Bayesian estimation using interval analysis

Luc Jaulin,
Laboratoire E^3I^2
(Extraction et Exploitation de l'Information en Environnements Incertains)
ENSIETA, 2 rue François Verny
29806 Brest Cédex 09
E-mail : luc.jaulin@ensieta.fr

Séminaire pour le groupe de travail

Calcul ensembliste
Jeudi 23 septembre 2004.

1 Motivations

Model : $\phi(\mathbf{p}, t) = p_1 e^{-p_2 t}$.

Parameters : p_1, p_2.

Sampling times : $t_1, t_2, \ldots, t_m$

Data bars : $[y_1^-, y_1^+], [y_2^-, y_2^+], \ldots, [y_m^-, y_m^+]$

Feasible set :

$$\begin{aligned} \mathbb{S} &= \left\{\mathbf{p} \in \mathbb{R}^2, \forall i \in \{1, \ldots, m\}, \phi(\mathbf{p}, t_i) \in [y_i^-, y_i^+]\right\}. \\ &= \boldsymbol{\phi}^{-1}([\mathbf{y}]). \end{aligned}$$

SetDemo (Guillaume Baffet)

2 Problem

Consider a function $f(\mathbf{p})$ positive for all $\mathbf{p} \in \mathbb{R}^n$, such as $\int_{\mathbb{R}^n} f(\mathbf{p}) d\mathbf{p}$ is finite and a real number $\alpha \in [0, 1]$. Characterize the set $\mathbb{S}_\alpha$ defined by

$$\begin{array}{ll} \text{(i)} & \mathbb{S}_\alpha = f^{-1}([s_\alpha, +\infty[), \\ \text{(ii)} & \dfrac{\int_{\mathbb{S}_\alpha} f(\mathbf{p}) d\mathbf{p}}{\int_{\mathbb{R}^n} f(\mathbf{p}) d\mathbf{p}} = \alpha. \end{array}$$

The set $\mathbb{S}_\alpha$ is the confidence region associated with the unnormalized pdf f.

It corresponds to the smallest set which contains $\mathbf{p}$ with a probability equal to α.

Example : Consider a random variable p, described by the unnormalized pdf:

$$f(p) = \exp\left(-\frac{p^2}{2}\right).$$

Let us compute its confidence region $\mathbb{S}_{0.95}$. Since,

$$\int_{-\infty}^{\infty} \exp\left(-\frac{p^2}{2}\right) dp = \sqrt{2\pi},$$

we should solve

$$\begin{array}{ll} \text{(i)} & \mathbb{S}_{0.95} = f^{-1}([s_\alpha, +\infty[), \\ \text{(ii)} & \frac{1}{\sqrt{2\pi}} \int_{\mathbb{S}_\alpha} f(\mathbf{p}) d\mathbf{p} = 0.95. \end{array}$$

For this example, $\mathbb{S}_{0.95} = [-b, b]$. Thus

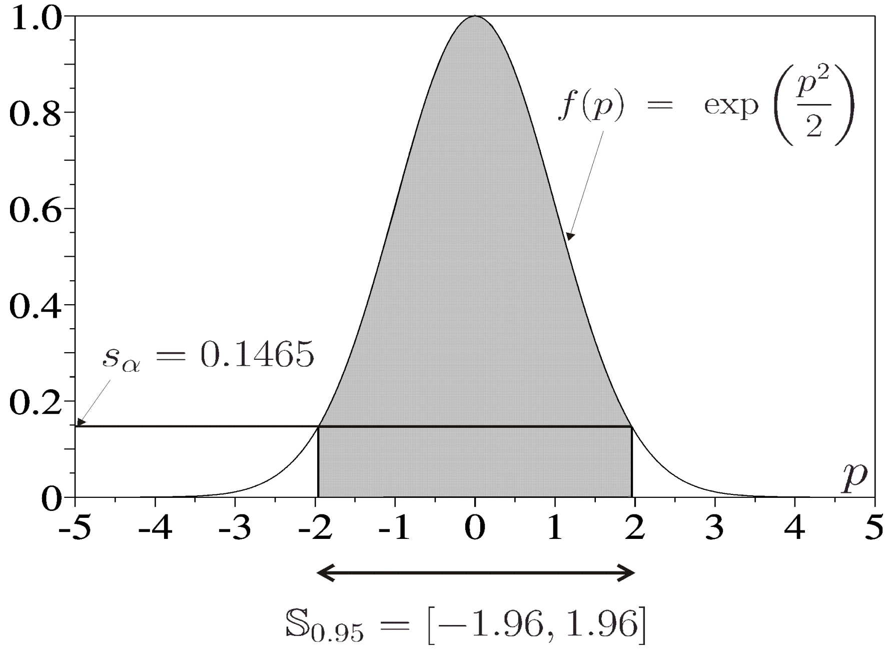
$$\begin{array}{ll} \text{(i)} & [-b, b] = f^{-1}([s_\alpha, +\infty[), \\ \text{(ii)} & \frac{1}{\sqrt{2\pi}} \int_{-b}^{b} \exp\left(-\frac{p^2}{2}\right) dp = 0.95. \end{array}$$

After integration

$$\begin{array}{ll} \text{(i)} & [-b,b] = f^{-1}([s_\alpha, +\infty[), \\ \text{(ii)} & \quad \mathrm{erf}\left(\tfrac{1}{2}b\sqrt{2}\right) = 0.95. \end{array}$$

We get $b = 1.96$. Finally,

$$\begin{aligned} \mathbb{S}_{0.95} &= [-1.96, 1.96]. \\ s_\alpha &= f(b) = \exp\left(-\frac{1.96^2}{2}\right) = 0.1465. \end{aligned}$$



1.0
0.8
0.6
0.4
0.2
0
$f(p) = \exp\left(\frac{p^2}{2}\right)$
$s_\alpha = 0.1465$
p
-5 -4 -3 -2 -1 0 1 2 3 4 5
$\mathbb{S}_{0.95} = [-1.96, 1.96]$

3 Lattices and intervals

A *lattice* $(\mathcal{E}, \leq)$ is a partially ordered set, closed under least upper and greatest lower bounds.

The least upper bound (*join*) of x and y is written $x \vee y$.

The greatest lower bound (*meet*) is written $x \wedge y$.

A lattice $\mathcal{E}$ is *complete* if for all subsets $\mathcal{A}$ of $\mathcal{E}$, $\vee\mathcal{A}$ and $\wedge\mathcal{A}$ belong to $\mathcal{E}$.

Example 1 : The set $\mathbb{R}$ is not a complete lattice whereas $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ is.

Example 2 : The set $\mathbb{R}^n$ is a lattice with respect to the partial order relation given by

$$\mathbf{x} \leq \mathbf{y} \Leftrightarrow \forall i \in \{\mathbf{1}, \ldots, n\}, x_i \leq y_i.$$

We have

$$\begin{aligned} \mathbf{x} \wedge \mathbf{y} &= (\min(x_1, y_1), \ldots, \min(x_n, y_n)) \text{ and} \\ \mathbf{x} \vee \mathbf{y} &= (\max(x_1, y_1), \ldots, \max(x_n, y_n)). \end{aligned}$$

An *interval* $[x]$ of a complete lattice $\mathcal{E}$ is a subset of $\mathcal{E}$ which satisfies

$$[x] = \{x \in \mathcal{E} \mid \wedge [x] \leq x \leq \vee [x]\}\,.$$

Both $\emptyset$ and $\mathcal{E}$ are intervals of $\mathcal{E}$.

The sets $[0,1]_{\bar{\mathbb{R}}}$ and $[0,\infty]_{\bar{\mathbb{R}}}$ are intervals of $\bar{\mathbb{R}}$.

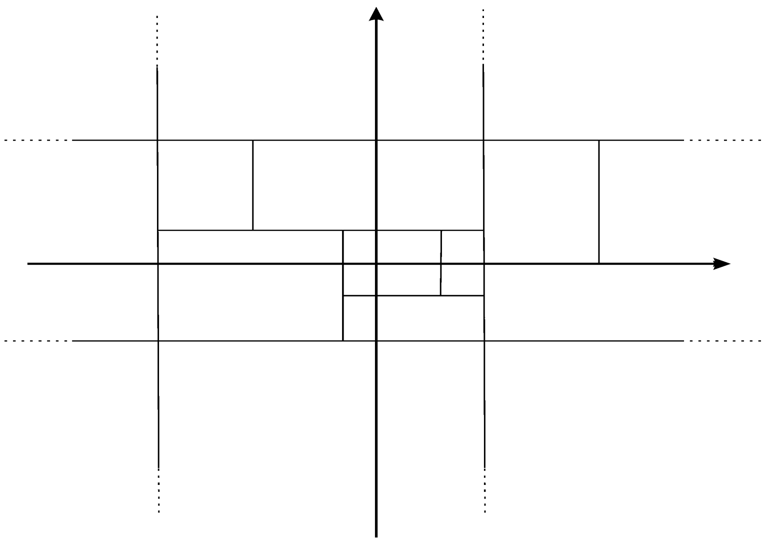
The set $\{2,3,4,5\} = [2,5]_{\bar{\mathbb{N}}}$ is an interval of $\bar{\mathbb{N}}$.

The set $\{4,6,8,10\} = [4,10]_{2\bar{\mathbb{N}}}$ is an interval of $2\bar{\mathbb{N}}$.

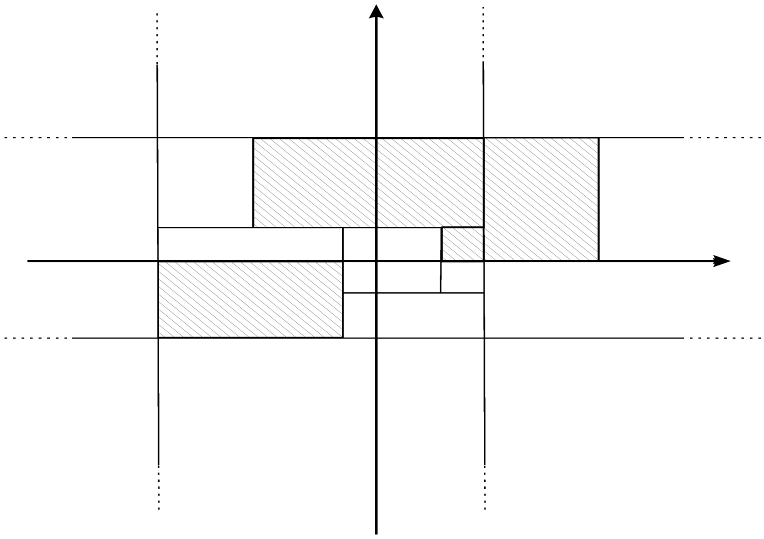
The set $[1,2] \times [3,4]) = [(1,3),(2,4)]_{\bar{\mathbb{R}}^2}$ is an interval of $\bar{\mathbb{R}}^2$.

4 Interval subpavings

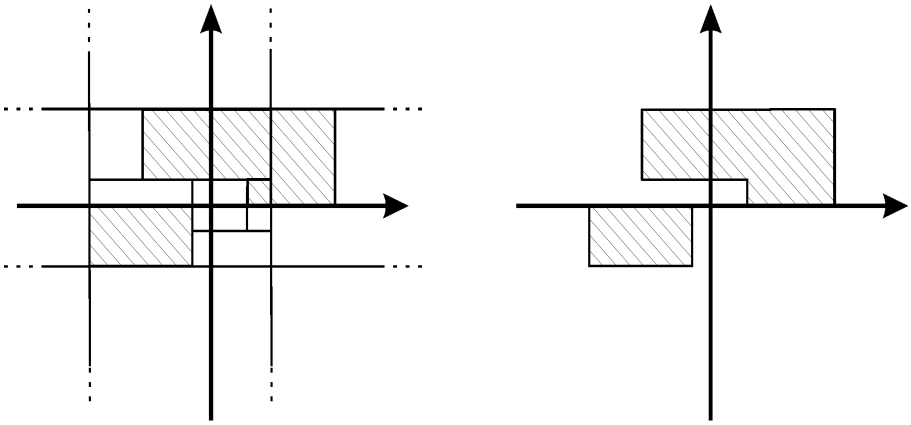
A *paving* $\mathcal{Q}$ of $\mathbb{R}^n$ is a set of nonoverlapping boxes covering $\mathbb{R}^n$.



A *subpaving* of $\mathcal{Q}$ is a subset of $\mathcal{Q}$.



The *support* $\{\mathcal{K}\} \subset \mathbb{R}^n$ of a subpaving $\mathcal{K}$ is the union of all boxes of $\mathcal{K}$.



If $\mathcal{P}(\mathcal{Q})$ denotes the set of all subpavings of $\mathcal{Q}$ then $(\mathcal{P}(\mathcal{Q}), \subset)$ is a complete lattice.

- The least upper bound (*join*) is the union:

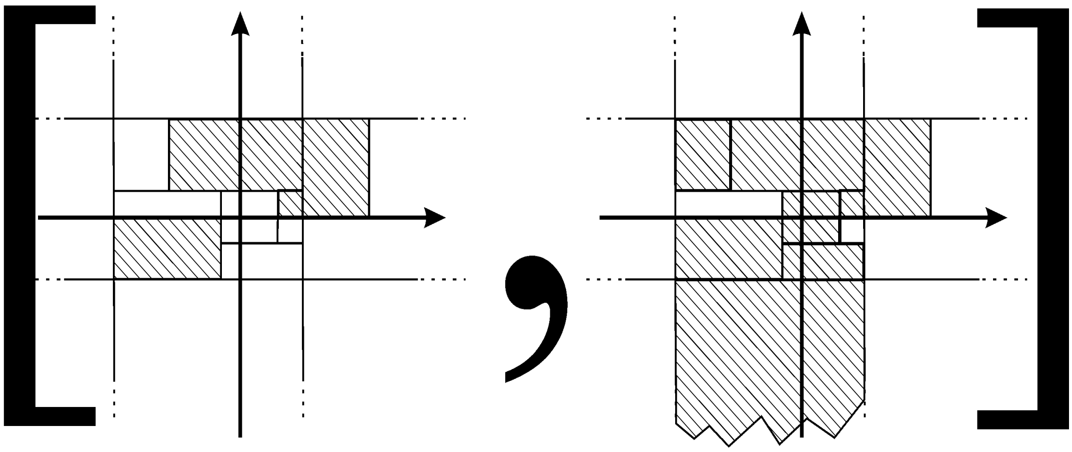
$$\mathcal{K}_1 \vee \mathcal{K}_2 = \mathcal{K}_1 \cup \mathcal{K}_2.$$

- The greatest lower bound (*meet*) is the intersection

$$\mathcal{K}_1 \wedge \mathcal{K}_2 = \mathcal{K}_1 \cap \mathcal{K}_2.$$

As a consequence intervals of $(\mathcal{P}(\mathcal{Q}), \subset)$ can be defined.

An interval subpaving $[\mathcal{K}^-, \mathcal{K}^+]$ of $\mathcal{Q}$ can be represented by pair of subpavings of $\mathcal{Q}$ such that $\mathcal{K}^- \subset \mathcal{K}^+$.



Example : Consider the paving of $\mathbb{R}$ defined by

$$\mathcal{Q} = \{]-\infty, 0], [0, 1], [1, 2], [2, 3], [3, \infty[\} \,.$$

Three possible interval subpavings of $\mathcal{Q}$ are

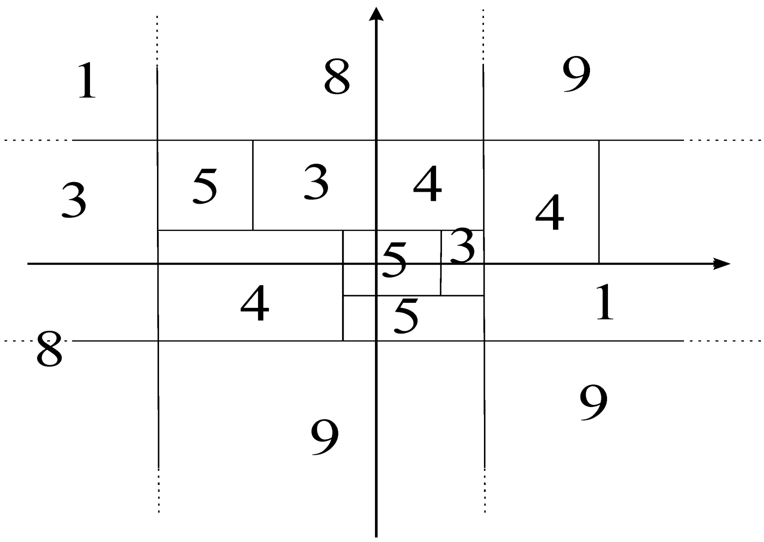
$$\begin{aligned}
[\mathcal{K}_1^-; \mathcal{K}_1^+] &= [\{]-\infty, 0], [1, 2]\}\,; \{]-\infty, 0], [1, 2], [2, 3]\}] \,, \\
[\mathcal{K}_2^-; \mathcal{K}_2^+] &= [\{[1, 2]\}\,; \{[1, 2]\}] \,, \\
[\mathcal{K}_3^-; \mathcal{K}_3^+] &= [\{\}\,; \{\}] \,.
\end{aligned}$$

Definition:

$$\mathbb{S} \in [\mathcal{K}^-, \mathcal{K}^+] \Leftrightarrow \left\{\mathcal{K}^-\right\} \subset \mathbb{S} \subset \left\{\mathcal{K}^+\right\}.$$

5 Interval staircase functions

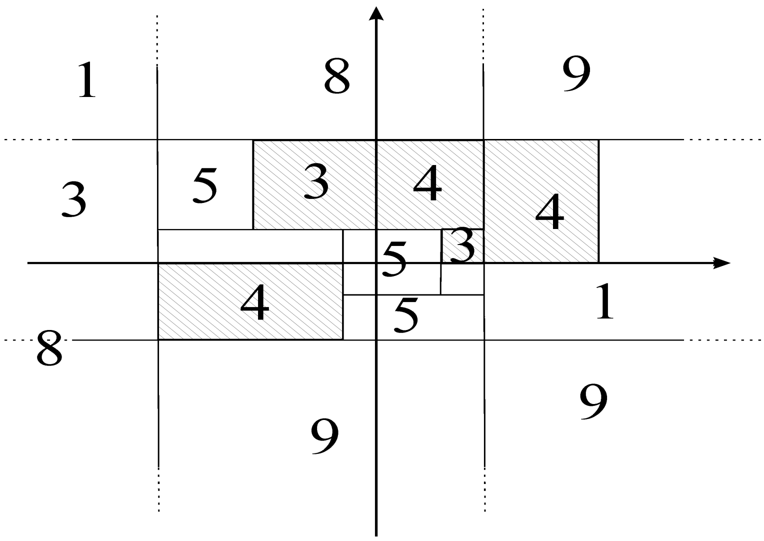
A *staircase function* $\hat{f}$ associated with a paving $\mathcal{Q}$ is a function from $\mathcal{Q}$ to $\bar{\mathbb{R}}$.



If $[s] = [s^-, s^+] \in \mathbb{IR}$, the *reciprocal image* of $[s]$ by $\hat{f}$ is the subpaving of $\mathcal{Q}$ defined by

$$\hat{f}^{-1}([s]) \triangleq \left\{[\mathbf{p}] \in \mathcal{Q} \mid \hat{f}([\mathbf{p}]) \in [s^-, s^+]\right\}.$$

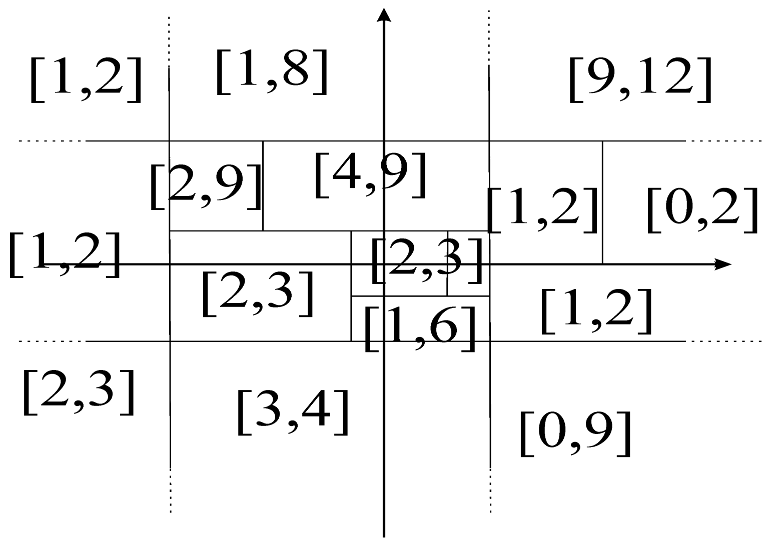
For instance, $\hat{f}^{-1}([2, 4])$ is represented as



The set of all staircase functions $(\hat{\mathcal{F}}, \leq)$ is a complete lattice. Interval staircase functions can thus be defined

An interval staircase function $[\hat{f}] = [\hat{f}^-, \hat{f}^+]$ can be represented a pair of two staircase functions such that

$$\forall [\mathbf{p}] \in \mathcal{Q}\ ,\ \hat{f}^-([\mathbf{p}]) \leq \hat{f}^+([\mathbf{p}]).$$



A function f from $\mathbb{R}^n \longrightarrow \mathbb{R}$ is said to belong to the interval staircase function $[\hat{f}]$ if

$$\forall [\mathbf{p}] \in \mathcal{Q}, \forall \mathbf{p} \in [\mathbf{p}], f(\mathbf{p}) \in [\hat{f}^-([\mathbf{p}]), \hat{f}^+([\mathbf{p}])].$$

An interval staircase function for a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ can easily be obtained by using interval techniques.

The *reciprocal image* of the interval $[s^-, s^+] \in \mathbb{IR}$ by the interval staircase function $[\hat{f}] = [\hat{f}^-, \hat{f}^+]$ is the interval subpaving of $\mathcal{Q}$ defined by

$$[\hat{f}]^{-1}([s^-, s^+]) \triangleq \Big[\Big\{[\mathbf{p}] \in \mathcal{Q} \mid [\hat{f}]([\mathbf{p}]) \subset [s^-, s^+]\Big\}\,. \\ \Big\{[\mathbf{p}] \in \mathcal{Q} \mid [\hat{f}]([\mathbf{p}]) \cap [s^-, s^+] \neq \emptyset\Big\}\Big]$$

Theorem

If f belongs to $[\hat{f}]$, then for all $[s^-, s^+] \in \mathbb{IR}$,

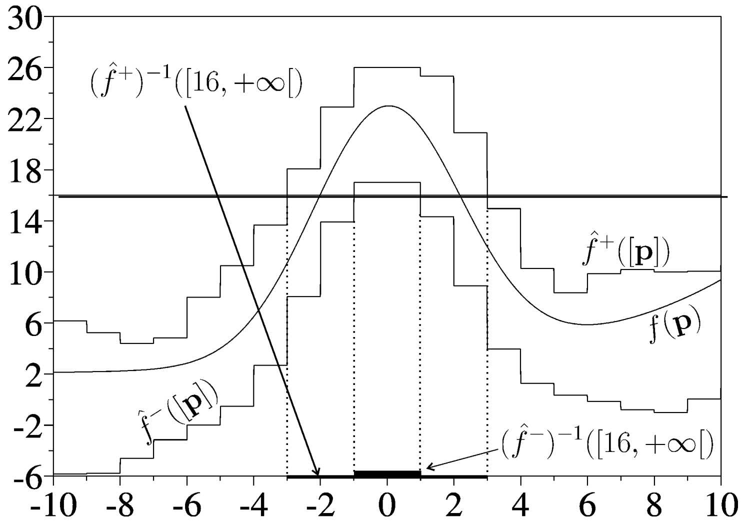
$$f^{-1}([s^-, s^+]) \in [\hat{f}]^{-1}([s^-, s^+]).$$

Example: If $[s^-, s^+] = [16, \infty[$ and $\mathcal{Q} = \{[i, i+1],\ i \in \mathbb{N}\}$, then

$$\{[\mathbf{p}] \in \mathcal{Q} \mid [\hat{f}]([\mathbf{p}]) \subset [s^-, s^+]\ \}$$
$$= \{[-1, 0], [0, 1]\} \equiv [-1, 1],$$
$$\{[\mathbf{p}] \in \mathcal{Q} \mid [\hat{f}]([\mathbf{p}]) \cap [s^-, s^+] \neq \emptyset\}$$
$$= \{[-3, -2], [-2, -1], [-1, 0], [0, 1], [1, 2], [2, 3]\}$$
$$\equiv [-3, 3].$$

We have

$$[-1, 1] \subset f^{-1}([16, \infty[) \subset [-3, 3].$$



If $\left[\mathcal{K}^{-},\mathcal{K}^{+}\right]$ is an interval subpaving of $\mathcal{Q}$ and if $[\hat{f}]$ is a positive interval staircase function, the *integral* of $[\hat{f}]$ over $\left[\mathcal{K}^{-},\mathcal{K}^{+}\right]$ is

$$\int_{[\mathcal{K}^{-},\mathcal{K}^{+}]}[\hat{f}](\mathbf{p})d\mathbf{p} \triangleq \left[\sum_{[\mathbf{p}]\in\mathcal{K}^{-}} \hat{f}^{-}([\mathbf{p}]).\text{volume}([\mathbf{p}])\;.\right.$$

$$\left.\sum_{[\mathbf{p}]\in\mathcal{K}^{+}} \hat{f}^{+}([\mathbf{p}]).\text{volume}([\mathbf{p}])\right]$$

Theorem

If $f \in [\hat{f}]$ and if $\mathbb{S} \in \left[\mathcal{K}^-, \mathcal{K}^+\right]$, then

$$\int_{\mathbb{S}} f(\mathbf{p})d\mathbf{p} \in \int_{[\mathcal{K}^-,\mathcal{K}^+]} [\hat{f}](\mathbf{p})d\mathbf{p}.$$

6 Algorithm

Equation in s_α to be solved

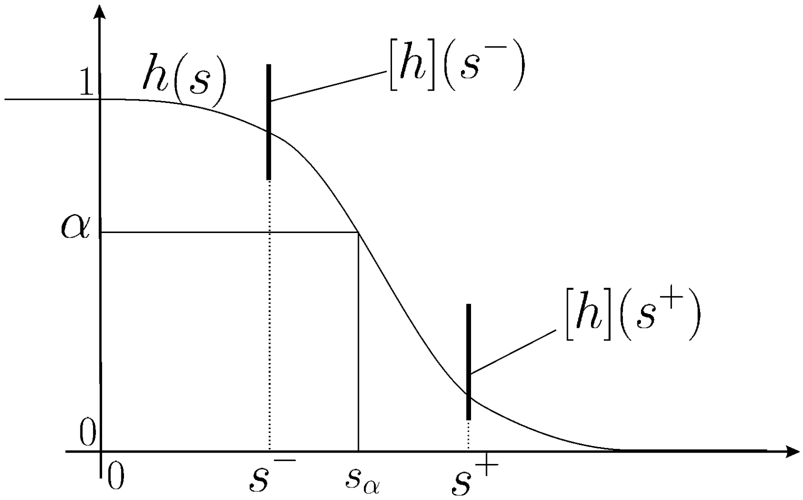
$$\alpha = h(s_\alpha) \triangleq \frac{\int_{f^{-1}([s_\alpha,\infty[)} f(\mathbf{p})d\mathbf{p}}{\int_{\mathbb{R}^n} f(\mathbf{p})d\mathbf{p}}$$

The function $h(s)$ is decreasing. Moreover,

$$h(s) \in [h](s) \triangleq \frac{\int_{[\hat{f}]^{-1}([s,\infty[)} [f](\mathbf{p})d\mathbf{p}}{\int_{\mathbb{R}^n} [\hat{f}](\mathbf{p})d\mathbf{p}}.$$

Thus

$$\begin{array}{lll} (a) & \alpha < lb([h](s^-)) & \Rightarrow s^- < s_\alpha \\ (b) & \alpha > ub([h](s^+)) & \Rightarrow s^+ > s_\alpha \end{array}$$



1. Take a paving $\mathcal{Q}$ of $\mathbb{R}^n$; $s^- := +\infty$; $s^+ := 0$;

2. Compute an interval staircase function $[\hat{f}]$ enclosing f;

3. Decrease s^- until $\alpha < lb([h](s^-))$

4. Increase s^+ until $\alpha > ub([h](s^+))$;

5. $\left[\mathcal{K}_\alpha^-, \mathcal{K}_\alpha^+\right] := ([\hat{f}] - [s^-, s^+])^{-1}([0, \infty[)$.

Theorem : After completion of this algorithm, we have

$$\mathbb{S}_\alpha \in \left[\mathcal{K}_\alpha^-, \mathcal{K}_\alpha^+\right] \text{ and } s_\alpha \in [s^-, s^+].$$

7 Application to Bayesian estimation

Model: $\mathbf{y} = \boldsymbol{\phi}(\mathbf{p}) + \mathbf{n}$, where $\mathbf{n}$ is the noise vector, $\mathbf{y}$ is the data vector and $\boldsymbol{\phi}$ is the model function.

Bayes rule:

$$\pi_{\text{post}}(\mathbf{p}) = \frac{\pi_n(\mathbf{y} - \boldsymbol{\phi}(\mathbf{p})).\pi_{\text{prior}}(\mathbf{p})}{\int_{\mathbf{p}\in\mathbb{R}^n} \pi_n(\mathbf{y} - \boldsymbol{\phi}(\mathbf{p})).\pi_{\text{prior}}(\mathbf{p})d\mathbf{p}}.$$

Unnormalized pdf:

$$f(\mathbf{p}) = \pi_n(\mathbf{y} - \phi(\mathbf{p})).\pi_{\text{prior}}(\mathbf{p}).$$

Model:

$$y(t) = p_1 \sin(p_2 t) + n(t)$$

where $n(t)$ is a white normal random signal with:

$$\pi_n(n) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{n^2}{2\sigma^2}\right),$$

where the standard deviation is $\sigma = \frac{1}{2}$.

Sampling times and data:

$$\begin{aligned}\mathbf{t} &= (1, 2, 3),\\ \mathbf{y} &= (0.8,\ 1.0,\ 0.2)^{\mathsf{T}}.\end{aligned}$$

Therefore

$$\underbrace{\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}}_{\mathbf{y}} = \underbrace{\begin{pmatrix} p_1 \sin(p_2) \\ p_1 \sin(2p_2) \\ p_1 \sin(3p_2) \end{pmatrix}}_{\phi(\mathbf{p})} + \underbrace{\begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}}_{\mathbf{n}}$$

Since $n(t)$ is white,

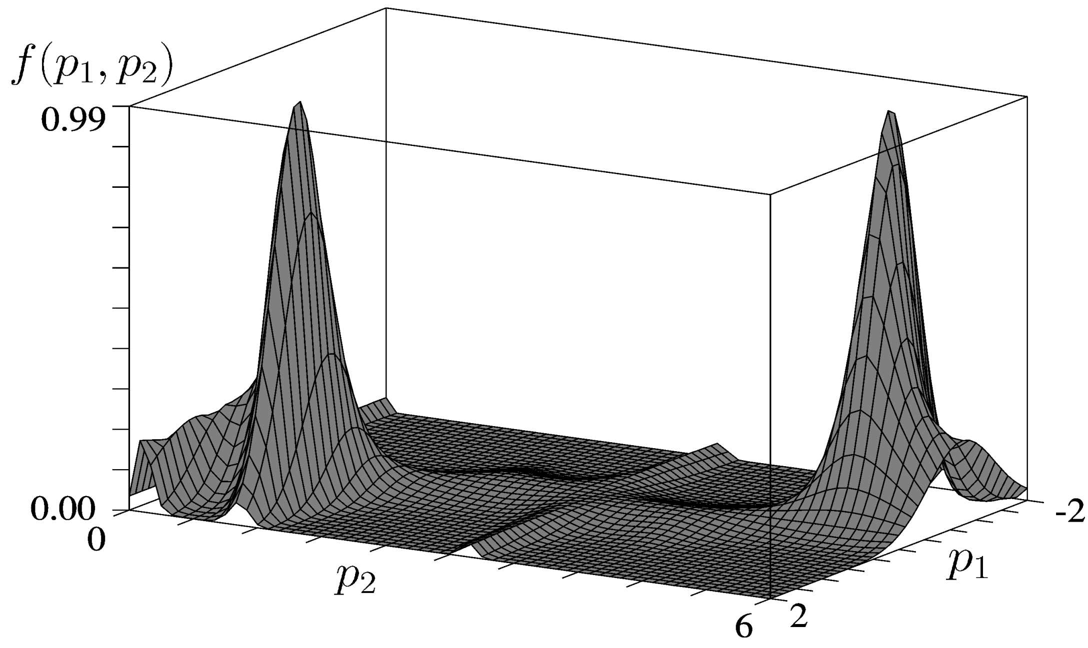
$$\begin{aligned}\pi_n(\mathbf{n}) &= \pi_n(n_1).\pi_n(n_2).\pi_n(n_3) \\ &= \frac{1}{\left(\sqrt{2\pi}\right)^3}\exp(-2n_1^2)\exp(-2n_2^2)\exp(-2n_3^2).\end{aligned}$$

If

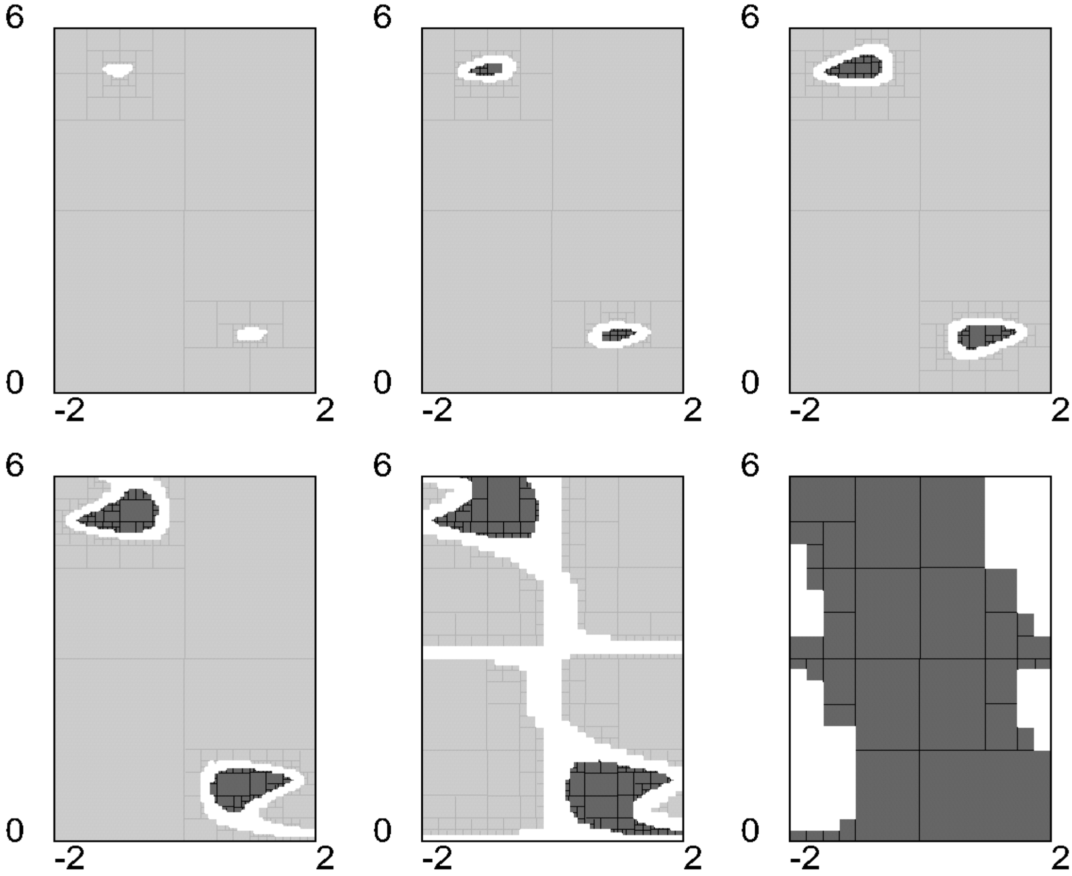
$$\pi_{\text{prior}}(\mathbf{p}) = \frac{\text{door}_{[-2,2]}(p_1).\text{door}_{[0,6]}(p_2)}{24},$$

The posterior unnormalized pdf for $\mathbf{p}$:

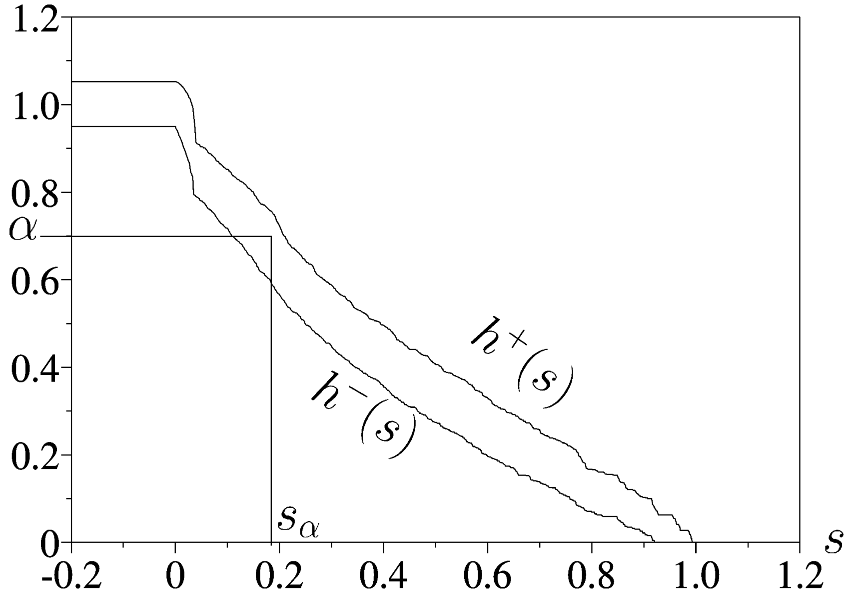
$$f(\mathbf{p}) = \left(\prod_{k=1}^{3} \exp(-2\,(y_k - p_1 \sin(kp_2))^2)\right) . \text{door}_{[-2,2]}(p_1) . \text{door}_{[0,6]}(p_2).$$



$f(p_1, p_2)$
0.99
0.00
0
p_2
6
-2
p_1
2



$$\left[\mathcal{K}_\alpha^-, \mathcal{K}_\alpha^+\right] \text{ for } \alpha \in \{0, 0.2, 0.4, 0.6, 0.8, 1\};$$



1.2
1.0
0.8
α
0.6
0.4
0.2
0
-0.2
0
0.2
0.4
0.6
0.8
1.0
1.2
s
s_α
$h^+(s)$
$h^-(s)$