) a_i\right)_{i \in \llbracket 1, d \rrbracket}\right) \end{aligned}$$

On a alors la définition générale suivante des fonctions polynômiales de Bernstein associées à une fonction définie sur un pavé de  $\mathbb{R}^d$ .

#### Définition B.5

Soit  $f \in \mathcal{C}(\prod_{i=1}^d [a_i, b_i], \mathbb{R})$  et  $\mathbf{n} \in (\mathbb{N}_1)^d$ . La **fonction polynômiale de Bernstein de paramètres  $d$  et  $\mathbf{n}$  associée à  $f$**  est définie pour tout  $\mathbf{x} \in \prod_{i=1}^d [a_i, b_i]$  :

$$B_{\mathbf{n}, f}^{(d)}(\mathbf{x}) := B_{\mathbf{n}, \mathfrak{h}(f)}^{(d)}\left(\left(\frac{x_i - a_i}{b_i - a_i}\right)_{i \in \llbracket 1, d \rrbracket}\right).$$

## Références

- [1] N. Bourbaki, *Éléments de mathématique, livre III : Topologie générale*, Chapitre 9, Springer, (2007).
- [2] Evans L. C., Gariépy R. F., *Measure Theory and Fine Properties of Functions*, CRC Press, (1991).
- [3] Hildebrandt T. H., Schoenberg I. J., *On linear functional operations and the moment problem for a finite interval in one or several dimensions*, Annals of Mathematics, 34 (2) : 327, (1933).
- [4] Stone M. H., *The Generalized Weierstrass Approximation Theorem*, Mathematics Magazine, 21 (4) : 167–184, (1948).
- [5] Weierstrass K., *Über die analytische Darstellbarkeit sogenannter willkürlicher Functionen einer reellen Veränderlichen*, Sitzber. K. Preuss. Akad. Wiss. Berlin, (1885).
- [6] Wikipédia, *Bernstein polynômial # Generalizations to higher dimension*.
- [7] Wikipédia, *Chebyshev's inequality # Extensions*.