

Représentation et synthèse de l'information incertaine

S. Destercke (1ère année)

IRSN/DPAM/SEMIC/LIMSI

Journée des thèses IRSN 2006

Tuteur IRSN : Eric Chojnacki (DPAM/SEMIC/LIMSI)
Directeur de thèse : Didier Dubois (IRIT, université Toulouse 3)
Thèse financée par l'IRSN



Contexte

Analyse d'incertitude répond à un besoin :

- ▶ de rigueur scientifique
- ▶ de crédibilité sociale

Nombreux domaines d'applications

- ▶ Radioprotection (GRNC), Sûreté nucléaire (BEMUSE, UMS)
- ▶ Environnement (Thèse cofinancée BRGM/IRIT/IRSN)

Retour d'Expérience (REX)

- ▶ La façon dont sont quantifiées/représentées les incertitudes sources a un impact majeur sur l'analyse d'incertitude
- ▶ Ces incertitudes sources proviennent souvent d'opinions d'experts

Objectifs

Globaux

Modéliser le plus précisément possible l'ensemble des incertitudes sur les paramètres d'un modèle pour garantir la pertinence du résultat final.

1ère phase

- ▶ Etude des théories mathématiques permettant de représenter l'incertitude
- ▶ Traitement de l'information experte

2 types d'incertitudes

Incertitude = **variabilité** + **imprécision**

- ▶ **variabilité** : caractéristique naturelle d'un phénomène (e.g. temps de défaillance, taille, ...) → incertitude aléatoire
- ▶ **imprécision** : due à un manque d'information, de données (e.g. pression ou température en un point, ...) → incertitude épistémique

Traitement des opinions expertes

Opinions d'experts généralement traitées en 3 étapes

1. Représentation des opinions
2. Pondération des experts
3. Synthèse des opinions

Représentation usuelle

Usuellement

- ▶ Experts communiquent leurs connaissances (imprécises)
- ▶ Sélection d'une distribution de probabilité suivant le principe de maximum d'entropie

Inconvénients

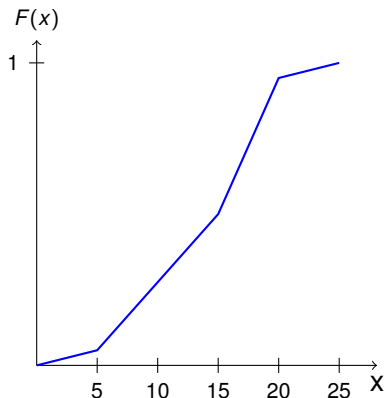
Utiliser le max. d'entropie revient à rajouter de l'information.

Comment y remédier ?

Considérer l'ensemble des distributions en utilisant, par exemple :

- ▶ Théorie des fonctions de croyance
- ▶ Théorie des possibilités

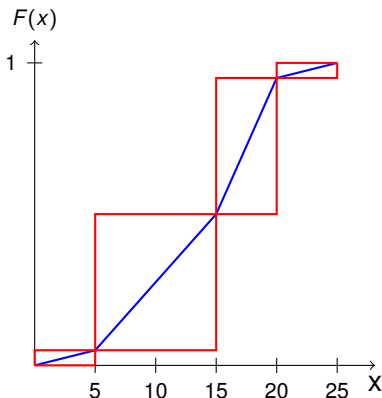
Exemple 1 : probabilité partiellement connue



Temps de lessivage d'un sol

- ▶ Expert fournit les valeurs suivantes pour les quantiles à 5%, 50%, 95% : 5, 15, 20
- ▶ Distribution par max. d'entropie
- ▶

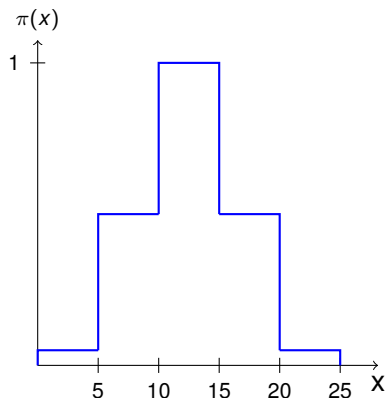
Exemple 1 : probabilité partiellement connue



Temps de lessivage d'un sol

- Expert fournit les valeurs suivantes pour les quantiles à 5%, 50%, 95% : 5, 15, 20
- Distribution par max. d'entropie
- Ensemble des distributions possibles

Exemple 2 : Valeur imprécisément connue



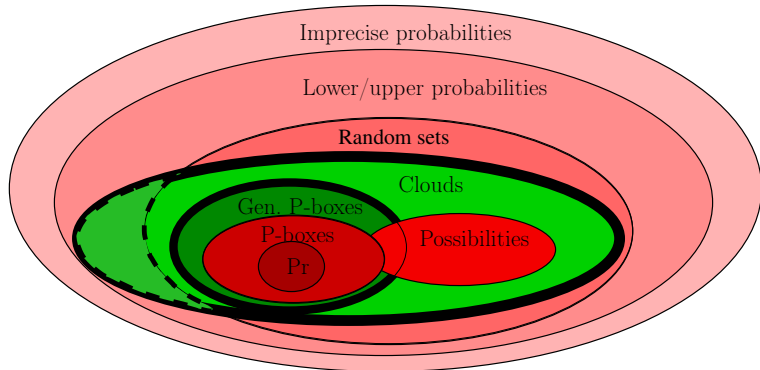
Intervalle de confiance

Expert donne intervalles de confiance suivants

- ▶ 50% $\in [10, 15]$
- ▶ 95% $\in [5, 20]$
- ▶ 100% $\in [0, 25]$

→ Distribution de possibilité

Bibliographie et développements théoriques issus du travail de thèse : synthèse, extension, et établissement de relations entre les théories de l'incertain



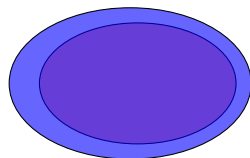
Pondération des experts

Principes de la pondération

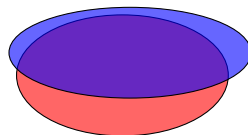
- ▶ Différencier les experts grâce à une évaluation objective
- ▶ Evaluation = **Calibration** + **Informativité**
 - ▶ **Calibration** : cohérence avec réalité connue (nécessité de variables témoins)
 - ▶ **Informativité** : précision de l'opinion.

Notion de calibration

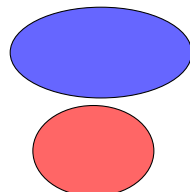
Connaissance expert
Réalité connue



$Cal = 1$
Bien calibré



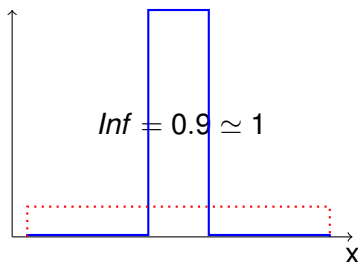
$\tilde{Cal} = 0.5$
Moyennement calibré



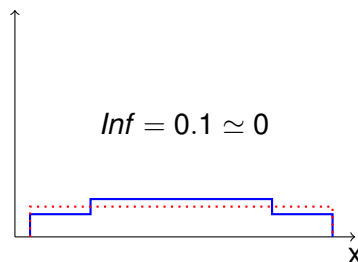
$Cal = 0$
Mal calibré

Notion d'informativité

Connaissance expert
Distribution non-informative



Expert Informatif



Expert non-informatif

Pondération → Bien étudiée pour les probabilités

Extension aux autres théories de l'incertain dans le cadre de la thèse

Distributions non-informatives : p^i, π^i, m^i ; expertes : p^e, π^e, m^e ; témoins : p^t, π^t, m^t

	Existant	Nos propositions (exemples)	
	Prob.	Poss.	Croy.
Calibration	$\sum_j p_j^e \ln\left(\frac{p_j^e}{p_j^t}\right)$	$\frac{ \min(\pi^e, \pi^t) }{ \pi^t }$	$\frac{IP(\mathcal{P}_{m^t} \cap \mathcal{P}_{m^e})}{IP(\mathcal{P}_{m^t})}$
Informativité	$\sum_j p_j^e \ln\left(\frac{p_j^e}{p_j^i}\right)$	$1 - \frac{ \pi^e }{ \pi^i }$	$1 - \frac{IP(\mathcal{P}_{m^e})}{IP(\mathcal{P}_{m^i})}$

$\mathcal{P}_{m^t}$: famille de probabilités modélisée par la fonction de croyance m^t

IP : mesure d'imprécision dérivée de l'inverse de Möbius

$$(m(E) = \sum_{A \subseteq E} -1^{|E-A|} \mu(A))$$

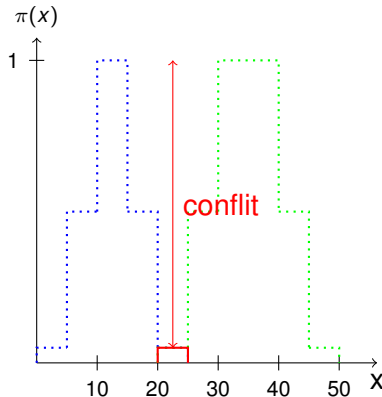
Méthodes de synthèse

But : prendre l'ensemble des opinions et les synthétiser de façon cohérente pour n'obtenir qu'une seule distribution.

2 approches principales

- ▶ Conjonction : Intersection des opinions (→ On conserve les valeurs communes à toutes les opinions) .
 - ▶ Résultat très précis / peu fiable en cas de désaccord
- ▶ Disjonction : Union des opinions (→ On considère l'ensemble des valeurs des opinions)
 - ▶ Résultat très peu précis / très fiable

Synthèse : conjonction



Intervalle de confiance

Intervalle à 50%, 95%, 100% :

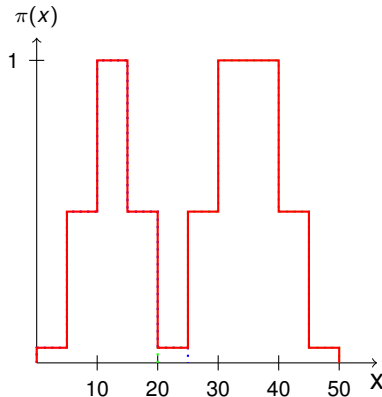
- ▶ Expert 1 :
[10, 15], [5, 20], [0, 25]
- ▶ Expert 2 :
[30, 40], [25, 45], [20, 50]

→ Conjonction (Intersection, connecteur logique ET)

$$\pi_{\cap} = \min(\pi_1, \pi_2)$$

→ Désaccord important!

Synthèse : disjonction



Intervalles de confiance

Intervalles à 50%, 95%, 100% :

- ▶ Expert 1 :
[10, 15], [5, 20], [0, 25]
- ▶ Expert 2 :
[30, 40], [25, 45], [20, 50]

→ Disjonction (Union,
connecteur logique OU)

$$\pi_U = \max(\pi_1, \pi_2)$$

→ Approche conservative

Autres méthodes de synthèse

- ▶ Moyennes pondérées
- ▶ Méthodes asymétriques
- ▶ Méthodes hiérarchiques,...

Travail en cours : construire méthode tirant profit de

- ▶ pondération
- ▶ conjonction
- ▶ disjonction

en utilisant la notion de **sous-ens. maximaux cohérents.**

Représentation et synthèse de l'information incertaine

Conclusions

Par rapport aux méthodes probabilistes usuelles, les probabilités imprécises permettent plus de liberté :

1. dans la modélisation de la connaissance
 2. dans la pondération des experts (ou de sources d'information en général)
 3. dans la manière de synthétiser les connaissances
- Objectif : obtenir des résultats plus fiables / plus pertinents

Représentation et synthèse de l'information incertaine

Perspectives

- ▶ Utilisation des méthodes sur un Benchmark OCDE (BEMUSE) concernant des logiciels de sûreté en Th. accidentelle. → comparaison des résultats fournis par les instituts
- ▶ Proposer de nouvelles méthodes dans le cas probabiliste imprécis pour les problèmes suivants
 - ▶ Représentation/synthèse des résultats
 - ▶ Analyse de sensibilité
 - ▶ Prise en compte de dépendances entre les variables / les sources
 - ▶ Problème inverse

Représentation et synthèse de l'information incertaine

publications

Journaux

- ▶ "Some links between practical representations of imprecise probabilities" En préparation, à soumettre.

Conférences avec comité de lecture

- ▶ "A unified view of some representations of imprecise probabilities" Third International Conference on Soft Methods in Probabilities and Statistics (SMPS 2006), Bristol, UK
- ▶ "Aggregation of expert opinions and uncertainty theories" Rencontres Francophones sur la Logique Floue et ses Applications (LFA 2006), Toulouse
- ▶ "Notes on clouds" 28th Seminar on fuzzy sets, 2007, Linz, Austria (prevision)

MERCI!

questions ?